

Mathematik 2

für

Bau- und Umweltingenieurwesen

Unterlagen zur Vorlesung an der TU Wien
im Sommersemester 2021 gehalten von
Reinhard Winkler
und zu den Übungen mit zahlreichen Aufgaben von
Gabriel Maresch

26. Februar 2021

Sehr geehrte Hörerinnen und Hörer!

Das vorliegende Skriptum zur *Mathematik 2 für Bauingenieurwesen* unterscheidet sich von der Ausgabe vom Februar 2019 durch einige Verbesserungen. Sinnstörende Druckfehler wurden korrigiert. Darüber hinaus reichende Verbesserungen wurden in den ersten beiden Kapiteln und in Teilen von Kapitel 3 vorgenommen. (Die ursprünglich für diese Ausgabe beabsichtigte Revision auch der übrigen Teile wurde durch die Corona-Pandemie und den durch sie bedingten unermesslich explodierenden organisatorischen Aufwand für alles und jedes, vor allem für Prüfungen, vereitelt.) Jedenfalls möchte ich die aktuelle Fassung auch jenen ans Herz legen, die schon eine frühere Ausgabe besitzen.

Aus dem erwähnten Grund muss ich um Nachsicht bitten für diverse immer noch vorhandene Unvollkommenheiten. Trotz der vielen Zeit, die ich schon an die Revision und Anpassung an meine Vorlesung investiert habe, konnte ich immer noch nicht allen Abschnitten jene Gestalt geben, die mir eigentlich vorschwebt. Beträchtliche Teile waren entstanden, bevor ich erstmals die Vorlesung hielt, und wurden von anderen Personen und unter anderen Rahmenbedingungen verfasst. Auch wenn ich die Autorenschaft einzelner Abschnitte gar nicht mehr eindeutig zuzuordnen weiß, gilt mein Dank all jenen, deren Arbeit ich hiermit weiter verwenden darf. Insbesondere stammen die zahlreichen (sehr aufwendig zu erstellenden) Grafiken durchwegs nicht von mir, sondern aus älteren Texten. Daraus erklären sich manche Redundanzen, uneinheitliche Notation und Terminologie, Unvollständigkeiten im Index und bei Querverweisen, spürbare Brüche zwischen Abschnitten unterschiedlicher Herkunft etc.

Der Stoff ist sehr umfangreich, und der zeitliche Rahmen wird es nicht erlauben, alle Inhalte des Skriptums auch in der Vorlesung durchzunehmen. (In diesem Punkt unterscheidet sich Mathematik 2 deutlich von Mathematik 1.) Was genau in der Vorlesung gemacht werden kann, ergibt sich teilweise erst aus dem Fortgang der Lehrveranstaltung während des Semesters. Nur Themen, die auch in der Vorlesung angesprochen werden, sind Prüfungsstoff. Folglich empfiehlt es sich, die Vorlesung möglichst regelmäßig zu verfolgen. Trotz der Distanzlehre, mit der das Semester beginnen (und voraussichtlich auch enden) wird, empfiehlt es sich, die Vorlesungen möglichst regelmäßig zu verfolgen oder wenigstens nachzuhören. Ich werde mich bemühen, Ersatz zu bieten für die Interaktion, die normalerweise im Hörsaal möglich ist, um die wir in Zeiten der Pandemie vom Schicksal aber weitgehend betrogen werden.

Wie schon in Mathematik 1 haben Herr Maresch und ich als Orientierungshilfe für den Umgang mit den Übungsaufgaben eine Zuordnung zu drei Kategorien T (Test), P (Prüfung) und E (Ergänzung) vorgenommen. Für die Vorbereitung auf die beiden Übungstests empfehlen wir vor allem die Beschäftigung mit den T-Aufgaben (T* bezeichnet Aufgaben, die zur Testvorbereitung empfohlen werden, die aber etwas länger oder aufwendiger sind, als es bei tatsächlichen Testaufgaben zu erwarten ist), für die Prüfung darüber hinaus mit den P-Aufgaben. Von den E-Aufgaben sind einige etwas anspruchsvoller. Lassen Sie sich nicht entmutigen, wenn Sie manche davon nicht lösen können. Wenn Sie sich aber bei jeder E-Aufgabe wenigstens klar machen, worin die Aufgabenstellung besteht, so wird das Ihr Verständnis ganz wesentlich vertiefen und daher

eine sehr sinnvolle Vorbereitung auf die Prüfung sein.

Für spezifische Anwendungen der Mathematik im Bauingenieurwesen verweise ich auf die Sammlung, die Spezialisten aus Ihrer Fakultät zusammengestellt haben und die ich Ihnen per TISS elektronisch zukommen lasse. Prüfungsstoff für Mathematik 2 sind diese Inhalte nicht, sie sollen vor allem Ihrer Motivation dienen. Es gilt, Mathematik nicht nur als ein wesentliches Instrument für Ihre eigene Wissenschaft zu begreifen, sondern auch als eine Sprache, ohne die viele für die Ingenieurwissenschaften grundlegende Sachverhalte vor allem aus der Physik nicht einmal formuliert werden könnten.

Einige Empfehlungen zur Prüfungsvorbereitung: Der erste Schritt in Ihrer Vorbereitung könnte darin bestehen, sich einen groben Überblick über den gesamten Stoff zu machen. Es ist sicher kein Luxus, sich die Themen der fünf Kapitel (oberste Gliederungsebene) und ihrer 6+4+3+4+3 Abschnitte (zweite Gliederungsebene) einzuprägen. Dabei könnte es hilfreich sein, zunächst die Einleitungen zu den Kapiteln und zu den Abschnitten zu lesen. Bei der genaueren Erarbeitung des Stoffes empfehle ich, von den Kurzfassungen am Beginn jedes Unterabschnitts (dritte Gliederungsebene) auszugehen. Diese bestehen in der Regel nur aus wenigen Zeilen, in denen die wichtigsten Schlagworte und Grundideen erwähnt werden. Fragen Sie sich zu jedem dieser Schlagworte nicht nur „Wie rechne ich damit?“, sondern vor allem auch: „Was ist das?“ Mathematiker stellen mit Vorliebe zuerst die zweite Frage, und das gilt insbesondere bei Prüfungen. Zwei Beispiele: Zum Schlagwort „Ableitung“ sollte Ihnen jedenfalls die Assoziation „lokale lineare Approximation“ kommen, noch bevor Sie darüber nachdenken, wie Sie partielle Ableitungen ausrechnen und zur Funktionalmatrix als Darstellung der Ableitung zusammensetzen. Oder: Zum Schlagwort „Wahrscheinlichkeit“ sollten Sie den Begriff des Wahrscheinlichkeitsraumes mit seinem wichtigsten Drumherum parat haben und nicht hilflos lediglich nach ein, zwei trivialen Beispielen suchen. In Verbindung mit den allgemeinen Begriffen jedoch sind für das Verständnis von Mathematik auch charakteristische Beispiele nicht nur nützlich, sondern eine ganz wesentliche Hilfe. Versuchen Sie daher, abstrakte Begriffe immer auch in Verbindung mit typischen Beispielen zu sehen.

Vielleicht fragen Sie sich, wie intensiv Sie sich mit den im Skriptum enthaltenen Beweisen auseinandersetzen sollen. Auch dazu ein paar Worte: Beweise sind für Fachmathematiker das Herz ihrer Wissenschaft. In Anwendungsdisziplinen wie dem Bauingenieurwesen kommt es aber nicht so sehr darauf an, dass man komplizierte Beweise mathematischer Sätze aktiv wiedergeben kann. Deshalb prüfe ich auch keine langen Beweise. Trotzdem habe ich zahlreiche Beweise im Skriptum aufgenommen, weil ich glaube, dass man das Bemerkenswerte von mathematischen Ergebnissen viel besser versteht, wenn man die Beweise wenigstens in ihren Grundideen studiert. Meine Empfehlung für die Prüfungsvorbereitung lautet daher: Versuchen Sie, auch möglichst viele Beweise lesend zu verstehen; lassen Sie sich aber nicht verunsichern, nur weil Ihnen das wahrscheinlich nicht bei allen zur Gänze gelingt. Nur wenn Sie viele Beweise überhaupt nicht verstehen, ist das ein Alarmsignal. Dann ist nämlich zu befürchten, dass Ihr Verständnis des Stoffes noch nicht ausreicht, um die Prüfung zu bestehen. Denn die Prüfung versuche ich gerade so zu gestalten, dass es auf ein sinnvolles (nicht notwendig detailreiches) Verständnis der

wichtigsten mathematischen Konzepte, Resultate und Methoden ankommt.

Wenn Sie unsicher sind, ob Sie sich bei Ihrer Prüfungsvorbereitung auf einem guten Weg befinden, steht es Ihnen jederzeit frei, sich an mich zu wenden. So lange die Corona-Pandemie persönliche Kontakte weitgehend unmöglich macht, kann ich keine Sprechstunden in üblicher Form, d.h. zu fixen wöchentlichen Terminen in meinem Dienstzimmer anbieten. Ich möchte aber Zoom-Konferenzen in ähnlicher Weise veranstalten wie schon zur Mathematik 1 im Wintersemester. Sollte sich die Situation im Laufe des Semester hinsichtlich persönlicher Begegnungen doch noch zum Positiven wenden – umso besser.

Ich hoffe, dass diese Präambel für Sie von Nutzen ist, insbesondere bei der Prüfungsvorbereitung. Für die Prüfung selbst wünsche ich Ihnen viel Erfolg!

Reinhard Winkler, im Februar 2021

Inhaltsverzeichnis

1	Lineare Algebra	9
1.1	Vektorräume und ihre Geometrie	10
1.1.1	Vektorraum als abstrakte Struktur	10
1.1.2	Unterraum, Linearkombination und Erzeugnis	12
1.1.3	Lineare Unabhängigkeit	15
1.1.4	Basen	18
1.1.5	Dimension	21
1.2	Lineare Abbildungen	25
1.2.1	Der Begriff der linearen Abbildung	25
1.2.2	Linearität und Unterräume	27
1.2.3	Lineare Abbildungen und Dimension	29
1.2.4	Die Festlegung einer linearen Abbildung auf einer Basis	30
1.2.5	Vektorräume linearer Abbildungen	32
1.2.6	Beispiele linearer Abbildungen und ihre geometrische Deutung	34
1.3	Matrizen	37
1.3.1	Matrizen als $n \times m$ -Vektoren (-Zahlenschemata)	37
1.3.2	Die Matrix einer linearen Abbildung bezüglich gegebener Basen	39
1.3.3	Das Matrizenprodukt	42
1.3.4	Matrizenkalkül	44
1.3.5	Transformation bei Basiswechsel	48
1.4	Multilinearität	49
1.4.1	Multilineare Abbildungen	50
1.4.2	Interessante Beispiele multilinearer Abbildungen	54
1.4.3	Skalarprodukt	56
1.4.4	Orthogonale Matrizen	61
1.4.5	Determinanten	65
1.5	Lineare Gleichungssysteme	72
1.5.1	Problemstellung	73
1.5.2	Ein einfaches Beispiel	74
1.5.3	Berechnung der inversen Matrix	76
1.5.4	Berechnung der Determinante	79
1.5.5	Lösung linearer Gleichungssysteme, der allgemeine Fall	82
1.5.6	Ergänzende Bemerkungen	88
1.6	Eigenwerttheorie	91
1.6.1	Definitionen und charakteristische Gleichung	91
1.6.2	Berechnungsmethoden	93
1.6.3	Ähnlichkeit und Diagonalform	96

2	Differentialrechnung in mehreren Variablen	101
2.1	Nichtlineare Geometrie im \mathbb{R}^n	101
2.1.1	Motivation	102
2.1.2	Höherdimensionale Folgen und stetige Funktionen	102
2.1.3	Reellwertige Funktionen – Funktionsgebirge	106
2.1.4	Kurven und Flächen	107
2.1.5	Koordinatentransformationen	110
2.1.6	Vektorfelder	111
2.2	Die Ableitung im Höherdimensionalen	112
2.2.1	Richtungs- und partielle Ableitungen	113
2.2.2	Definition der Ableitung als lineare Approximation	115
2.2.3	Funktionalmatrix und -determinante	117
2.2.4	Entfaltung der Differentialrechnung	122
2.2.5	Geometrische Interpretation der Ableitung	127
2.2.6	Potentialfelder (Gradientenfelder)	131
2.3	Nichtlineare Gleichungssysteme	136
2.3.1	Übersicht	136
2.3.2	Mehrdimensionales Newtonverfahren	137
2.3.3	Lokale Umkehrfunktionen	139
2.3.4	Implizite Funktionen	141
2.4	Extremwertbestimmung in mehreren Variablen	144
2.4.1	Problemstellung und Lösungsstrategie im Überblick	145
2.4.2	Die Rolle der zweiten Ableitungen	147
2.4.3	Taylor, quadratische Formen und Hauptachsentransformation	149
2.4.4	Extrema unter Nebenbedingungen	156
2.4.5	Strategie zur Lösung von Extremwertaufgaben	158
2.4.6	Anwendungen: Ausgleichrechnung und Regressionsgerade	162
3	Differentialgleichungen (DGLen)	165
3.1	Allgemeine Theorie	165
3.1.1	DGLen als Funktionalgleichungen	166
3.1.2	Bemerkungen zur Notation	168
3.1.3	Gewöhnliche und partielle Differentialgleichungen	169
3.1.4	Ordnung versus Dimension – gekoppelte Systeme	172
3.1.5	Existenz und Eindeutigkeit von Lösungen expliziter DGLen	173
3.1.6	Qualitative Untersuchungen autonomer DGLen	179
3.2	Lineare Differentialgleichungen	181
3.2.1	Allgemeine Vorbemerkungen	182
3.2.2	Speziellere und allgemeinere lineare DGLen – Überblick	185
3.2.3	Einfache homogene lineare DGLen mit konstanten Koeffizienten	186
3.2.4	Inhomogene Gleichungen mit konstanten Koeffizienten	190
3.2.5	Einfache lineare DGLen mit nichtkonstanten Koeffizienten	193
3.2.6	Gekoppelte lineare DGLen erster Ordnung	196
3.2.7	Homogene Systeme höherer Ordnung	206

3.2.8	Inhomogene Systeme	208
3.2.9	Das Anfangswertproblem	213
3.2.10	Lösungsmethode bei konstanten Koeffizienten	214
3.2.11	Das Anfangswertproblem für lineare DG	218
3.3	Lösungsverfahren für ausgewählte nichtlineare DGLen	220
3.3.1	Potenzreihenansatz	220
3.3.2	Implizite Lösungen vermittelt Trennung der Variablen	221
3.3.3	Exakte DGLen	223
3.3.4	Integrierende Faktoren	226
3.3.5	Variablensubstitution	229
3.3.6	Numerische Methoden	230
4	Integralrechnung in mehreren Variablen	233
4.1	Bereichsintegrale	233
4.1.1	Ein gemeinsamer Rahmen	233
4.1.2	Der Satz von Fubini	236
4.1.3	Funktionaldeterminante	239
4.1.4	Die Substitutionsregel	240
4.1.5	Definition und Grundeigenschaften der Doppelintegrale	243
4.1.6	Methoden zur Berechnung und weitere Eigenschaften	246
4.1.7	Gaußsches Fehlerintegral	249
4.1.8	Anwendungen	250
4.1.9	Dreifachintegrale	251
4.2	Kurvenintegrale	252
4.2.1	Kurvenintegrale erster Art (über Funktionen)	252
4.2.2	Kurvenintegrale zweiter Art (über Vektorfelder)	254
4.2.3	Wegunabhängigkeit von Kurvenintegralen in Potentialfeldern	257
4.3	Oberflächenintegrale	259
4.3.1	Vorbemerkungen zur Geometrie von Flächen	259
4.3.2	Oberflächenintegrale erster Art (über Funktionen)	261
4.3.3	Oberflächenintegrale erster Art, Ergänzungen	262
4.3.4	Oberflächenintegrale zweiter Art (über Vektorfelder)	264
4.3.5	Oberflächenintegrale zweiter Art, Ergänzungen	267
4.4	Vektoranalysis und Integralsätze	270
4.4.1	Gradient	271
4.4.2	Rotation	272
4.4.3	Divergenz	274
4.4.4	Gemeinsamkeiten der Integralsätze	277
4.4.5	Integralsatz von Green im \mathbb{R}^2	279
4.4.6	Integralsatz von Stokes	284
4.4.7	Integralsatz von Gauß	286
4.4.8	Ergänzungen zu den Integralsätzen	288

5	Stochastik	291
5.1	Grundbegriffe	291
5.1.1	Wahrscheinlichkeit und Wahrscheinlichkeitsräume	292
5.1.2	W-Räume und Zufallsexperimente, grundlegende Beispiele	294
5.1.3	Diskrete und reelle Verteilungen, Verteilungsfunktionen und Dichten	300
5.1.4	Zufallsgrößen und ihre Verteilung	303
5.1.5	Erwartungswert	305
5.1.6	Varianz	308
5.1.7	Bedingte Wahrscheinlichkeit, Unabhängigkeit und Produkträume .	309
5.1.8	Die Formel von Bayes	312
5.2	Wichtige Ergebnisse der Wahrscheinlichkeitstheorie	313
5.2.1	Rechenregeln für Erwartungswert und Varianz	314
5.2.2	Beispiele diskreter Verteilungen und ihre Erwartungswerte	316
5.2.3	Beispiele stetiger Verteilungen und ihre Erwartungswerte	319
5.2.4	Das Gesetz der großen Zahlen	321
5.2.5	Der Hauptsatz der mathematischen Statistik	325
5.2.6	Der zentrale Grenzwertsatz	326
5.3	Statistik	331
5.3.1	Beschreibende Statistik – Merkmale und Skalen	331
5.3.2	Kennzahlen der beschreibenden Statistik	333
5.3.3	Allgemeine Vorbemerkungen zur Beurteilenden Statistik	336
5.3.4	Punktschätzungen	337
5.3.5	Konfidenzintervalle	342
5.3.6	Hypothesentests	344

1 Lineare Algebra

Unter *Linearer Algebra* versteht man jenes Teilgebiet der Mathematik, in dem vor allem Vektorräume als Grundstruktur fungieren. Aus dem ersten Semester sind bereits die Standardvektorräume \mathbb{R}^n bekannt. Für uns sind sie auch weiterhin sehr wichtige (wenn nicht gar die wichtigsten) Beispiele. Zu einem besseren Verständnis und für spätere Anwendungen ist eine systematische Theorie, auch mit gewissen Abstraktionen, unerlässlich. Sie ist Gegenstand dieses Kapitels.

Die allgemeine Theorie beschäftigt sich mit Vektorräumen über beliebigen Körpern, während wir uns hier vor allem für reelle Vektorräume interessieren, die sich auch hervorragend für die Beschreibung geometrischer Sachverhalte eignen. Die wichtigste Kennzahl eines Vektorraums ist seine Dimension. Diese kann eine natürliche Zahl oder auch unendlich (genauer: eine unendliche Kardinalzahl) sein. Im Fall endlicher Dimension und bei vorgegebener Basis ist die Beschreibung von Vektoren als Linearkombinationen der Basisvektoren besonders einfach und gehorcht sehr vertrauten Mustern. Abschnitt 1.1 behandelt die grundlegenden Konzepte.

So wie in der Mathematik generell Funktionen/Abbildungen überall eine entscheidende Rolle spielen, sind es für Vektorräume speziell *lineare* Abbildungen. Sie sind, wie man sagt, mit der Vektorraumstruktur verträglich und ermöglichen ein wesentlich besseres Verständnis von Vektorräumen. Sie werden in Abschnitt 1.2 eingeführt und begleiten uns durch das gesamte Kapitel.

Die Darstellung linearer Abbildungen bezüglich gegebener Basen führt fast zwangsläufig zum Matrizenkalkül (1.3), dem höherdimensionalen Analogon des elementaren Rechnens mit Zahlen mit Hilfe der üblichen Zifferndarstellung zur Basis 10.

Hängen Funktionen von mehr als einer Variablen (von mehr als einem Vektor) in jeweils linearer Weise ab, so spricht man von multilinearen Abbildungen, im skalarwertigen Fall von Multilinearformen. Die wichtigsten Beispiele sind Skalarprodukte und Determinanten, womit die wichtigsten Themen von Abschnitt (1.4) benannt sind.

Die häufigste Anwendung höherdimensionaler Vektorräume jenseits unmittelbarer Anschaulichkeit (also mit Dimension $n > 3$) sind lineare Gleichungssysteme. Ihre praktische Bedeutung ist offensichtlich. Bei ihrer Untersuchung erweisen sich die Begriffe aus der allgemeinen Theorie (Dimension, lineare Abbildung etc.) als die geeigneten Instrumente. Mit dem Gaußschen Eliminationsverfahren, das im Zentrum von Abschnitt 1.5 steht, gibt es einen Algorithmus, der sowohl für die Lösung linearer Gleichungssysteme, als auch für die Berechnung der Determinanten und der Inversen einer quadratischen Matrix keine Wünsche offen lässt.

Der letzte Abschnitt dieses Kapitels (1.6) beschäftigt sich mit den besonders für Anwendungen besonders wichtigen Eigenvektoren und Eigenwerten sowie mit der damit zusammenhängenden Transformation von quadratischen Matrizen auf Diagonalgestalt.

1.1 Vektorräume und ihre Geometrie

Die Definition eines Vektorraums ergibt sich in natürlicher Weise als Abstraktion vertrauter Rechenregeln (siehe 1.1.1). Sehr naheliegend sind dann auch die Begriffe Unterraum, Linearkombination und Erzeugnis oder lineare Hülle (1.1.2). Letzteres führt in Verbindung mit dem für die weitere Theorie besonders wichtigen Begriff der linearen Unabhängigkeit von Vektoren (1.1.3) zu dem der Basis (1.1.4) und, nachdem man erkannt hat, dass je zwei Basen ein und desselben Vektorraums stets gleich viele Elemente enthalten, zu dem der Dimension (1.1.5). Je zwei Vektorräume mit derselben Dimension haben nämlich dieselbe algebraische Struktur, während sich solche mit unterschiedlicher Dimension strukturell unterscheiden.

1.1.1 Vektorraum als abstrakte Struktur

Inhalt in Kurzfassung: Definition des Begriffs (reeller) *Vektorraum*, die Standardvektorräume \mathbb{R}^n als wichtige und bereits aus Mathematik 1 vertraute Beispiele, einfache Rechenregeln in beliebigen Vektorräumen

Von den bekannten Standardvektorräumen \mathbb{R}^n abstrahieren wir zu folgender allgemeinen Definition.

Definition 1.1.1.1. *Ein (reeller) Vektorraum ist gegeben erstens durch eine Menge V (ihre Elemente \mathbf{x} heißen **Vektoren**, die wir mit fetten Buchstaben schreiben, gelegentlich auch als \vec{x}) auf der eine **Vektoraddition** $+$, ein **Nullvektor** \mathbf{o} und eine (additive) **Inversenbildung** (Subtraktion) $-$ derart definiert sind, dass eine abelsche Gruppe vorliegt. (Zur Erinnerung: Abelsche Gruppe zu sein, bedeutet in diesem Fall, dass für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ das Assoziativgesetz $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$, das Kommutativgesetz $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$, das Nullelementgesetz $\mathbf{x} + \mathbf{o} = \mathbf{x}$ und das Inversengesetz $\mathbf{x} - \mathbf{x} := \mathbf{x} + (-\mathbf{x}) = \mathbf{o}$ gelten.) Zweitens muss eine **Skalarmultiplikation** (Zahl mal Vektor) definiert sein, die jedem $r \in \mathbb{R}$ und jedem $\mathbf{x} \in V$ den (r -fachen) Vektor $r\mathbf{x}$ zuordnet, so dass für alle $r, s \in \mathbb{R}$ und $\mathbf{x}, \mathbf{y} \in V$*

die Distributivgesetze $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$ und $(r + s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$,

das Assoziativgesetz $(rs)\mathbf{x} = r(s\mathbf{x})$ und

das Einselementgesetz $1\mathbf{x} = \mathbf{x}$

gelten. Beziehen sich die Skalare r, s nicht auf reelle Zahlen, sondern auf rationale, komplexe etc. (jedenfalls müssen sie einem Körper – Definition siehe Mathematik 1 – entnommen sein), spricht man statt von einem reellen von einem rationalen, komplexen etc. Vektorraum.

Wir kennen aus Mathematik 1 die **Standardvektorräume** \mathbb{R}^n . Die Vektoren sind dabei die n -tupel $\mathbf{x} = (x_1, \dots, x_n)$ reeller Zahlen. Dabei nennt man x_i die i -te **Komponente** oder **Koordinate** von \mathbf{x} . Oft schreiben wir Vektoren aus \mathbb{R}^n (später auch Vektoren aus einem beliebigen Vektorraum) bezüglich einer Basis lieber als Spaltenvektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Alle Operationen werden komponentenweise definiert:

$$\text{Vektoraddition} \quad \mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

$$\text{Nullvektor} \quad \mathbf{o} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$\text{negativer Vektor} \quad - \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix},$$

$$\text{Skalarmultiplikation} \quad r\mathbf{x} = r \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ rx_2 \\ \vdots \\ rx_n \end{pmatrix}.$$

Dass für die Standardvektorräume die für einen Vektorraum geforderten Gesetze gelten, ist unmittelbar einzusehen und wurde bereits in Mathematik 1 festgehalten. Also sind die bekannten Standardvektorräume tatsächlich auch Vektorräume im abstrakten Sinn. Auch viele vertraute Rechenregeln gelten in beliebigen Vektorräumen, weil sie allein aus der Definition 1.1.1.1 folgen:

Proposition 1.1.1.2. *Sei V ein Vektorraum, $\mathbf{x}, \mathbf{y} \in V$ und $r, s \in \mathbb{R}$. Dann gilt:*

1. $0\mathbf{x} = \mathbf{o}$
2. $r\mathbf{o} = \mathbf{o}$
3. $(-1)\mathbf{x} = -\mathbf{x}$
4. Aus $r\mathbf{x} = \mathbf{o}$ folgt $r = 0$ oder $\mathbf{x} = \mathbf{o}$.

Das nachzurechnen ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 1. (E) *Beweisen Sie Proposition 1.1.1.2, indem Sie ausschließlich die durch Definition 1.1.1.1 gegebenen Gesetze verwenden.*

1 Lineare Algebra

Neben den Standardvektorräumen \mathbb{R}^n seien noch weitere Beispiele von Vektorräumen erwähnt:

- die Menge \mathbb{C} aller komplexen Zahlen $a + ib$ (indem man (a, b) für $a + ib$ schreibt, kann man \mathbb{C} mit \mathbb{R}^2 identifizieren)
- die Menge $\mathbb{R}[x]$ aller reellen Polynome (Beispiel eines unendlichdimensionalen Vektorraums)
- die Menge $\mathbb{R}[x]_n \subseteq \mathbb{R}[x]$ aller reellen Polynome vom Grad $\leq n$ (Beispiel eines $n + 1$ -dimensionalen Vektorraums)
- die Menge $C([a, b])$ aller stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$ (ebenfalls unendlichdimensional)
- Die Lösungsmenge einer homogenen linearen Differentialgleichung (siehe 3.2.1)

Übungsaufgabe 2. (P) Erklären Sie genau, warum \mathbb{C} , $\mathbb{R}[x]$ und $C([a, b])$ Vektorräume sind. Dazu müssen Sie erklären, was Addition, was Nullvektor, was negativer Vektor und was Skalarmultiplikation ist. Außerdem sind die Gesetze aus Definition 1.1.1.1 zu verifizieren.

1.1.2 Unterraum, Linearkombination und Erzeugnis

Inhalt in Kurzfassung: Unter den Teilmengen U eines Vektorraums V sind jene von besonderem Interesse, die selbst einen Vektorraum bilden, die sogenannten Unterräume. $U = \{\mathbf{o}\}$ und $U = V$ sind stets Unterräume (die sogenannten trivialen). Zu einer beliebigen Menge $T \subseteq V$, auch wenn sie selbst kein Unterraum ist, gibt es stets einen kleinsten Unterraum U , der T enthält: das Erzeugnis oder die lineare Hülle von T . Sie besteht aus allen Linearkombinationen von Vektoren aus T . Beispielsweise ist der Standardvektorraum \mathbb{R}^n das Erzeugnis der kanonischen Basis.

Eine Teilmenge U eines Vektorraums V ist dann von besonderem Interesse, wenn U mit den gegebenen Operationen selbst wieder ein Vektorraum ist. Dabei kommt es nur darauf an, dass die auf V definierten Operationen, wenn man sie auf Vektoren aus U anwendet, nicht aus der Menge U hinausführen:

Definition 1.1.2.1. Ist V ein Vektorraum, dann heißt eine Teilmenge $U \subseteq V$ **Unterraum** von V , (manchmal schreibt man dafür symbolisch auch $U \leq V$), wenn gilt:

Für alle $\mathbf{x}, \mathbf{y} \in U$ und $r \in \mathbb{R}$ liegen auch die Vektoren $\mathbf{x} + \mathbf{y}$ und $r\mathbf{x}$ in U , außerdem der Nullvektor: $\mathbf{o} \in U$.

Der Durchschnitt U von zwei oder auch beliebig vielen Unterräumen U_i eines Vektorraumes V ist wieder ein Unterraum von V , denn: Liegen \mathbf{x}, \mathbf{y} in U , so in allen U_i . Weil die U_i Unterräume sind, liegen daher auch $\mathbf{x} + \mathbf{y}$ und $r\mathbf{x}$ für beliebige $r \in \mathbb{R}$ in allen U_i und somit auch in deren Durchschnitt U . Also erfüllt U die Bedingungen in Definition 1.1.2.1.

Diese Beobachtung hat zur Folge, dass es auch zu jeder Teilmenge $M \subseteq V$ eines Vektorraums einen kleinsten Unterraum U von V mit $M \subseteq U$ gibt, nämlich den Durchschnitt aller Unterräume, die M als Teilmenge enthalten. (V selbst ist sicher so ein Unterraum.) Dieser Unterraum U (den wir auch mit $\text{LH}(M)$ bezeichnen werden) lässt sich wie folgt auch anders beschreiben. Dazu holen wir etwas weiter aus.

Ist V ein Vektorraum, $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ und $r_1, \dots, r_n \in \mathbb{R}$, so heißt der Vektor

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{x}_i$$

eine **Linearkombination** (Abkürzung LK) der \mathbf{x}_i . Die Zahlen r_i heißen in diesem Zusammenhang auch **Koeffizienten**. Offenbar sind Unterräume gerade so definiert, dass sie abgeschlossen sind gegenüber LKen, d.h. wenn die \mathbf{x}_i in einem bestimmten Unterraum U liegen, so auch \mathbf{x} . Umgekehrt bildet zu jeder Menge $M \subseteq V$ von Vektoren die sogenannte **lineare Hülle**

$$\text{LH}(M) := \left\{ \sum_{i=1}^n r_i \mathbf{x}_i : n \in \mathbb{N}, r_i \in \mathbb{R}, \mathbf{x}_i \in M \right\}$$

von M , also die Menge aller LKen aus Vektoren in M , einen Unterraum. Folglich ist $\text{LH}(M)$ der kleinste Unterraum, der M enthält, und heißt deshalb auch der von M **erzeugte Unterraum**, (oder das **Erzeugnis** von M oder der von M **aufgespannte Unterraum**, symbolisch oft auch $\text{span}(M)$). Dieser Sachverhalt lässt sich auch so formulieren: $\text{LH}(M)$ ist der Durchschnitt aller Unterräume von V , die M als Teilmenge enthalten, als Formel:

$$\text{LH}(M) = \bigcap \{U \leq V : M \subseteq U\}$$

Ist $\text{LH}(M) = V$, so heißt M ein **Erzeugendensystem** von V . Klarerweise gilt für $M_1 \subseteq M_2 \subseteq V$ stets $\text{LH}(M_1) \subseteq \text{LH}(M_2)$. Für ein Erzeugendensystem M ist folglich auch jede Menge M' mit $M \subseteq M' \subseteq V$ ein Erzeugendensystem. Ist $M = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ endlich, schreiben wir für $\text{LH}(M)$ auch $\text{LH}(\mathbf{a}_1, \dots, \mathbf{a}_n)$ (ohne Mengenklammern).

Übungsaufgabe 3. (P) Begründen Sie all das sorgfältig, indem Sie für einen Vektorraum V folgende drei Behauptungen überprüfen:

1. Ist U ein Unterraum von V mit $\mathbf{x}_1, \dots, \mathbf{x}_n \in U$, außerdem $r_1, \dots, r_n \in \mathbb{R}$, so folgt $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{x}_i \in U$.
2. Sei $M \subseteq V$ irgendeine Teilmenge von V . Dann bildet die Menge $\text{LH}(M)$ aller LKen aus Vektoren von M einen Unterraum von V .
3. Sei wieder $M \subseteq V$ irgendeine Teilmenge von V und U ein Unterraum von V mit $M \subseteq U$. Dann gilt $\text{LH}(M) \subseteq U$.

Jeder Vektorraum hat die Unterräume $\{\mathbf{o}\}$ (der **Nullraum**, der nur aus dem Nullvektor besteht) und V selbst, die beiden sogenannten **trivialen Unterräume**. Im Fall des

1 Lineare Algebra

zweidimensionalen Standardvektorraums $V = \mathbb{R}^2$ kommen zum Nullraum $\{\mathbf{o}\}$ (nulldimensional) und zu $V = \mathbb{R}^2$ (zweidimensional) noch die (eindimensionalen) Unterräume dazu. Jeder von ihnen lässt sich als Gerade durch den Ursprung auffassen. Der dreidimensionale Standardvektorraum $V = \mathbb{R}^3$ enthält außerdem noch zahlreiche nichttriviale zweidimensionale Unterräume, von denen sich jeder als Ebene im Raum durch den Ursprung \mathbf{o} deuten lässt. Warum genau das alle sind, wird die nun zu entwickelnde Theorie zeigen. Zuvor aber noch einige Übungsaufgaben.

Übungsaufgabe 4. (P) Begründen Sie, warum für jeden Vektorraum V der Nullraum $\{\mathbf{o}\}$ ein Unterraum von V ist, indem Sie die Eigenschaften aus Definition 1.1.2.1 nachprüfen.

Übungsaufgabe 5. (T) Welche der folgenden Mengen sind Vektorräume (mit den üblichen Operationen auf \mathbb{R}^2 bzw. \mathbb{R}^3)? Falls kein Vektorraum vorliegt, geben Sie an, welche Eigenschaft verletzt ist!

1. $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : 3x_1 = 2x_2 = x_3\}$
2. $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \cdot x_2 = 0\}$
3. $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = 0\}$
4. $\{(\lambda, \lambda^2) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\}$
5. $\{(\lambda^2, \lambda^2) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\}$
6. $\{(\lambda^3, \lambda^3) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\}$
7. $\{(\mu + \lambda, \mu - \lambda) \in \mathbb{R}^2 : \mu, \lambda \in \mathbb{R}\}$
8. $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq x_2\}$

Übungsaufgabe 6. (T)

Welche der folgenden Mengen von reellen Funktionen sind Vektorräume (mit den üblichen punktweise definierten Operationen)? Falls kein Vektorraum vorliegt, geben Sie an, welche Eigenschaft verletzt ist!

1. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f(x) = f(-x)\}$
2. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \geq 0\}$
3. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ besitzt mindestens eine Nullstelle}\}$
4. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ besitzt höchstens eine Nullstelle}\}$
5. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ stetig}\}$
6. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ stetig und } \int_{-1}^1 f(x)dx = 0\}$
7. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ monoton steigend oder monoton fallend}\}$
8. $\{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ differenzierbar}\}$

Eine besondere Rolle spielen im Standardvektorraum \mathbb{R}^n die sogenannten **kanonischen Einheitsvektoren**

$$\mathbf{e}_1 = \mathbf{e}_1^{(n)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \mathbf{e}_2^{(n)} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \mathbf{e}_n^{(n)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Den oberen Index werden wir oft weglassen, sofern aus dem Kontext klar ist, dass ein Einheitsvektoren im \mathbb{R}^n gemeint ist. Der i -te Einheitsvektor $\mathbf{e}_i = (0, 0, \dots, 0, 1, 0, \dots, 0)$ hat 1 als i -te Komponente, alle anderen Komponenten von \mathbf{e}_i sind 0. Gemeinsam bilden die \mathbf{e}_i

ein Erzeugendensystem von \mathbb{R}^n , weil sich jeder beliebige Vektor $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ in (sogar eindeutiger Weise) als LK der \mathbf{e}_i schreiben lässt, nämlich:

$$\mathbf{x} = (x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \sum_{i=1}^n x_i \mathbf{e}_i$$

Übungsaufgabe 7. (P) Zeigen Sie, dass die Menge aller reellen Zahlen \mathbb{R} einen rationalen Vektorraum bildet.

1.1.3 Lineare Unabhängigkeit

Inhalt in Kurzfassung: Der Begriff der linearen Unabhängigkeit von Vektoren bereitet erfahrungsgemäß manchen Anfängern Schwierigkeiten. Deshalb ist ihm hier ein ganzer Unterabschnitt gewidmet. Die kanonische Basis des \mathbb{R}^n liefert das einfachste Beispiel, mit Hilfe dessen man sich auch die allgemeine Situation besser einprägen kann.

Der Begriff der linearen (Un-)Abhängigkeit von Vektoren, genauer: einer Familie von Vektoren, ist grundlegend für die Theorie der Vektorräume. Im dreidimensionalen Raum beispielsweise sind drei Vektoren linear abhängig, wenn sie in einer Ebene liegen oder, äquivalent dazu und allgemeiner gesprochen, wenn einer von ihnen sich von den anderen erzeugen lässt. Dabei bedeutet *erzeugt werden*: sich als LK darstellen lassen. Insbesondere ist jede LK $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{x}_i$ der Vektoren \mathbf{x}_i von diesen linear abhängig. Bringt man auch \mathbf{x} auf die rechte Seite dieser Gleichung, so ist der Nullvektor $\mathbf{o} = -\mathbf{x} + \sum_{i=1}^n r_i \mathbf{x}_i$ eine LK der Vektoren $\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_n$, die nichttrivial in dem Sinn ist, dass der Vektor \mathbf{x} auch wirklich auftritt, weil vor ihm der Koeffizient $-1 \neq 0$ zu denken ist. Wenn man will, kann man auch sagen: Die lineare Abhängigkeit einer Menge von Vektoren bedeutet, dass sie in Bezug auf den Erzeugungsprozess redundant ist, weil sie wenigstens einen Vektor enthält, der weggelassen werden könnte, ohne dass deshalb die lineare Hülle kleiner würde. Eine exakte Definition lautet wie folgt:

Definition 1.1.3.1. Ein n -Tupel $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ ($n \in \mathbb{N}$, $n \geq 1$) von Vektoren \mathbf{x}_i eines Vektorraumes V heißt **linear abhängig**, wenn es Zahlen $r_1, \dots, r_n \in \mathbb{R}$ gibt, die nicht alle $= 0$ sind und für die

$$\sum_{i=1}^n r_i \mathbf{x}_i = \mathbf{o}$$

gilt. Eine Menge $M \subseteq V$ von Vektoren in V heißt **linear abhängig**, wenn es ein $n \in \mathbb{N}$ gibt und paarweise verschiedene Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n \in M$, so dass $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ linear abhängig ist.

Ein n -Tupel $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ mit $\mathbf{x}_i \in V$ bzw. eine Teilmenge $M \subseteq V$ heißt **linear unabhängig**, wenn es bzw. sie nicht linear abhängig ist. (Die leere Menge ist definitionsgemäß linear unabhängig.)

Wie die dieser Definition vorangegangene Diskussion zeigt, lässt sich – vielleicht leichter fasslich – auch formulieren:

Proposition 1.1.3.2. *Eine Menge $M \subseteq V$ von Vektoren eines Vektorraumes V ist genau dann linear abhängig, wenn einer der Vektoren aus M sich als Linearkombination anderer Vektoren aus M darstellen lässt, formal:*

Es gibt ein $n \in \mathbb{N}$ und paarweise verschiedene Vektoren $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in M$ sowie Zahlen r_1, r_2, \dots, r_n mit

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{x}_i.$$

Eine triviale Beobachtung lautet: Jede Menge von Vektoren, die den Nullvektor enthält, ist linear abhängig (insbesondere der Nullvektor für sich). Denn dann kann man $\mathbf{x}_1 := \mathbf{o}$ und $n = 0$ setzen, weil die leere Summe ja immer den Nullvektor darstellt.

In Definition 1.1.3.1 mag die Unterscheidung zwischen Mengen und n -Tupeln etwas verwirren. Dazu ein paar Worte zur Erklärung und teilweise auch zur Entwarnung beziehungsweise Vereinfachung der Sprechweise: Ob ein n -tupel $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ linear abhängig oder unabhängig ist, hängt nicht von der Reihenfolge der Komponenten ab, weil man mit den \mathbf{x}_i ja auch die Koeffizienten r_i in einer LK umnummerieren kann. Deshalb sagt man auch einfach die Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ seien linear abhängig bzw. unabhängig. Dennoch bedeutet das nicht unbedingt, dass das Gleiche auch für die Menge $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ gilt. Die Tücke, deretwegen wir so sorgfältig unterscheiden müssen, liegt ausschließlich daran, dass zwei der \mathbf{x}_i gleich sein könnten. Beispiel: Das 2-Tupel (Paar) $(\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{x}, \mathbf{x})$ ist nämlich linear abhängig, weil für die Koeffizienten $r_1 = 1$ und $r_2 = -1$ die LK $r_1 \mathbf{x}_1 + r_2 \mathbf{x}_2 = \mathbf{x} - \mathbf{x} = \mathbf{o}$ ergibt. Hingegen ist die Menge $M = \{\mathbf{x}_1, \mathbf{x}_2\} = \{\mathbf{x}, \mathbf{x}\} = \{\mathbf{x}\}$ linear unabhängig, sofern $\mathbf{x} \neq \mathbf{o}$, denn dann ist $r\mathbf{x} = \mathbf{o}$ nur für $r = 0$ möglich (siehe Übungsaufgabe 1, Teil 4). Außerdem wird später vor allem bei den Matrizendarstellungen die Reihenfolge der zugrunde gelegten (linear unabhängigen) Basisvektoren wichtig sein, weshalb die Darstellung als Tupel der als (ungeordnete) Menge vorzuziehen ist. Der entscheidende (weil erfahrungsgemäß von vielen als schwierig empfundene) Punkt in der Definition der linearen Unabhängigkeit besteht darin, dass *nur* die triviale LK (mit allen Koeffizienten = 0) den Nullvektor darstellt. (Dass die triviale LK den Nullvektor darstellt, ist hingegen trivial, weil es auch für linear abhängige Vektoren gilt, und verdient keine besondere Hervorhebung.)

Übungsaufgabe 8. (P) Begründen Sie für Teilmengen $M_1 \subseteq M_2 \subseteq V$ des Vektorraums V : Ist M_1 linear abhängig, so auch M_2 . Ist M_2 linear unabhängig, so auch M_1 .

Eine für das Weitere typische Situation besteht darin, Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ auf lineare Abhängigkeit überprüfen zu müssen. Zur Übung spielen wir diese Situation nun für kleine Werte von n durch.

Für $n = 0$ geht es um die leere Menge, die definitionsgemäß linear unabhängig ist.

Für $n = 1$ haben wir schon oben gesehen: Ist $\mathbf{x}_1 = \mathbf{o}$ der Nullvektor, so liegt lineare Abhängigkeit vor, jeder andere Vektor ist für sich linear unabhängig.

Für $n = 2$ bedeutet lineare Abhängigkeit, dass es Zahlen r_1, r_2 gibt, von denen wenigstens eine, z.B. $r_2 \neq 0$ von 0 verschieden ist, und $r_1 \mathbf{x}_1 + r_2 \mathbf{x}_2 = \mathbf{o}$ gilt. Dann können wir umformen zu $\mathbf{x}_2 = (-r_2^{-1} r_1) \mathbf{x}_1$, was bedeutet, dass \mathbf{x}_2 ein Vielfaches (eine LK) von \mathbf{x}_1 ist. Anschaulich lässt sich sagen: \mathbf{x}_1 und \mathbf{x}_2 liegen auf einer Geraden, sind also **kollinear**. Sind \mathbf{x}_1 und \mathbf{x}_2 hingegen linear unabhängig, so spannen sie eine Ebene auf.

Für $n = 3$ überlegt man sich ähnlich zum Fall $n = 2$, dass im linear abhängigen Fall einer der drei Vektoren (dessen Koeffizient von 0 verschieden ist) als LK der anderen beiden darstellbar ist, d.h. in der von diesen aufgespannten Ebene liegt. Drei linear unabhängige Vektoren hingegen haben nicht Platz in einer Ebene, sondern spannen, wie wir uns gleich noch genauer überlegen werden, einen dreidimensionalen Raum auf.

Für $n \geq 4$ entzieht sich die geometrische Illustration n -dimensionaler Räume unserer unmittelbaren Anschauung. Dennoch lohnt es, sich gewisse Intuitionen dafür zu erarbeiten, weil vor allem beim Lösen linearer Gleichungssysteme in mehreren Variablen entsprechende Situationen auch in der Praxis auftreten.

Ein wichtiges Beispiel linear unabhängiger Vektoren in \mathbb{R}^n bilden die kanonischen Einheitsvektoren \mathbf{e}_i . Denn der Nullvektor lässt sich als LK

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \mathbf{o} = \sum_{i=1}^n r_i \mathbf{e}_i = r_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + r_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + r_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}$$

nur dann schreiben, wenn $r_1 = r_2 = \dots = r_n = 0$ ist.

Übungsaufgabe 9. (T) Zeigen Sie, warum die folgenden Teilmengen des \mathbb{R}^3 linear abhängig sind, indem sie jeweils einen Vektor dieser Menge (nicht den Nullvektor) als Linearkombination der anderen darstellen.

1. $\{(a^2, b^2, c^2) : a, b, c \in \mathbb{R}\}$
2. $\{(1, 2, n) : n = 2, 3, 4\}$
3. $\{(1, n, n^2) : n = 1, 2, \dots, 5\}$
4. $\{(t, 2t, 3t) : 0 \leq t \leq 1\}$

Übungsaufgabe 10. (T) Geben Sie für die Mengen aus Aufgabe 9 jeweils den erzeugten Unterraum von \mathbb{R}^3 , gegebenenfalls in Form einer Ebenen- bzw. Geradengleichung an.

In den folgenden Aufgaben steht 1 auch für die konstante Funktion mit dem Wert 1:

Übungsaufgabe 11. (T*) Sei $\mu, \lambda \neq 0$ fest. Ist $\{1, e^{\mu x}, e^{\lambda x}\}$ eine linear unabhängige Teilmenge des Vektorraums $C[-1, 1]$?

Anleitung: Unterscheiden Sie die beiden Fälle $\mu = \lambda$ und $\mu \neq \lambda$ (einer davon ist sehr leicht). Nehmen Sie an, es gibt eine Linearkombination des Nullvektors

$$a \cdot 1 + b \cdot e^{\mu x} + c \cdot e^{\lambda x} = 0$$

und differenzieren Sie diese Gleichung! Dividieren Sie anschließend durch $e^{\mu x}$ (oder durch $e^{\lambda x}$).

Übungsaufgabe 12. (T^*) Wie Aufgabe 11 aber mit $\{1, \sin \mu x, \sin \lambda x\} \subseteq C[0, 2\pi]$. Hinweis: Euler'sche Formel.

Übungsaufgabe 13. (T) Ist die Menge $\{1, \cos^2 x, \cos 2x\} \subseteq C[0, 2\pi]$ linear unabhängig?

Übungsaufgabe 14. (T) Wie Aufgabe 13 aber mit $\{1, \sin^2 x, \cos 2x\} \subseteq C[0, 2\pi]$.

1.1.4 Basen

Inhalt in Kurzfassung: Eine Basis B eines Vektorraums V ist definiert als linear unabhängiges Erzeugendensystem von V . Äquivalent dazu sind folgende drei Aussagen: 1. B ist ein minimales Erzeugendensystem. 2. B ist eine maximale linear unabhängige Menge. 3. Jeder Vektor in V lässt sich in eindeutiger Weise als LK von Vektoren aus B schreiben. Die dritte Charakterisierung rechtfertigt es, von den (eindeutig bestimmten) Koordinaten oder Komponenten eines (beliebigen) Vektors aus V bezüglich B zu sprechen. Beispiel: Die Komponenten eines Vektors aus dem Standardvektorraum \mathbb{R}^n sind gleichzeitig seine Komponenten bezüglich der kanonischen Basis.

Ist V wieder ein Vektorraum und $M \subseteq V$, so wissen wir bereits: Ist M ein Erzeugendensystem von V so auch jede Obermenge M' mit $M \subseteq M' \subseteq V$. Ist M linear unabhängig, so auch jede Teilmenge M' mit $M' \subseteq M \subseteq V$. Also sind Erzeugendensysteme in einem gewissen Sinn groß, linear unabhängige Teilmengen klein. Von besonderem Interesse sind Teilmengen $B \subseteq V$, die beide Eigenschaften in sich vereinen:

Definition 1.1.4.1. Jedes linear unabhängige Erzeugendensystem eines Vektorraumes V heißt eine **Basis** von V .

Bei einer Basis kommt es oft auf die Reihenfolge der Vektoren an, weshalb man statt $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ (mit paarweiser verschiedenen \mathbf{b}_i) manchmal genauer $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ schreiben sollte.

Das wichtigste Beispiel einer Basis des Standardvektorraums \mathbb{R}^n kennen wir schon. Denn die kanonischen Einheitsvektoren $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ bilden sowohl ein Erzeugendensystem (siehe 1.1.2) als auch eine linear unabhängige Menge (siehe 1.1.3). Man nennt $\{\mathbf{e}_1 = \mathbf{e}_1^{(n)}, \mathbf{e}_2 = \mathbf{e}_2^{(n)}, \dots, \mathbf{e}_n = \mathbf{e}_n^{(n)}\}$ die **kanonische Basis** von \mathbb{R}^n .

Basen eines Vektorraumes lassen sich auf vielfache Art charakterisieren.

Satz 1.1.4.2. Für eine Teilmenge $B \subseteq V$ eines Vektorraums V sind die folgenden Aussagen äquivalent:

1. B ist eine Basis.
2. Jeder Vektor $\mathbf{x} \in V$ hat eine eindeutige (!) Darstellung als LK

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i$$

von Vektoren $\mathbf{b}_i \in B$ d.h.

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i = \sum_{i=1}^n s_i \mathbf{b}_i \quad \text{impliziert} \quad r_i = s_i \quad \text{für} \quad i = 1, \dots, n.$$

Das bedeutet genauer: Zu jedem $\mathbf{x} \in V$ gibt es genau eine endliche Teilmenge $B_0 := \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq B$ (die \mathbf{b}_i paarweise verschieden), so dass sich \mathbf{x} wie oben mit $r_i \neq 0$ darstellen lässt. Die zugehörigen Koeffizienten r_i sind durch \mathbf{x} eindeutig bestimmt. (Für $\mathbf{x} = \mathbf{o}$ ist $B_0 = \emptyset$ leer und die LK die leere Summe.)

3. B ist ein minimales Erzeugendensystem, d.h.: Entfernt man irgendeinen Vektor $\mathbf{b} \in B$ aus B , so ist die verbleibende Menge $B \setminus \{\mathbf{b}\}$ kein Erzeugendensystem von V .
4. B ist ein maximales linear unabhängiges System, d.h.: Fügt man zu B irgendeinen Vektor $\mathbf{x} \in V \setminus B$ hinzu, so ist $B \cup \{\mathbf{x}\}$ linear abhängig.

Beweis. Wir führen den Beweis der Äquivalenz der vier Aussagen zyklisch, indem wir die Implikationen $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$ beweisen:

$1 \rightarrow 2$: Sei B eine Basis. Nach Definition 1.1.4.1 ist B insbesondere ein Erzeugendensystem, weshalb jeder Vektor \mathbf{x} eine Darstellung als LK der \mathbf{b}_i hat. Wir beweisen nun die Eindeutigkeit der Darstellung. Wir nehmen an, zwei Darstellungen

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i = \sum_{j=1}^{n'} r'_j \mathbf{b}'_j$$

mit $n, n' \in \mathbb{N}$, $\mathbf{b}_i, \mathbf{b}'_j \in B$ und $r_i, r'_j \in \mathbb{R}$ seien gegeben. Sollten nicht alle \mathbf{b}_i unter den \mathbf{b}'_j vorkommen oder vice versa, ergänzen wir die Summen rechts und links entsprechend mit Koeffizienten = 0, so dass die Menge der \mathbf{b}_i mit jener der \mathbf{b}'_j übereinstimmt. Durch Ummummerierung erreichen wir $\mathbf{b}_i = \mathbf{b}'_i$. Wir können also Darstellungen

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i = \sum_{i=1}^{n'} r'_i \mathbf{b}_i$$

annehmen (wobei jetzt auch $r_i = 0$ oder $r'_i = 0$ erlaubt ist). Die Differenz

$$\sum_{i=1}^n r_i \mathbf{b}_i - \sum_{i=1}^{n'} r'_i \mathbf{b}_i = \sum_{i=1}^n (r_i - r'_i) \mathbf{b}_i = \mathbf{o}$$

stellt den Nullvektor dar. Wegen der linearen Unabhängigkeit von B ist das nur möglich, wenn $r_i - r'_i = 0$, also $r_i = r'_i$ für alle $i = 1, 2, \dots, n$ gilt. Damit ist die Eindeutigkeit der Darstellung bewiesen.

$2 \rightarrow 3$: Wir setzen voraus, dass sich jeder Vektor $\mathbf{x} \in V$ in der beschriebenen Weise eindeutig als LK von Vektoren aus B darstellen lässt. Jeder Vektor $\mathbf{b} \in B$ lässt sich in trivialer Weise als Summe von sich selbst als einzigem Summanden mit Koeffizienten 1

darstellen: $\mathbf{b} = \sum_{i=1}^1 r_1 \mathbf{b}_1$ mit $\mathbf{b}_1 = \mathbf{b}$ und $r_1 = 1$. Wegen der vorausgesetzten Eindeutigkeit der Darstellung gibt es keine Darstellung von \mathbf{b} als LK von Vektoren aus $B \setminus \{\mathbf{b}\}$. Folglich ist $B \setminus \{\mathbf{b}\}$ kein Erzeugendensystem von V .

3 \rightarrow 4: Nun lautet unsere Voraussetzung, dass B ein minimales Erzeugendensystem ist. Daraus müssen wir ableiten, dass B eine maximale linear unabhängige Menge ist.

Zunächst zur linearen Unabhängigkeit: Wäre B linear abhängig, so ließe sich ein $\mathbf{b} \in B$ als LK anderer $\mathbf{b}_1, \dots, \mathbf{b}_n$ schreiben. Wir behaupten, dass dann auch $B \setminus \{\mathbf{b}\}$ ein Erzeugendensystem von V wäre. Denn ein beliebiger Vektor $\mathbf{x} \in V$ lässt sich als LK gewisser $\mathbf{b}'_1, \dots, \mathbf{b}'_m \in B$ schreiben. Kommt \mathbf{b} unter diesen nicht vor, haben wir bereits $\mathbf{x} \in L(\{\mathbf{b}'_1, \dots, \mathbf{b}'_m\}) \subseteq B \setminus \{\mathbf{b}\}$. Andernfalls ersetzen wir jenes \mathbf{b}'_i mit $\mathbf{b}'_i = \mathbf{b}$ durch die Darstellung von \mathbf{b} als LK der $\mathbf{b}_1, \dots, \mathbf{b}_n \neq \mathbf{b}$. Auch in diesem Fall erhalten wir eine Darstellung von \mathbf{x} , die $\mathbf{x} \in L(B \setminus \{\mathbf{b}\})$ belegt. Also ist wie behauptet $B \setminus \{\mathbf{b}\}$ ein Erzeugendensystem von V . B wäre also nicht minimal mit dieser Eigenschaft, was im Widerspruch zur Voraussetzung steht.

B ist unter den linear unabhängigen Teilmengen von V maximal: Sei $\mathbf{x} \in V \setminus B$. Weil B ein Erzeugendensystem ist, gilt $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i$ mit geeigneten $\mathbf{b}_i \in B$ und $r_i \in \mathbb{R}$. Damit ist $\mathbf{o} = -\mathbf{x} + \sum_{i=1}^n r_i \mathbf{b}_i$ eine nichttriviale Darstellung des Nullvektors als LK von Vektoren aus $B \cup \{\mathbf{x}\}$. Also ist $B \cup \{\mathbf{x}\}$ linear abhängig und (weil $\mathbf{x} \in V \setminus B$ beliebig war) B als linear unabhängige Menge maximal.

4 \rightarrow 1: Sei nun B eine maximale linear unabhängige Teilmenge von V . Um zu zeigen, dass B eine Basis ist, müssen wir einen beliebigen Vektor \mathbf{x} als LK von geeigneten Vektoren aus B darstellen. Ist $\mathbf{x} = \mathbf{b} \in B$, so ist $\mathbf{x} = 1\mathbf{b}$ bereits so eine Darstellung (mit nur einem Summanden). Für $\mathbf{x} \notin B$ ist $B \cup \{\mathbf{x}\}$ eine echte Obermenge von B , laut der vorausgesetzten Maximalität von B als linear unabhängige Menge also linear abhängig, weil ja B maximal linear unabhängig ist. Somit gibt es eine nichttriviale Darstellung des Nullvektors als LK $\mathbf{o} = r\mathbf{x} + \sum_{i=1}^n r_i \mathbf{b}_i$, also $r\mathbf{x} = -\sum_{i=1}^n r_i \mathbf{b}_i$. Wäre $r = 0$, so müsste wegen der Nichttrivialität der Darstellung einer der Koeffizienten r_i von 0 verschieden sein. Somit läge eine nichttriviale Darstellung von \mathbf{o} als LK der $\mathbf{b}_i \in B$ vor, was der linearen Unabhängigkeit von B widerspräche. Also ist $r \neq 0$, und wir können umformen zu $\mathbf{x} = r^{-1}r\mathbf{x} = \sum_{i=1}^n -r^{-1}r_i \mathbf{b}_i$, eine LK der \mathbf{b}_i , was zu zeigen war. \square

In der eindeutigen Darstellung

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i$$

eines Vektors $\mathbf{x} \in V$ als LK der Basisvektoren \mathbf{b}_i , wie sie in Aussage 2 von Satz 1.1.4.2 beschrieben ist, heißen die r_i auch die **Koordinaten** oder **Komponenten** von \mathbf{x} bezüglich der Basis $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. Man beachte, dass die Indizes i die Zuordnung der r_i zu \mathbf{b}_i beinhalten. Deshalb wäre es präziser, als Basis das n -tupel $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ statt der (ungeordneten) Menge $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ anzusehen. In diesem Fall schließen wir uns aber der traditionellen Sprechweise an, die zwar etwas ungenau, aber weniger umständlich ist.

Übungsaufgabe 15. (T) Welches der folgenden Systeme $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ bzw. $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$

stellt eine Basis im \mathbb{R}^3 dar?

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix};$$

$$\mathbf{b}_1 = \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} -2 \\ -1 \\ 3 \end{pmatrix}.$$

Übungsaufgabe 16. (T) Seien n, m, k die ersten drei Ziffern Ihrer Matrikelnummer und $\mathbf{b}_1 = (n, m, k)$, $\mathbf{b}_2 = (-m, n, 0)$. Lässt sich das System $(\mathbf{b}_1, \mathbf{b}_2)$ zu einer Basis des \mathbb{R}^3 ergänzen?

Übungsaufgabe 17. Ist die Menge der Elemente $\mathbf{x} \in \mathbb{R}^3$, die der Bedingung $\mathbf{n} \cdot \mathbf{x} = 0$ (Skalarprodukt) mit $\mathbf{n} = (-1, 7, 2)$ genügen ein Vektorraum? Bestimmen Sie gegebenenfalls ein minimales Erzeugendensystem und fertigen Sie eine Skizze an! Welche Rolle spielt der Vektor \mathbf{n} ?

Übungsaufgabe 18. (P) Zeigen Sie, dass die Menge $\mathbb{R}[x]_3$ aller reellen Polynome vom Grad höchstens 3 einen reellen Vektorraum bildet. Geben Sie ein minimales Erzeugendensystem von $\mathbb{R}[x]_3$ an.

Übungsaufgabe 19. (P) Zeigen Sie, dass die Menge aller komplexen Zahlen \mathbb{C} einen reellen Vektorraum bildet. Geben Sie ein minimales Erzeugendensystem dieses reellen Vektorraums an.

Übungsaufgabe 20. (T) Betrachten Sie die Menge

$$E := \{(x, y, z) \in \mathbb{R}^3 : x - y = z\}. \quad (1.1)$$

Ist E (mit den üblichen Operationen) ein Vektorraum? Wenn ja, finden Sie eine maximale linear unabhängige Teilmenge von E .

1.1.5 Dimension

Inhalt in Kurzfassung: Der Austauschsatz von Steinitz zeigt (wenigstens für endlichdimensionale, d.h. von endlich vielen Vektoren erzeugte Vektorräume), dass jeder Vektorraum V eine Basis hat und dass je zwei Basen von V gleich viele Elemente haben. Deshalb ist es sinnvoll, die Dimension von V zu definieren als Kardinalität einer (beliebigen) Basis von V . Die kanonische Basis von \mathbb{R}^n besteht aus n Vektoren. Deshalb hat der Standardvektorraum \mathbb{R}^n auch im Sinne dieser Definition die Dimension n .

Für den n -dimensionalen Standardvektorraum \mathbb{R}^n besteht die kanonische Basis aus genau n Vektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$. Ist es möglich, den \mathbb{R}^n durch geschickte Wahl auch mit weniger Vektoren aufzuspannen? Die Antwort lautet *nein*. Die entscheidende Überlegung wird sich auf Satz 1.1.5.2 beziehen, der wiederum auf folgender einfachen Beobachtung beruht:

Hilfssatz 1.1.5.1. *Ist V ein Vektorraum und*

$$\mathbf{x} = \sum_{i=1}^n r_i \mathbf{a}_i$$

mit $n \in \mathbb{N}$, $r_i \in \mathbb{R}$, $\mathbf{x}, \mathbf{a}_1, \dots, \mathbf{a}_n \in V$, wobei $r_1 \neq 0$ gelte. Dann gilt:

$$\text{LH}(\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}) = \text{LH}(\{\mathbf{x}, \mathbf{a}_2, \dots, \mathbf{a}_n\})$$

Beweis. Nach Voraussetzung gilt $\mathbf{x} \in \text{LH}(\{\mathbf{a}_1, \dots, \mathbf{a}_n\})$, folglich ist die rechte lineare Hülle in der linken enthalten. Es bleibt die umgekehrte Inklusion \subseteq zu zeigen, wofür wiederum der Nachweis von $\mathbf{a}_1 \in \text{LH}(\{\mathbf{x}, \mathbf{a}_2, \dots, \mathbf{a}_n\})$ genügt. Der ergibt sich aber unmittelbar, weil sich die vorausgesetzte Gleichung $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{a}_i$ wegen $r_1 \neq 0$ zur LK

$$\mathbf{a}_1 = r_1^{-1}(\mathbf{x} - r_2 \mathbf{a}_2 - \dots - r_n \mathbf{a}_n)$$

umschreiben lässt. Also ist tatsächlich $\mathbf{a}_1 \in \text{LH}(\{\mathbf{x}, \mathbf{a}_2, \dots, \mathbf{a}_n\})$. \square

Damit lässt sich rasch der **Austauschsatz von Steinitz** beweisen, der die Grundlage der eindeutigen Definition der Dimension eines Vektorraums sein wird:

Satz 1.1.5.2. *In jedem Erzeugendensystem $E = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ eines Vektorraums V und zu jedem linear unabhängigen k -Tupel $(\mathbf{b}_1, \dots, \mathbf{b}_k) \subseteq V$ mit $\mathbf{b}_i \in V$ lassen sich die \mathbf{a}_i so umnummerieren, dass $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n\}$ ebenfalls ein Erzeugendensystem von V ist. Insbesondere ist $k \leq n$, d.h. in der linearen Hülle von n Vektoren können nie mehr als n Vektoren linear unabhängig sein.*

Beweis. Wir gehen mittels Induktion nach k vor. Für $k = 0$ ist nichts zu zeigen. Für den Induktionsschritt haben wir unter der Voraussetzung, dass die Behauptung für irgendein $k \in \mathbb{N}$ gilt, zu schließen, dass sie auch für $k + 1$ gilt.

Sei dazu $(\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{b}_{k+1})$ linear unabhängig in V . Nach Induktionsvoraussetzung lassen sich die \mathbf{a}_i so umnummerieren, dass $E_k := \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+1}, \dots, \mathbf{a}_n\}$ ein Erzeugendensystem von V ist. Insbesondere lässt sich $\mathbf{b}_{k+1} \in V$ als LK aus E_k schreiben, also

$$\mathbf{b}_{k+1} = \sum_{i=1}^k r_i \mathbf{b}_i + \sum_{i=k+1}^n s_i \mathbf{a}_i$$

mit $r_i, s_i \in \mathbb{R}$. Wären alle $s_i = 0$, so wäre \mathbf{b}_{k+1} eine LK der übrigen \mathbf{b}_i , was der linearen Unabhängigkeit der \mathbf{b}_i widerspräche. Also ist mindestens eines der s_i von 0 verschieden. Wir nummerieren die \mathbf{a}_i so um, dass das für $i = k + 1$ der Fall ist. Hilfssatz 1.1.5.1 besagt, dass dann $V = \text{LH}(E_k) = \text{LH}(E_{k+1})$ mit $E_{k+1} := \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_k, \mathbf{b}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n\}$ gilt, was zu zeigen war. \square

Ein Vektorraum heißt **endlichdimensional**, wenn er ein endliches Erzeugendensystem hat, andernfalls **unendlichdimensional**. Mit dieser Sprechweise können wir nun den folgenden wichtigen Satz beweisen:

Satz 1.1.5.3. *Sei V ein endlichdimensionaler Vektorraum. Dann gilt:*

1. V hat eine endliche Basis B .
2. Jedes beliebig vorgegebene Erzeugendensystem E von V enthält eine Basis.
3. Jede beliebig vorgegebene linear unabhängige Teilmenge von V kann zu einer Basis erweitert werden.
4. Je zwei Basen haben gleich viele Elemente.

Beweis. Im ersten Schritt überlegen wir uns, dass V eine Basis hat, die Teilmenge eines beliebig vorgegebenen Erzeugendensystems $E = \{\mathbf{a}_1, \dots, \mathbf{a}_m\}$ von V ist. Sollte E linear abhängig sein, dann lässt sich einer der Vektoren als LK der anderen darstellen. Nach geeigneter Umnummerierung sei das der Vektor $\mathbf{a}_m = \sum_{i=1}^{m-1} r_i \mathbf{a}_i$. Dann können wir \mathbf{a}_m aus E streichen, ohne dass dadurch die lineare Hülle der verbleibenden Vektoren kleiner würde. Also ist $E_1 := \{\mathbf{a}_1, \dots, \mathbf{a}_{m-1}\}$ weiterhin ein Erzeugendensystem von V . Diese Prozedur setzen wir so lange fort, bis ein linear unabhängiges Erzeugendensystem, also eine Basis $B := \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ mit $n \leq m$ übrig bleibt. Also ist eine Basis $B \subseteq E$ mit n Elementen gefunden. Damit ist Aussage 2 bewiesen und, weil V als endlichdimensionaler Vektorraum ein (endliches) Erzeugendensystem enthält, auch Aussage 1.

Ist eine linear unabhängige Menge $\{\mathbf{b}_1, \dots, \mathbf{b}_k\} \subseteq V$ gegeben, und ist die Menge $B := \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ eine Basis von V (wie sie nach 1 existiert), so lassen sich wieder nach dem Austauschsatz 1.1.5.2 die \mathbf{a}_i so umnummerieren, dass $\{\mathbf{b}_1, \dots, \mathbf{b}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \dots, \mathbf{a}_n\}$ eine Basis von V ist, die also alle \mathbf{b}_j enthält. Das beweist Aussage 3.

Es bleibt Aussage 4 zu zeigen, nämlich dass auch jede andere Basis B' genau n Elemente hat. Weil V die lineare Hülle von B ist, folgt aus dem Austauschsatz 1.1.5.2 $|B'| \leq n$, also $B' = \{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ mit $m \leq n$. Doch auch mit vertauschten Rollen von B und B' sind die Voraussetzungen erfüllt, weil B' ein Erzeugendensystem von V und B linear unabhängig ist. Folglich gilt auch $n \leq m$, insgesamt also $m = n$. \square

Satz 1.1.5.3 gilt auch für unendlichdimensionale Vektorräume. Für einen strengen Beweis sind mengentheoretische Methoden erforderlich, die bereitzustellen hier den Rahmen sprengen würde. Im Wesentlichen geht es aber schlicht darum zu rechtfertigen, dass jede linear unabhängige Teilmenge A eines beliebigen (möglicherweise also auch unendlichdimensionalen Vektorraums V) so lange durch linear unabhängige Elemente ergänzt werden kann, bis es keine weiteren gibt. Auf diese Weise erhält man eine maximale linear unabhängige Menge B , die nach Satz 1.1.4.2 eine Basis von V ist und überdies A enthält. Weil man auch noch zeigen kann, dass es zwischen je zwei Basen B_1 und B_2 eine bijektive Abbildung $f : B_1 \rightarrow B_2$ gibt, gilt (hier ohne formalen Beweis) insgesamt:

Satz 1.1.5.4. *Jede linear unabhängige Teilmenge A eines (eventuell auch unendlichdimensionalen) Vektorraums V kann zu einer Basis $B \supseteq A$ von V erweitert werden. Insbesondere ($A = \emptyset$ setzen) enthält jeder Vektorraum eine Basis. Je zwei Basen von V haben gleich viele Elemente.*

Für unsere Zwecke wird die endlichdimensionale Fassung aber vollauf genügen. Durch Satz 1.1.5.3 bekommt die folgende Definition ihren Sinn:

Definition 1.1.5.5. Die Dimension $\dim V$ eines Vektorraumes V ist definiert als die Anzahl $|B|$ der Elemente einer (und somit, nach Satz 1.1.5.3 jeder beliebigen) Basis B von V , welche nach Satz 1.1.5.4 auch existiert.

Der Vektorraum \mathbb{R}^n hat die n -elementige Teilmenge $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ als Basis. Also gilt im Sinn von Definition 1.1.5.5 $\dim \mathbb{R}^n = n$, so wie wir das durch die Sprechweise schon vorweggenommen haben. Der 1-dimensionale Standardvektorraum \mathbb{R}^1 ist identisch mit den reellen Zahlen \mathbb{R} , aufgefasst als Vektorraum (bezüglich der üblichen Operationen). Seine kanonische Basis besteht aus einem einzigen Vektor \mathbf{e}_1 , nämlich der Zahl $\mathbf{e}_1 = 1$. Ihre LKs $r\mathbf{e}_1 = r1 = 1$ sind genau sämtliche reellen Zahlen $r \in \mathbb{R}$. Es gibt sogar 0-dimensionale Vektorräume. Diese bestehen nur aus einem Nullvektor \mathbf{o} . Basis ist die leere Menge $B = \emptyset$. Die einzige LK, die sich daraus bilden lässt, ist die leere Summe, die definitionsgemäß den Nullvektor darstellt. Also ist auch $\dim \mathbb{R}^0 = \dim\{\mathbf{o}\} = 0$. Am vertrautesten sind die Standardvektorräume \mathbb{R}^2 und \mathbb{R}^3 der Dimensionen 2 und 3, die wir uns als Ebene bzw. als Anschauungsraum vorstellen.

Der Begriff der Dimension spielt für Vektorräume eine ähnliche Rolle wie der der Kardinalität oder Mächtigkeit für Mengen ohne weitere Struktur. Eine Analogie zu $|A| \leq |B|$ für eine Teilmenge $A \subseteq B$ zeigt folgende Übungsaufgabe:

Übungsaufgabe 21. (E) Zeigen Sie: Ist U ein Unterraum von V , so folgt $\dim U \leq \dim V$, wobei Gleichheit der Dimensionen nur im Fall von $U = V$ gilt.

Ein weiteres Beispiel ist der Dimensionssatz, dessen Analogie zum Inklusions-Exklusionsprinzip ($|A \cup B| + |A \cap B| = |A| + |B|$ für beliebige Mengen A, B) ins Auge springt:

Satz 1.1.5.6. Sind U_1 und U_2 Unterräume eines Vektorraums V . Dann gilt:

$$\dim(\text{LH}(U_1 \cup U_2)) + \dim(U_1 \cap U_2) = \dim(U_1) + \dim(U_2)$$

Beweis. Wir führen den Beweis für den endlichdimensionalen Fall. Für den unendlichdimensionalen verläuft er sehr ähnlich, allerdings verwendet er hier nicht zur Verfügung stehende unendliche Kardinalitäten.

Sei $U_0 := U_1 \cap U_2$, $d_i := \dim(U_i)$ für $i = 0, 1, 2$, $U := \text{LH}(U_1 \cup U_2) \geq d_1, d_2$ und $d := \dim U$. Dann gibt es eine Basis $B_0 = \{\mathbf{b}_1, \dots, \mathbf{b}_{d_0}\}$ von U_0 . Diese lässt sich durch Vektoren $\mathbf{b}_{d_0+1}, \dots, \mathbf{b}_{d_1}$ zu einer Basis B_1 von U_1 ergänzen, analog durch $\mathbf{b}'_{d_0+1}, \dots, \mathbf{b}'_{d_2}$ zu einer Basis B_2 von U_2 . Wir behaupten, dass dann $B := B_1 \cup B_2$ eine Basis von U ist. Unter Verwendung des Inklusions-Exklusionsprinzips folgt daraus die Behauptung

$$\dim(U) + \dim(U_0) = |B| + |B_0| = |B_1 \cup B_2| + |B_1 \cap B_2| = |B_1| + |B_2| = \dim(U_1) + \dim(U_2).$$

Klarerweise ist B ein Erzeugendensystem von U . Zum Nachweis der linearen Unabhängigkeit müssen wir für eine LK $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1 + \mathbf{u}_2$ mit

$$\mathbf{u}_0 = \sum_{i=1}^{d_0} r_i \mathbf{b}_i \in U_0, \quad \mathbf{u}_1 = \sum_{j=d_0+1}^{d_1} s_j \mathbf{b}_j \in U_1, \quad \mathbf{u}_2 = \sum_{k=d_0+1}^{d_2} s'_k \mathbf{b}'_k \in U_2$$

zeigen, dass $\mathbf{u} = \mathbf{o}$ nur für $r_i = s_j = s'_k = 0$ möglich ist. Ist $\mathbf{u} = \mathbf{u}_0 + \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{o}$, so bedeutet das $\mathbf{u}_1 = -(\mathbf{u}_0 + \mathbf{u}_2)$. Weil sowohl $\mathbf{u}_0 \in U_0 \subseteq U_2$ als auch \mathbf{u}_2 in U_2 liegen, bedeutet das $\mathbf{u}_1 \in U_2$. Außerdem gilt $\mathbf{u}_1 \in U_1$ gilt, also $\mathbf{u}_1 \in U_1 \cap U_2 = U_0$. So ein Vektor hat eine Darstellung als LK aus $B_0 \subseteq B_1$. Wegen der linearen Unabhängigkeit von B_1 muss die Darstellung von \mathbf{u}_1 als LK $\mathbf{u}_1 = \sum_{j=d_0+1}^{d_1} s_j \mathbf{b}_j$ trivial sein: $s_j = 0$ für $j = d_0 + 1, \dots, d_1$, also $\mathbf{u}_1 = \mathbf{o}$. Somit gilt $\mathbf{u}_2 = -\mathbf{u}_0 \in U_0$, was aus analogen Gründen (lineare Unabhängigkeit von $B_2 \supseteq B_0$) nur dann möglich ist, wenn auch $r_i = 0$ für $i = 1, \dots, d_0$ und $s'_k = 0$ für $k = d_0 + 1, \dots, d_2$ gilt. Nichts anderes war zu zeigen. \square

1.2 Lineare Abbildungen

Zwischen Vektorräumen V_1, V_2 sind solche Abbildungen $f : V_1 \rightarrow V_2$ von besonderem Interesse, die mit den Vektorraumoperationen verträglich sind. Solche Abbildungen nennt man linear. In (1.2.1) werden erste einfache Eigenschaften behandelt, in 1.2.2 ihre Wirkung auf Unterräume. Dass lineare Abbildungen insbesondere Linearkombinationen (mit Hilfe derer ja lineare lineare Abhängigkeit und daran anschließende Begriffe definiert sind) in entsprechende Linearkombinationen überführen, hat wichtige Zusammenhänge zwischen den durch sie miteinander in Beziehung gesetzten involvierten Vektorräumen und ihren Dimensionen zur Folge (1.2.3). Jede lineare Abbildung ist durch ihre Werte auf einem Erzeugendensystem B_1 von V_1 eindeutig bestimmt. Ist B_1 sogar eine Basis, so gibt es zu jeder Vorgabe der Werte $f(\mathbf{b}) \in V_2$, $\mathbf{b} \in B_1$, genau eine lineare Abbildung $f : V_1 \rightarrow V_2$ mit diesen Werten (1.2.4). Die Menge aller linearen $f : V_1 \rightarrow V_2$ bildet in natürlicher Weise selbst wieder einen Vektorraum (1.2.5).

1.2.1 Der Begriff der linearen Abbildung

Inhalt in Kurzfassung: In Mathematik 1 wurden zwei wichtige Eigenschaften von Drehungen der Ebene um den Ursprung beobachtet (und für die Ableitung der Additionstheoreme für Cosinus und Sinus verwendet), nämlich ihre Verträglichkeit mit Addition (Parallelogramme gehen wieder in Parallelogramme über) und mit der Streckung von Vektoren. Genau diese Eigenschaften nehmen wir zum Anlass für die Definition linearer Abbildungen $f : V_1 \rightarrow V_2$ zwischen Vektorräumen V_1 und V_2 . Bilder und Urbilder von Unterräumen unter einer linearen Abbildung f sind wieder Unterräume. Das Urbild des Nullvektors nennt man den Kern von f . Verkettungen und Umkehrungen (sofern vorhanden) linearer Abbildungen sind wieder linear.

Definition 1.2.1.1. Sind V_1 und V_2 (reelle) Vektorräume, so heißt eine Abbildung $f : V_1 \rightarrow V_2$ **linear**, wenn für alle $\mathbf{x}, \mathbf{y} \in V_1$ und alle $r \in \mathbb{R}$ die beiden Gleichungen

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \quad \text{und} \quad f(r\mathbf{x}) = rf(\mathbf{x})$$

gelten. (Im Falle von Vektorräumen über einem anderen Körper als \mathbb{R} , etwa \mathbb{Q} oder \mathbb{C} , ist r aus diesem zu wählen.) Ist $V_1 = V_2 = V$ heißt f auch eine **lineare Transformation** von V .

1 Lineare Algebra

Die Menge aller linearen Abbildungen $f : V_1 \rightarrow V_2$ bezeichnen wir mit $L(V_1, V_2)$, im Falle $V_1 = V_2 = V$ auch mit $L(V)$.

Ein sehr einfaches Beispiel einer linearen Abbildung ist die identische Abbildung $\text{id}_V : V \rightarrow V, \mathbf{x} \mapsto \mathbf{x}$ auf einem Vektorraum V . Weniger triviale Beispiele sind die bereits erwähnten Drehungen, aber auch Projektionen oder, allgemeiner gewisse (eben lineare) Verzerrungen. Letztere kann man sich beispielsweise in der Ebene so vorstellen, dass das von den beiden Einheitsvektoren \mathbf{e}_1 und \mathbf{e}_2 erzeugte Gitter (Schachbrettmuster) in jenes schiefwinkelige Gitter übergeht, das von den Bildvektoren $f(\mathbf{e}_1)$ von \mathbf{e}_1 und $f(\mathbf{e}_2)$ von \mathbf{e}_2 unter der linearen Abbildung erzeugt wird. Wir kennen aber auch schon lineare Abbildungen, die sich nicht so leicht geometrisch veranschaulichen lassen. Die Differentiation $f \mapsto f'$, definiert auf dem (unendlichdimensionalen) Vektorraum V aller auf \mathbb{R} (oder einer geeigneten Teilmenge $D \subseteq \mathbb{R}$) definierten reellen Funktionen f ist wegen der Differentiationsregeln $(f + g)' = f' + g'$ und $(cf)' = cf'$ ($f, g \in V, c \in \mathbb{R}$) linear. Analoges gilt wegen $\int_a^b (f+g) = \int_a^b f + \int_a^b g$ und $\int_a^b (cf) = c \int_a^b f$ auch für bestimmte Integrale. Achtung: Bei unbestimmten Integralen liegen die Dinge wegen der additiven Konstanten etwas komplizierter.

Übungsaufgabe 22. (E) Überlegen Sie sich eine Möglichkeit, die Vektorräume V_1 und V_2 so zu definieren, dass die Bildung der Stammfunktion $f \mapsto \int f$ (unbestimmte Integration) als lineare Abbildung von V_1 nach V_2 aufgefasst werden kann. Anleitung: Die Elemente von V_2 müssen statt Funktionen f Mengen von Funktionen der Gestalt $f + c$ mit $c \in \mathbb{R}$ sein. Eine Alternative Möglichkeit besteht darin, sich auf Stammfunktionen F mit der zusätzlichen Eigenschaft $F(0) = 0$ zu beschränken.

Die Verknüpfung (Hintereinanderausführung, Verkettung) linearer Abbildungen ist wieder linear:

Proposition 1.2.1.2. Sind V_1, V_2, V_3 Vektorräume und $f : V_1 \rightarrow V_2$ und $g : V_2 \rightarrow V_3$ lineare Abbildungen, so ist auch ihre Verkettung $g \circ f : V_1 \rightarrow V_3, \mathbf{x} \mapsto g(f(\mathbf{x}))$ linear.

Beweis. Es gilt sowohl

$$\begin{aligned}(g \circ f)(\mathbf{x} + \mathbf{y}) &= g(f(\mathbf{x} + \mathbf{y})) = g(f(\mathbf{x}) + f(\mathbf{y})) = g(f(\mathbf{x})) + g(f(\mathbf{y})) = \\ &= (g \circ f)(\mathbf{x}) + (g \circ f)(\mathbf{y})\end{aligned}$$

als auch

$$(g \circ f)(r\mathbf{x}) = g(f(r\mathbf{x})) = g(rf(\mathbf{x})) = rg(f(\mathbf{x})) = r(g \circ f)(\mathbf{x})$$

für alle $\mathbf{x}, \mathbf{y} \in V_1$ und $r \in \mathbb{R}$. □

Auch Umkehrabbildungen injektiver linearer Abbildungen sind linear. Genauer:

Proposition 1.2.1.3. Ist die Abbildung $f : V_1 \rightarrow V_2$ zwischen den Vektorräumen V_1 und V_2 linear und injektiv, also bijektiv als Abbildung $f : V_1 \rightarrow f(V_1)$, so ist auch die Umkehrabbildung $f^{(-1)} : f(V_1) \rightarrow V_1$ linear.

Beweis. Wegen Proposition 1.2.2.1 ist $f(V_1)$ ein Unterraum von V_2 , also selbst ein Vektorraum. Somit ist es überhaupt sinnvoll von Linearität der Umkehrabbildung $f^{(-1)} : f(V_1) \rightarrow V_1$ zu sprechen. Seien irgendwelche $\mathbf{x}, \mathbf{y} \in f(V_1)$ gegeben. Dann gibt es $\mathbf{x}', \mathbf{y}' \in V_1$ mit $f(\mathbf{x}') = \mathbf{x}$ und $f(\mathbf{y}') = \mathbf{y}$. Weil f linear ist, gilt $f(\mathbf{x}' + \mathbf{y}') = f(\mathbf{x}') + f(\mathbf{y}') = \mathbf{x} + \mathbf{y}$, folglich

$$f^{(-1)}(\mathbf{x} + \mathbf{y}) = f^{(-1)}(f(\mathbf{x}' + \mathbf{y}')) = (f^{(-1)} \circ f)(\mathbf{x}' + \mathbf{y}') = \mathbf{x}' + \mathbf{y}' = f^{(-1)}(\mathbf{x}) + f^{(-1)}(\mathbf{y}),$$

die erste Linearitätseigenschaft für $f^{(-1)}$. Die zweite, nämlich $f^{(-1)}(r\mathbf{x}) = rf^{(-1)}(\mathbf{x})$ für alle $r \in \mathbb{R}$ sieht man ganz ähnlich ein (Übung). \square

Übungsaufgabe 23. (P) Beweisen Sie die im Beweis von Proposition 1.2.1.3 verwendete Beziehung $f^{(-1)}(r\mathbf{x}) = rf^{(-1)}(\mathbf{x})$ für alle $r \in \mathbb{R}$.

Äquivalent zu Definition 1.2.1.1 könnte man lineare Abbildungen auch dadurch definieren, dass sie beliebige LK erhalten, genauer: Für Vektorräume V_1, V_2 ist eine Abbildung $f : V_1 \rightarrow V_2$ genau dann linear, wenn für alle $n \in \mathbb{N}$, alle Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ und alle Zahlen $r_1, \dots, r_n \in \mathbb{R}$ gilt:

$$f\left(\sum_{i=1}^n r_i \mathbf{x}_i\right) = \sum_{i=1}^n r_i f(\mathbf{x}_i)$$

Übungsaufgabe 24. (E) Beweisen Sie, dass diese Bedingung tatsächlich äquivalent ist zur Definition 1.2.1.1. Anleitung: Es sind zwei Implikationen zu zeigen. Um von der Bedingung mit den Linearkombinationen auf die beiden Bedingungen in Definition 1.2.1.1 zu schließen, hat man lediglich zweimal zu spezialisieren. Für die andere Richtung gehe man mittels Induktion nach n vor. Für den Induktionsanfang $n = 0$ verwende man, dass die leere Summe definitionsgemäß den Nullvektor darstellt, und Übungsaufgabe 26.

Übungsaufgabe 25. (T) Welche der folgenden Abbildungen sind \mathbb{R} -linear?

1. $\mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x + y, y)$
2. $\mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b$ für $a, b \in \mathbb{R}$ fest gewählt
3. $\mathbb{R}^3 \rightarrow \mathbb{R}^2, (x, y, z) \mapsto (x - z, y - x)$
4. $\mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z) \mapsto |x| + |y| + |z|$
5. $\mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$
6. $C[0, 1] \rightarrow C[0, 1], f(x) \mapsto \int f(x) dx$
7. $C[0, 1] \rightarrow \mathbb{R}, f(x) \mapsto \int_0^1 f(x) dx$

1.2.2 Linearität und Unterräume

Inhalt in Kurzfassung: Linearität von Abbildungen verträgt sich sehr gut mit den Strukturelementen von Vektorräumen, die wir bereits untersucht haben: Sie bilden den Nullvektor wieder auf den Nullvektor ab, Unterräume auf Unterräume, und auch die Urbilder von Unterräumen unter linearen Abbildungen sind wieder Unterräume.

Für eine lineare Abbildung $f : V_1 \rightarrow V_2$ mit Nullvektoren $\mathbf{o}_i \in V_i$ für $i = 1, 2$ (diese und ähnliche selbsterklärende Schreibweisen werden wir häufig verwenden, auch ohne sie jedes Mal ausdrücklich zu erklären) gilt stets $f(\mathbf{o}_1) = \mathbf{o}_2$.

Übungsaufgabe 26. (E) Beweisen Sie die behauptete Beziehung $f(\mathbf{o}_1) = \mathbf{o}_2$. Anleitung: Von den beiden Gleichungen in Definition 1.2.1.1 müssen sie nur eine verwenden, welche können Sie sich aussuchen.

Eine lineare Abbildung ist verträglich nicht nur, wie wir in Übungsaufgabe 24 gesehen haben, mit LKen, sondern auch mit Unterräumen in dem Sinn, dass Bilder und Urbilder von Unterräumen unter linearen Abbildungen selbst Unterräume sind:

Proposition 1.2.2.1. Sind V_1, V_2 Vektorräume, $f : V_1 \rightarrow V_2$ linear sowie $U_1 \leq V_1$ und $U_2 \leq V_2$ Unterräume, dann gilt:

1. Das Bild $f(U_1) = \{f(\mathbf{a}) : \mathbf{a} \in U_1\}$ von U_1 unter f ist ein Unterraum von V_2 .
2. Das Urbild $f^{-1}(U_2) = \{\mathbf{a} \in V_1 : f(\mathbf{a}) \in U_2\}$ von U_2 unter f ist ein Unterraum von V_1 .

Beweis. Für die erste Behauptung ist von Elementen in $f(U_1)$, also solchen der Form $f(\mathbf{a}), f(\mathbf{b})$ mit $\mathbf{a}, \mathbf{b} \in U_1$ auszugehen und zu zeigen, dass auch Summe $f(\mathbf{a}) + f(\mathbf{b})$ und jedes Vielfache $rf(\mathbf{a})$ mit $r \in \mathbb{R}$ wieder in $f(U_1)$ liegen. Wegen der Linearität von f gilt $f(\mathbf{a}) + f(\mathbf{b}) = f(\mathbf{a} + \mathbf{b})$ und $rf(\mathbf{a}) = f(r\mathbf{a})$. Weil U_1 ein Unterraum ist, liegen mit \mathbf{a}, \mathbf{b} auch $\mathbf{a} + \mathbf{b}$ und $r\mathbf{a}$ in U_1 . Also folgt tatsächlich $f(\mathbf{a}) + f(\mathbf{b}) = f(\mathbf{a} + \mathbf{b}) \in f(U_1)$ und $rf(\mathbf{a}) = f(r\mathbf{a}) \in f(U_1)$.

Der Beweis der zweiten Behauptung erfolgt ganz ähnlich, lediglich ist von zwei Vektoren $\mathbf{a}, \mathbf{b} \in f^{-1}(U_2)$, d.h. mit $f(\mathbf{a}), f(\mathbf{b}) \in U_2$ auszugehen. Die Linearität von f zeigt wieder $f(\mathbf{a} + \mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b})$ und $f(r\mathbf{a}) = rf(\mathbf{a})$. Wegen der Unterraumeigenschaft von U_2 liegen beide Elemente in U_2 , also ist auch $\mathbf{a} + \mathbf{b} \in f^{-1}(U_2)$ und $r\mathbf{a} \in f^{-1}(U_2)$ gezeigt. \square

Besondere Hervorhebung verdienen die Fälle $U_1 = V_1$ und $U_2 = \{\mathbf{o}_2\}$. Nach 1.2.2.1 sind $f(V_1)$ und $f^{-1}(\{\mathbf{o}_2\})$ Unterräume von V_2 bzw. von V_1 . Man schreibt für das **Bild** $f(V_1)$ der linearen Abbildung f symbolisch auch $\text{Im}(f)$.¹ Das Urbild $f^{-1}(\{\mathbf{o}_2\})$ des Nullraums in V_2 nennt man **Kern der linearen Abbildung** f , symbolisch $\ker f := f^{-1}(\{\mathbf{o}_2\})$. Nützlich ist die folgende Beobachtung:

Proposition 1.2.2.2. Eine lineare Abbildung $f : V_1 \rightarrow V_2$ zwischen den Vektorräumen V_1 und V_2 ist genau dann injektiv, wenn $\ker f = \{\mathbf{o}_1\}$ (mit dem Nullvektor \mathbf{o}_1 von V_1).

Beweis. Sei zunächst f injektiv. Für die behauptete Mengengleichheit $\ker f = \{\mathbf{o}\}$ sind zwei Inklusionen zu zeigen. Weil $\ker f$, wie soeben bemerkt, ein Unterraum von V_1 ist, muss er insbesondere den Nullvektor $\mathbf{o}_1 \in V_1$ enthalten, also $\{\mathbf{o}_1\} \subseteq \ker f$. Es gilt aber auch die umgekehrte Inklusion. Zum Nachweis sei $\mathbf{x} \in \ker f$, also $f(\mathbf{x}) = \mathbf{o}_2$ (Nullvektor von V_2). Gleichzeitig gilt wegen der Linearität von f aber auch $f(\mathbf{o}_1) = \mathbf{o}_2$ (siehe Übungsaufgabe 26). Die vorausgesetzte Injektivität von f erzwingt deshalb $\mathbf{x} = \mathbf{o}_1$, womit auch $\ker f \subseteq \{\mathbf{o}_1\}$ bewiesen ist.

Sei nun $\ker f = \{\mathbf{o}_1\}$ vorausgesetzt. Um daraus die Injektivität von f zu folgern, haben wir aus $f(\mathbf{x}) = f(\mathbf{y})$ für irgendwelche $\mathbf{x}, \mathbf{y} \in V_1$ die Gleichheit $\mathbf{x} = \mathbf{y}$ abzuleiten.

¹Im für englisch *image*.

In der Tat erzwingt die Linearität von f , dass $f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = \mathbf{o}_2$ gilt, somit $\mathbf{x} - \mathbf{y} \in \ker f = \{\mathbf{o}_1\}$, folglich $\mathbf{x} - \mathbf{y} = \mathbf{o}_1$ oder $\mathbf{x} = \mathbf{y}$. \square

1.2.3 Lineare Abbildungen und Dimension

Inhalt in Kurzfassung: Lineare Abbildungen übertragen Linearkombinationen und somit lineare Abhängigkeiten. Das hat zur Folge, dass die Dimension des Bildes einer linearen Abbildung nie die des Ausgangsraumes übersteigen kann – analog wie unter beliebigen Abbildungen Kardinalitäten nicht größer, sondern nur kleiner werden können. Die Begriffe Rang und Defekt ermöglichen eine sehr konzise Formulierung der Zusammenhänge in der Rangformel.

Die Linearität einer Abbildung $f : V_1 \rightarrow V_2$ zwischen Vektorräumen V_1 und V_2 zieht gewisse Beziehungen zwischen den Dimensionen nach sich. Denn jede lineare Abhängigkeit $\sum_{i=1}^n r_i \mathbf{a}_i = \mathbf{o}$ zwischen Vektoren $\mathbf{a}_i \in V_1$ mittels Koeffizienten r_i geht durch f über in die entsprechende lineare Abhängigkeit

$$\sum_{i=1}^n r_i f(\mathbf{a}_i) = f\left(\sum_{i=1}^n r_i \mathbf{a}_i\right) = f(\mathbf{o}) = \mathbf{o}$$

zwischen den $f(\mathbf{a}_i) \in V_2$. (Die Linearität von f wurde für die erste Gleichheit verwendet.) Sind die Vektoren \mathbf{a}_i linear abhängig, so sind es folglich auch ihre Bilder $f(\mathbf{a}_i)$. Mit anderen Worten: Im Bild $f(V_1)$ einer linearen Abbildung kann es nicht mehr linear unabhängige Vektoren geben als in V_1 , die Dimension kann bei der Anwendung einer linearen Abbildung also nicht zunehmen. Zur Präzisierung dieser Zusammenhänge definieren wir:

Definition 1.2.3.1. Seien V_1 und V_2 Vektorräume und $f : V_1 \rightarrow V_2$ linear. Dann heißt die Dimension $\dim f(V_1)$ des Bildes von f **Rang** von f , symbolisch $\operatorname{rg} f := \dim \operatorname{Im} f$, die Dimension $\dim f^{-1}(\{\mathbf{o}\})$ des Kerns von f **Defekt** von f , symbolisch $\operatorname{def} f := \dim \ker f$. Ist $V_1 = V_2 = V$ und $f : V \rightarrow V$ bijektiv, so nennt man f auch **regulär**.

Damit können wir das Ergebnis unserer Überlegung zur Ungleichung $\operatorname{rg} f \leq \dim V_1$ zusammenfassen. (Man beachte die Analogie zur Formel $|f(A)| \leq |A|$, die für beliebige Abbildungen gilt.) Noch genauere Auskunft gibt die **Rangformel**:

Satz 1.2.3.2. Seien V_1, V_2 Vektorräume und $f : V_1 \rightarrow V_2$ linear. Dann gilt

$$\dim V_1 = \operatorname{def} f + \operatorname{rg} f.$$

Beweis. Sei B_0 eine Basis von $\ker f$, also $|B_0| = \operatorname{def} f$. Wegen Satz 1.1.5.3 gibt es eine Basis B von V_1 , die B_0 enthält. (Wir haben diesen Satz nur für endlichdimensionales V_1 bewiesen, er gilt aber allgemein, siehe auch Satz 1.1.5.4.) Sei $B_1 := B \setminus B_0$. Dann gilt $\dim V_1 = |B| = |B_0 \cup B_1| = |B_0| + |B_1| = \operatorname{def} f + |B_1|$. Der Beweis ist fertig, wenn wir $|B_1| = \operatorname{rg} f$ zeigen können. Das wiederum folgt, wenn die $f(\mathbf{b})$ mit $\mathbf{b} \in B_1$ eine Basis von $\operatorname{Im} f$ bilden. Dafür sind zwei Behauptungen zu zeigen, nämlich dass diese Vektoren ganz $\operatorname{Im} f$ erzeugen und dass sie linear unabhängig sind.

Die $f(\mathbf{b})$, $\mathbf{b} \in B_1$, erzeugen $\text{Im } f$: Jeder Vektor aus $\text{Im } f$ hat die Gestalt $f(\mathbf{a})$ mit einem $\mathbf{a} \in V_1$. Weil B eine Basis ist, gibt es eine Darstellung von \mathbf{a} als LK $\mathbf{a} = \sum_{i=1}^n r_i \mathbf{b}_i$ mit geeigneten \mathbf{b}_i , wobei wir die \mathbf{b}_i so nummerieren, dass die ersten k davon zu B_1 gehören, die restlichen $n - k$ zu B_0 . Dann ist $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_0$, wobei \mathbf{a}_1 eine LK ausschließlich von Basisvektoren aus B_1 ist und \mathbf{a}_0 eine von Basisvektoren aus B_0 . Insbesondere ist $\mathbf{a}_0 \in \ker f$, also $f(\mathbf{a}_0) = \mathbf{o}$, woraus

$$f(\mathbf{a}) = f(\mathbf{a}_1) + f(\mathbf{a}_0) = f(\mathbf{a}_1) = \sum_{i=1}^k r_i f(\mathbf{b}_i)$$

folgt, Das ist eine Darstellung von $f(\mathbf{a})$ als LK von gewissen $f(\mathbf{b})$ mit $\mathbf{b} \in B_1$, wie behauptet.

Die $f(\mathbf{b})$, $\mathbf{b} \in B_1$, sind linear unabhängig: Angenommen, die $f(\mathbf{b})$, $\mathbf{b} \in B_1$, wären linear abhängig. Dann gäbe es gewisse paarweise verschiedene $\mathbf{b}_i \in B_1$ und Skalare $r_i \in \mathbb{R} \setminus \{0\}$ mit $\sum_{i=1}^k r_i f(\mathbf{b}_i) = \mathbf{o}$. Wegen der Linearität von f bedeutet das

$$f\left(\sum_{i=1}^k r_i \mathbf{b}_i\right) = \sum_{i=1}^k r_i f(\mathbf{b}_i) = \mathbf{o},$$

also $\mathbf{a} := \sum_{i=1}^k r_i \mathbf{b}_i \in \ker f$. Weil B_0 eine Basis von $\ker f$ ist, hat jeder Vektor aus $\ker f$ eine Darstellung als LK der $\mathbf{b} \in B_0$; weil $B = B_0 \cup B_1$ eine Basis von V_1 ist, ist so eine Darstellung auch unter Einbeziehung der $\mathbf{b} \in B_1$ eindeutig. Für die Koeffizienten r_i von \mathbf{a} bezüglich der $\mathbf{b}_i \in B_1$ folgt daraus $r_i = 0$, was die behauptete lineare Unabhängigkeit der $f(\mathbf{b})$, $\mathbf{b} \in B_1$, beweist. \square

1.2.4 Die Festlegung einer linearen Abbildung auf einer Basis

Inhalt in Kurzfassung: Die große Bedeutung von Basen liegt zu einem guten Teil daran, dass jede auf ihnen definierte Abbildung in einen anderen Vektorraum ein eindeutige lineare Fortsetzung hat. Dabei liegt die Eindeutigkeit daran, dass jede Basis ein Erzeugendensystem ist. Die Existenz folgt aus der linearen Unabhängigkeit, weil die Linearitätsbedingung für LKs dann zu keinen Widersprüchen führen kann. Im Abschnitt 1.3 wird uns diese Eigenschaft von Basen zur Darstellung linearer Abbildungen durch Matrizen führen. Für die Theorie werden wir schon unmittelbar die Konsequenz ziehen, dass zwei Vektorräume genau dann isomorph (strukturgleich) sind, wenn sie dieselbe Dimension haben.

Das Zusammenspiel zwischen linearen Abbildungen einerseits und Erzeugendensystemen, linear unabhängigen Mengen bzw. Basen andererseits wird durch den sogenannten **Fortsetzungssatz** beschrieben.

Satz 1.2.4.1. *Seien V_1, V_2 Vektorräume und $A \subseteq V_1$ eine Teilmenge von V_1 .*

1. *Sind $f, g : V_1 \rightarrow V_2$ lineare Abbildungen, A ein Erzeugendensystem von V_1 und $f(\mathbf{a}) = g(\mathbf{a})$ für alle $\mathbf{a} \in A$, so folgt $f = g$. (Mit anderen Worten: Lineare Abbildungen sind durch ihre Werte auf einem Erzeugendensystem eindeutig bestimmt.)*

2. Ist A eine Basis von V_1 , so lässt sich jede Abbildung $f_0 : A \rightarrow V_2$, in eindeutiger Weise zu einer linearen Abbildung $f : V_1 \rightarrow V_2$ fortsetzen.

Beweis. 1. Weil A ein Erzeugendensystem von V_1 ist, lässt sich ein beliebiger Vektor $\mathbf{x} \in V_1$ als LK $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{a}_i$ mit geeigneten $r_i \in \mathbb{R}$ und $\mathbf{a}_i \in A$ schreiben. Wegen der Linearität von f und g folgt daraus

$$f(\mathbf{x}) = f\left(\sum_{i=1}^n r_i \mathbf{a}_i\right) = \sum_{i=1}^n r_i f(\mathbf{a}_i) = \sum_{i=1}^n r_i g(\mathbf{a}_i) = g\left(\sum_{i=1}^n r_i \mathbf{a}_i\right) = g(\mathbf{x}).$$

2. Wegen Teil 1 genügt es, ein lineares $f : V_1 \rightarrow V_2$ mit $f(\mathbf{b}) = f_0(\mathbf{b})$ für alle $\mathbf{b} \in A$ zu konstruieren. Nach Satz 1.1.5.4 lässt sich A zu einer Basis B erweitern. Zunächst setzen wir f_0 fort zu einer Abbildung $f_1 : B \rightarrow V_2$, indem wir $f_1(\mathbf{a}) := f_0(\mathbf{a})$ für alle $\mathbf{a} \in A$ und $f_1(\mathbf{b}) := \mathbf{o}$ für alle $\mathbf{b} \in B \setminus A$ setzen. Jeder Vektor $\mathbf{x} \in V_1$ hat eine Darstellung als LK $\mathbf{x} = \sum_{i=1}^n r_i \mathbf{b}_i$ mit geeigneten $r_i \in \mathbb{R}$ und paarweise verschiedenen $\mathbf{b}_i \in B$. Deshalb ist auch die Festsetzung

$$f(\mathbf{x}) := \sum_{i=1}^n r_i f_1(\mathbf{b}_i)$$

eindeutig. Für $\mathbf{b} \in A$ kann $n = 1$, $r_1 = 1$ und $\mathbf{b}_1 = \mathbf{b}$ gewählt werden, folglich ist $f(\mathbf{b}) = f_1(\mathbf{b}) = f_0(\mathbf{b})$. Zum Nachweis der Linearität ist $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ und $f(r\mathbf{x}) = rf(\mathbf{x})$ für alle $\mathbf{x}, \mathbf{y} \in V_1$ und $r \in \mathbb{R}$ zu zeigen, was man leicht als Übungsaufgabe nachrechnet. □

Übungsaufgabe 27. (E) Rechnen Sie die Linearität der Abbildung f in Teil 2 von Satz 1.2.4.13 nach.

Übungsaufgabe 28. (T^*) Sei P ein Punkt im \mathbb{R}^2 mit Vektordarstellung $\begin{pmatrix} x \\ y \end{pmatrix}$. Betrachten Sie die lineare Transformation f welche P auf den Punkt P' mit den Koordinaten $x' = ax + by$ und $y' = bx - ay$ abbildet (dabei seien a, b reelle Parameter, die die Bedingung $a^2 + b^2 = 1$ erfüllen).

Berechnen Sie für die Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ das Bild unter der Transformation $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Zeigen Sie weiters, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ eine Basis von \mathbb{R}^2 bilden und entscheiden Sie, ob auch $f\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und $f\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ eine Basis bilden. Ist f injektiv? (Anmerkung: Wir werden in 1.3 systematische Methoden zur Behandlung derartiger Aufgaben kennen lernen. Zur Vorbereitung sollen Sie sich schon hier mit den bisherigen Mitteln anhand dieser noch recht einfachen Aufgabe selbständig eine mögliche Vorgangsweise überlegen.)

Haben zwei Vektorräume V_1, V_2 dieselbe Dimension $n = \dim V_1 = \dim V_2$, so gibt es Basen $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ von V_1 und $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ von V_2 . Die zweite Aussage von Satz 1.2.4.1 garantiert, dass es genau eine lineare Abbildung $f : V_1 \rightarrow V_2$ mit $f(\mathbf{a}_i) = \mathbf{b}_i$ für $i = 1, \dots, n$ gibt. Weil $f(V_1)$ ein Unterraum von V_2 ist, der alle Basisvektoren $\mathbf{b}_1 = f(\mathbf{a}_1), \dots, \mathbf{b}_n = f(\mathbf{a}_n)$ enthält, muss dieser ganz V_2 sein, also $\text{Im } f = V_2$ und somit

$\operatorname{rg} f = \dim \operatorname{Im} f = \dim V_2 = n$. Der einzige Unterraum von V_2 der Dimension n ist V_2 selbst, also $f(V_1) = V_2$. Das bedeutet, dass f surjektiv ist. Die Rangformel (1.2.3.2) liefert $n = \dim V_1 = \operatorname{def} f + \operatorname{rg} f = \operatorname{def} f + n$, woraus $\dim \ker f = \operatorname{def} f = 0$ folgt. Der einzige Unterraum der Dimension 0 ist der Nullraum, also $\ker f = \{\mathbf{o}\}$, was wegen Proposition 1.2.2.2 die Injektivität von f bedeutet. Also ist f sogar bijektiv. Es gibt somit eine Umkehrabbildung $f^{(-1)} : V_2 \rightarrow V_1$, die nach Proposition 1.2.1.3 ebenfalls linear ist. Man beachte die Analogie zu bijektiven Abbildungen ganz generell: Bijektive Abbildungen lassen Kardinalitäten unverändert, bijektive lineare Abbildungen auch Dimensionen.

Eine bijektive lineare Abbildung $f : V_1 \rightarrow V_2$ mit Umkehrabbildung $f^{(-1)} : V_2 \rightarrow V_1$ wie im obigen Absatz ermöglicht es, zwischen LKen in V_1 und V_2 nach Belieben hin und her zu wechseln. In diesem Sinn haben die Vektorräume V_1 und V_2 die gleiche Struktur. Man spricht deshalb von einem **Isomorphismus** f und nennt die beiden Vektorräume V_1 und V_2 **isomorph**, symbolisch $V_1 \cong V_2$. Damit haben wir im vorangegangenen Absatz für den endlichdimensionalen Fall bewiesen, dass zwei Vektorräume derselben Dimension isomorph sind. Ganz ähnliche Argumente funktionieren auch für unendlichdimensionale Räume. Umgekehrt haben zwei isomorphe Vektorräume V_1 und V_2 aber auch dieselbe Dimension, weil durch einen Isomorphismus $f : V_1 \rightarrow V_2$ jede Basis B von V_1 in eine gleich mächtige Basis $f(B)$ von V_2 übergeht (siehe Übung 29). Damit ist insgesamt bewiesen:

Satz 1.2.4.2. *Je zwei Vektorräume V_1 und V_2 sind genau dann isomorph, wenn $\dim V_1 = \dim V_2$. Insbesondere ist jeder n -dimensionale (reelle) Vektorraum V isomorph zu \mathbb{R}^n . (Analoges gilt auch für Vektorräume über anderen Körpern.)*

Für jede bijektive Abbildung $f_0 : B_1 \rightarrow B_2$ zwischen einer Basis B_1 von V_1 und einer Basis B_2 von V_2 ist ein Isomorphismus durch die (laut Fortsetzungssatz 1.2.4.1 eindeutige) lineare Fortsetzung $f : V_1 \rightarrow V_2$ von f_0 gegeben, die durch

$$f \left(\sum_{i=1}^n r_i \mathbf{b}_i \right) = \sum_{i=1}^n r_i f_0(\mathbf{b}_i)$$

für alle $\mathbf{b}_i \in B_1$ und $r_i \in \mathbb{R}$ charakterisiert ist.

Dieser Satz lässt sich auch so verstehen: Über die Struktur eines Vektorraums wissen wir alles, wenn wir seine Dimension kennen. Bei Fragen über (reelle) Vektorräume genügt es deshalb oft, die Situation im Raum \mathbb{R}^n zu studieren, weil die Antwort dann auf jeden zu \mathbb{R}^n isomorphen Vektorraum, also auf jeden (reellen) Vektorraum der Dimension n übertragen werden kann.

Übungsaufgabe 29. (E) Zeigen Sie folgende weiter oben verwendete Behauptung: Ist $f : V_1 \rightarrow V_2$ ein Isomorphismus zwischen den Vektorräumen V_1 und V_2 und $B \subseteq V_1$ eine Basis von V_1 , dann ist $f(B)$ eine Basis von V_2 .

1.2.5 Vektorräume linearer Abbildungen

Inhalt in Kurzfassung: Ist V ein Vektorraum und X irgendeine Menge, so bildet auch die Menge $A(X, V)$ aller Abbildungen $f : X \rightarrow V$ in natürlicher Weise einen Vektorraum der

Dimension $|X| \cdot \dim V$. Ist X selbst ein Vektorraum, so bilden die linearen Abbildungen einen Unterraum $L(X, V)$ von $A(X, V)$.

Für Abbildungen $f, g : X \rightarrow V$ von einer beliebigen Menge X in einen Vektorraum V (man beachte auch den Spezialfall $V = \mathbb{R} = \mathbb{R}^1$) lassen sich die Abbildungen $f + g$ und rf in natürlicher Weise durch

$$(f + g)(\mathbf{x}) := f(\mathbf{x}) + g(\mathbf{x}) \quad \text{und} \quad (rf)(\mathbf{x}) := rf(\mathbf{x})$$

definieren. Unmittelbar rechnet die Gesetze einer abelschen Gruppe nach: $(f + g) + h = f + (g + h)$, $f + g = g + f$, $f + \mathbf{o}_{A(X, V)} = f$ mit der Nullfunktion $\mathbf{o}_{A(X, V)}$ auf X und $f + (-f) = \mathbf{o}_{A(X, V)}$ für $(-f)(\mathbf{x}) := -f(\mathbf{x})$. Gleiches gilt für die Vektorraumgesetze:

$$r(f + g) = rf + rg, \quad (r + s)f = rf + sf, \quad (rs)f = r(sf) \quad \text{und} \quad 1f = f$$

Die Menge $A(X, V)$ aller $f : X \rightarrow V$ bildet mit dieser Addition und Skalarmultiplikation daher selbst einen Vektorraum.

Übungsaufgabe 30. (P) Überprüfen Sie die Vektorraumgesetze auf $A(X, V)$ im Detail. Machen Sie insbesondere deutlich, was genau gezeigt werden muss und was Sie in Ihrer Rechnung wo verwenden.

Ist X endlich, so ist $\dim A(X, V) = |X| \cdot \dim V$: Ist nämlich B eine Basis von V , so können wir für jedes Paar $(x, \mathbf{b}) \in X \times B$ ein $f_{(x, \mathbf{b})} \in A(X, V)$ definieren, indem wir $f_{(x, \mathbf{b})}(x) = \mathbf{b}$ setzen und $f_{(x, \mathbf{b})}(x') = \mathbf{o}$ für alle $x' \neq x$ aus X . Dann bilden (Übung) sämtliche $f_{(x, \mathbf{b})}$, $(x, \mathbf{b}) \in X \times B$ eine Basis von $A(X, V)$, und wie behauptet ist ihre Anzahl $\dim A(X, V) = |X| \cdot \dim V$. Übrigens ist uns ein Spezialfall bereits bekannt: Für $V = \mathbb{R}$ (eindimensional) und $X = \{1, 2, \dots, n\}$, so ist jede Abbildung $f : X \rightarrow V = \mathbb{R}$ durch ihre Werte gegeben, die wir als $x_1 = f(1), x_2 = f(2), \dots, x_n = f(n)$ schreiben und zu einem n -Tupel $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ zusammenfassen können. Also ist $A(X, V) = A(\{1, \dots, n\}, \mathbb{R})$ im Wesentlichen der n -dimensionale Standardvektorraum \mathbb{R}^n .

Besonders interessieren wir uns für die Situation, dass auch die Menge X ein Vektorraum ist, und für lineare Abbildungen f . Wir betrachten also die Menge $A(V_1, V_2)$ und darin die Teilmenge $L(V_1, V_2)$ aller linearen Abbildungen $f : V_1 \rightarrow V_2$ zwischen zwei Vektorräumen V_1 und V_2 . Für $f, g \in L(V_1, V_2)$ und $r \in \mathbb{R}$ sind, wie man sofort nachrechnet (Übung), auch $f + g$ und rf linear, folglich auch beliebige LKen linearer Abbildungen.

Übungsaufgabe 31. (P) Prüfen Sie die Linearität von $f + g$ und rf ausführlich nach. Geben Sie in jedem Schritt sorgfältig an, was Sie dabei verwenden. Schließen Sie daraus, dass auch beliebige LKen $\sum_{i=1}^n r_i f_i$ linearer Abbildungen f_i linear sind. (Hinweis: Betrachten Sie zunächst $n = 2$ Summanden und gehen Sie dann mittels Induktion nach n vor.)

Das zeigt, dass $L(V_1, V_2)$ ein Unterraum von $A(V_1, V_2)$, also selbst ein Vektorraum ist. Zur deutlichen Unterscheidung schreiben wir \mathbf{o}_1 und \mathbf{o}_2 für die Nullvektoren in V_1 bzw. V_2 . Der Nullvektor $\mathbf{o} = \mathbf{o}_{L(V_1, V_2)}$ in $L(V_1, V_2)$ ist die konstante Abbildung $\mathbf{x} \mapsto \mathbf{o}_2$ (die

natürlich selbst linear ist), das additive Inverse $-f$ einer linearen Abbildung $f : V_1 \rightarrow V_2$ ist gegeben durch die gleichfalls lineare Abbildung $-f = (-1)f : \mathbf{x} \mapsto -f(\mathbf{x})$. Wir fassen zusammen:

Proposition 1.2.5.1. *Für eine beliebige Menge X und einen beliebigen Vektorraum V bildet die Menge $A(X, V)$ aller Abbildungen $f : X \rightarrow V$ bezüglich der punktweise definierten Operationen einen Vektorraum. Ist $X = V_1$ ebenfalls ein Vektorraum und $V = V_2$, so bildet die Teilmenge $L(V_1, V_2)$ aller linearen $f \in A(V_1, V_2)$ einen Unterraum, symbolisch: $L(V_1, V_2) \leq A(V_1, V_2)$.*

Um auch die Dimension $d := \dim L(V_1, V_2)$ zu bestimmen, erweist sich die schon mehrmals angekündigte Matrizendarstellung linearer Abbildungen (siehe 1.3) als hilfreich. Zuvor widmen wir uns aber der geometrischen Deutung linearer Abbildungen.

1.2.6 Beispiele linearer Abbildungen und ihre geometrische Deutung

Inhalt in Kurzfassung: Beispiele linearer Abbildungen in den Standardvektorräumen, gegliedert nach der Dimension.

Zur besseren Veranschaulichung des abstrakten Begriffs der linearen Abbildung lohnt es, sich einige Beispiele anhand der Standardvektorräume in verschiedenen Dimensionen vor Augen zu halten. Sei also $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ linear. Die Nullvektoren seien mit $\mathbf{o}_m \in \mathbb{R}^m$ bzw. $\mathbf{o}_n \in \mathbb{R}^n$ bezeichnet. Kommen Basen in Spiel, so wählen wir die kanonischen.

$m = 0$ oder $n = 0$: In diesem Fall ist $\mathbb{R}^m = \{\mathbf{o}_m\}$ oder $\mathbb{R}^n = \{\mathbf{o}_n\}$, und es gibt nur eine einzige lineare Abbildung, nämlich die konstante Nullabbildung $f : \mathbf{x} \mapsto \mathbf{o}_n$ für alle $\mathbf{x} \in \mathbb{R}^m$.

$m = n = 1$: Eine lineare Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ im Sinn der Linearen Algebra ist eindeutig bestimmt durch den Wert $r := f(1)$ auf dem kanonischen Basisvektor $\mathbf{e}_1 = 1 \in \mathbb{R}$. Dann haben wir es mit $f(x) := rx$, der homogenen linearen reellen Funktion mit Steigung r zu tun. Damit erklärt sich auch die Bezeichnung *linear*: Im eindimensionalen Fall handelt es sich bei linearen Funktionen genau um jene, deren Graph eine (gerade) *Linie* (durch den Koordinatenursprung) ist.

$m = 1, n$ beliebig: Wieder ist jedes lineare $f : \mathbb{R} \rightarrow \mathbb{R}^n$ eindeutig bestimmt durch den Wert $\mathbf{x} := f(1) \in \mathbb{R}^n$ an der Stelle 1, nämlich als Abbildung $r \mapsto r\mathbf{x}$.

m beliebig, $n = 1$, d.h. $f : \mathbb{R}^m \rightarrow \mathbb{R}$: In diesem Fall spricht man von **Linearformen**. Ist $f(\mathbf{e}_1) = r_1, \dots, f(\mathbf{e}_m) = r_m$, so hat f die Gestalt

$$f : \mathbb{R}^m \rightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto \sum_{j=1}^m r_j x_j.$$

Handelt es sich bei f nicht um die Nullabbildung (d.h. sind nicht alle r_j gleich 0), so ist $0 < \operatorname{rg} f \leq \dim \mathbb{R} = 1$, also $\operatorname{rg} f = 1$. Die Rangformel (1.2.3.2) besagt in diesem Fall $m = \dim \mathbb{R}^m = \operatorname{def} f + \operatorname{rg} f = \operatorname{def} f + 1$, woraus $\dim \ker f = \operatorname{def} f = m - 1$ folgt.

Für $m = 2$ lässt sich daraus eine sehr ansprechende geometrische Veranschaulichung ableiten. Denken wir uns den Definitionsbereich \mathbb{R}^2 in üblicher Weise mit Hilfe eines zweidimensionalen Koordinatensystems dargestellt und interpretieren wir zur Veranschaulichung gleichzeitig (zum Beispiel) die x -Achse als Zielmenge \mathbb{R} . Dann werden genau die \mathbf{x} aus dem Kern $\ker f$, der laut obiger Überlegung ein $m - 1 = 2 - 1 = 1$ -dimensionaler Unterraum von \mathbb{R}^2 auf $\mathbf{o}_2 = 0$ abgebildet. Allgemeiner gilt $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ genau dann, wenn $f(\mathbf{x}_1 - \mathbf{x}_2) = f(\mathbf{x}_1) - f(\mathbf{x}_2) = 0$, also wenn für die Differenz $\mathbf{x}_1 - \mathbf{x}_2 \in \ker f$ gilt. $\ker f$ ist als eindimensionaler Unterraum von $V_1 = \mathbb{R}^2$ eine Gerade g durch den Ursprung. Geometrisch bedeutet das: Das Bild $f(\mathbf{x})$ eines Punktes \mathbf{x} der Ebene erhalten wir, indem wir durch \mathbf{x} eine Parallele zu g legen und diese mit der x -Achse schneiden. Man spricht von einer (im Allgemeinen nicht orthogonalen Parallel-) **Projektion** des zweidimensionalen Raumes \mathbb{R}^2 auf \mathbb{R}^1 . Im allgemeinen Fall ist der Zielbereich aber kein Unterraum, weshalb er in einem anderen Koordinatensystem gedacht werden muss. Für $m = 3$ ist die Situation analog, nur dass der Kern von f zweidimensional, also eine Ebene im Raum durch den Ursprung ist, parallel zu der auf eine Achse projiziert wird. Später werden noch einige Übungsaufgaben dazu folgen.

$m = n = 2$: Als wichtige Beispiele haben wir bereits in Mathematik 1 die Drehungen d_α der Ebene um einen Winkel α kennengelernt. Der allgemeine Fall einer linearen Abbildung lässt sich geometrisch am besten wie folgt veranschaulichen. Man stelle sich das von den Einheitsvektoren $\mathbf{e}_1 = (1, 0)$ und $\mathbf{e}_2 = (0, 1)$ aufgespannte Einheitsquadrat Q vor, welches wie ein Schachbrettmuster periodisch fortgesetzt wird und die gesamte Ebene \mathbb{R}^2 überzieht. Durch f wird Q zu einem Parallelogramm verzerrt, dessen Erzeugende die Vektoren $f(\mathbf{e}_1)$ und $f(\mathbf{e}_2)$ sind. Dieser Verzerrung wird die gesamte Ebene unterworfen, so dass das Schachbrett mit quadratischen Elementen in ein Muster aus Parallelogrammen übergeht. Ein Sonderfall liegt vor, wenn $f(\mathbf{e}_1)$ und $f(\mathbf{e}_2)$ linear abhängig sind, wenn also $\operatorname{def} f > 0$. Dann ist $\operatorname{rg} f < 2$ und das Bild nicht mehr eine Ebene, sondern bei $\operatorname{rg} f = 1$ eine Gerade, auf die die Ebene durch f zusammengequetscht wird, bei $\operatorname{rg} f = 0$ die Nullabbildung.

$m = 2, n = 3$: Eine lineare Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ kann man sich vorstellen als eine Verzerrung der Ebene, wie wir sie für $m = n = 2$ besprochen haben und wo das quadratische Schachbrettmuster in eines aus Parallelogrammen übergeht, verknüpft mit einer Schrägstellung der Ebene im Raum \mathbb{R}^3 . Die Erzeugenden der Parallelogramme sind die (nun dreidimensionalen) Vektoren $f(\mathbf{e}_1)$ und $f(\mathbf{e}_2)$. Auch hier sind die entsprechenden Sonderfälle bei $\operatorname{def} f > 0$ möglich.

$m = 3, n = 2$: Für $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ muss $\ker f$ wegen der Rangformel 1.2.3.2 nichttrivial, also $\operatorname{def} f > 0$ sein. Eine typische Veranschaulichung ist ein Schrägriss. Man gibt sich eine Richtung (die von $\ker f$) im \mathbb{R}^3 vor. Ein beliebiger Punkt $\mathbf{x} \in \mathbb{R}^3$ im Raum wird auf die Basisebene \mathbb{R}^2 projiziert, indem man eine Parallele zu $\ker f$ durch \mathbf{x} zieht. Der Schnitt-

punkt dieser Geraden mit der Grundebene ist dann $f(\mathbf{x})$. Diese Veranschaulichung ist für $\text{def } f = 1$, also $\text{rg } f = 2$ sinnvoll. Für $\text{rg } f = 1$ ist das Bild unter f eine Gerade in der Grundebene \mathbb{R}^2 , auf die parallel zu einer Ebene, dem 2-dimensionalen Kern von f projiziert wird. Bei $\text{rg } f = 0$ werden alle Punkte auf $\mathbf{o}_2 \in \mathbb{R}^2$ abgebildet. So wie für $m = 2, 3$ und $n = 1$ ist auch hier zu bemerken, dass im allgemeinen Fall Definitions- und Zielbereich jedoch in verschiedenen Koordinatensystemen zu denken sind.

$m = n = 3$: Lineare Abbildungen $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ veranschaulicht man sich am besten in Analogie zur Situation bei $m = n = 2$. So wie dort ein quadratisches Schachbrettmuster zu einem unendlichen Muster aus gleichförmigen Parallelogrammen verzerrt wird, ist es hier eine dreidimensionale Unterteilung des Raumes \mathbb{R}^3 in Würfel, die zu einer solchen aus Parallelepipeden (Spaten) verzerrt wird.

Höhere Dimensionen: Lineare Abbildungen $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ mit $m > 3$ oder $n > 3$ entziehen sich naturgemäß ähnlich einfachen Veranschaulichungen wie den bisher behandelten. Im Laufe der Beschäftigung mit den abstrakten Begriffen kann man aber eine recht verlässliche Intuition dafür entwickeln, inwiefern man sich dabei von den anschaulicheren Fällen leiten lassen kann. Sicherheit geben einem dabei die allgemeinen Sätze aus der Theorie.

Zur Förderung der Intuition folgen nun einige Übungsaufgaben, insbesondere zu Projektionen:

Definition 1.2.6.1. Eine lineare Abbildung $p : V \rightarrow V$ heisst **Projektion**, falls $p^2 := p \circ p = p$.

Übungsaufgabe 32. (E) Sei $p : V \rightarrow V$ eine Projektion.

1. Finden Sie jeweils ein konkretes Beispiel einer Projektion für $V = \mathbb{R}^2$ und $V = \mathbb{R}^3$.
2. Zeigen Sie, dass wenn p eine Projektion ist, auch $\text{id}_V - p$ eine Projektion ist.
3. Bestimmen Sie für Ihre Beispiele aus 1. jeweils auch die Projektion $\text{id}_V - P$.

Übungsaufgabe 33. (E) Gegeben sind zwei Projektionen $p_1, p_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, mit $p_1 \circ p_2 = p_2 \circ p_1$. Zeigen Sie, dass dann auch $p_1 \circ p_2$ eine Projektion ist. Worauf projiziert $p_1 \circ p_2$?

Übungsaufgabe 34. (E) Finden Sie zwei Projektionen $p_1, p_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, sodass $p_1 \circ p_2 \neq p_2 \circ p_1$.

Übungsaufgabe 35. (E) Die Menge V aller unendlich oft differenzierbaren Funktionen $f : [-1, 1] \rightarrow \mathbb{R}$ bildet mit den üblichen punktweise definierten Operationen einen Vektorraum. V enthält insbesondere alle Monomfunktionen $m_n(x) = x^n$ mit $n = 0, 1, 2, \dots$. Auf V wird nun für festes $n \in \mathbb{N}$ eine lineare Abbildung $p_n : V \rightarrow V$ definiert durch

$$p_n f(x) := \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k.$$

Begründen Sie, warum p_n eine Projektion ist! Auf welchen Unterraum von V wird dabei projiziert?

1.3 Matrizen

Matrizen lassen sich zunächst ohne Bezugnahme auf algebraische Zusammenhänge als rechteckige Zahlenschemata (mit gewissen Spalten- und Zeilenzahlen m und n) definieren. Sofort fällt aber die Möglichkeit ins Auge, damit zu rechnen wie mit Vektoren in einem Standardvektorraum der Dimension mn (1.3.1). Für die Lineare Algebra wird das interessant, weil mit solchen Matrizen offenbar in eindeutiger Weise die linearen Abbildungen von einem m - in einen n -dimensionalen Vektorraum beschrieben werden können, sofern Basen dieser Räume gegeben sind (1.3.2). Das Anliegen, damit auch die Verkettung linearer Abbildungen zu erfassen, führt zum Matrizenprodukt (1.3.3). Die zahlreichen Rechenregeln, die insbesondere für lineare Abbildungen gelten, führen zu entsprechenden Rechenregeln für Matrizen, wodurch ein sehr handlicher Rechenkalkül entsteht (1.3.4). Den Abschluss des Abschnitts bildet die Anwendung dieses Kalküls auf die Beschreibung des Transformationsverhaltens von Koordinaten bei einem Basiswechsel (1.3.5).

1.3.1 Matrizen als $n \times m$ -Vektoren (-Zahlenschemata)

Inhalt in Kurzfassung: Matrizen werden als rechteckige Zahlenschemata mit n Zeilen und m Spalten eingeführt. Zunächst kann man das verstehen als lediglich eine andere Notation für Vektoren im mn -dimensionalen Standardvektorraum.

Am einfachsten kann man eine $n \times m$ -**Matrix** definieren als rechteckiges Zahlenschema

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix}$$

mit n Zeilen und m Spalten, bestehend aus sogenannten **Einträgen**, **Elementen** oder auch **Komponenten** $a_{i,j} \in \mathbb{R}$. Wir schreiben auch

$$A = (a_{i,j}) \quad \text{bzw. ausführlicher} \quad A = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}.$$

Die Zahl $a_{i,j}$ heißt in diesem Zusammenhang der **Eintrag** von A in der i -ten **Zeile** und j -ten **Spalte**. Der Vektor

$$(a_{i,1}, a_{i,2}, \dots, a_{i,m}) \in \mathbb{R}^m$$

heißt der i -te **Zeilenvektor** und

$$\begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix} \in \mathbb{R}^n$$

der j -te **Spaltenvektor** von A . Stimmen die Anzahl n der Zeilen und die Anzahl m der Spalten von A überein, so heißt A eine **quadratische Matrix**.

1 Lineare Algebra

Oft schreibt man a_{ij} statt $a_{i,j}$. Um den Doppelindex aber vom Produkt ij der Indizes zu unterscheiden, bevorzugen wir (wenigstens vorläufig) die zweite Schreibweise.

Eine Matrix A ist also durch mn reelle Zahlen gegeben, so wie ein Vektor im Standardvektorraum $\mathbb{R}^{n \times m}$ der Dimension mn . Wollen wir diese Analogie weiter ausnutzen, liegt es nahe, auf der Menge $\mathcal{M}(n, m)$ aller $n \times m$ -Matrizen Vektorraumoperationen komponentenweise durch

$$A + B = (a_{i,j}) + (b_{i,j}) := (a_{i,j} + b_{i,j}) \quad \text{und} \quad rA = r(a_{i,j}) := (ra_{i,j})$$

für $A = (a_{i,j}), B = (b_{i,j}) \in \mathcal{M}(n, m)$ und $r \in \mathbb{R}$ zu definieren. Sei

$$\varphi : \mathcal{M}(n, m) \rightarrow \mathbb{R}^{n \cdot m}$$

die Abbildung mit

$$\varphi : A = (a_{i,j}) \mapsto (a_{1,1}, a_{1,2}, \dots, a_{1,m}, a_{2,1}, a_{2,2}, \dots, a_{2,m}, \dots, a_{n,1}, a_{n,2}, \dots, a_{n,m}),$$

die aus einer $n \times m$ -Matrix A jenen Vektor mit nm Komponenten, in dem alle Komponenten von A (in einer bestimmten, zum Beispiel der hier angedeuteten Reihenfolge) enthält. Offensichtlich ist φ bijektiv und wegen

$$\varphi(A + B) = \varphi(A) + \varphi(B) \quad \text{und} \quad \varphi(rA) = r\varphi(A)$$

auch linear, d.h. (bezüglich der oben definierten Matrizenoperationen) ein Isomorphismus zwischen $\mathcal{M}(n, m)$ und $\mathbb{R}^{n \times m}$. Folglich (siehe Satz 1.2.4.2) ist $\mathcal{M}(n, m)$ selbst ein mn -dimensionaler Vektorraum. Die kanonische Basis von $\mathcal{M}(n, m)$ ist gegeben durch die Menge der Matrizen $A_{i,j}$ ($1 \leq i \leq n, 1 \leq j \leq m$). Dabei hat $A_{i,j}$ nur einen einzigen von 0 verschiedenen Eintrag, und zwar den Eintrag 1 in der i -ten Zeile und j -ten Spalte.

Jede Matrix (quadratisch oder nicht) lässt sich **transponieren**, indem man Zeilen und Spalten vertauscht, genauer: Ist $A = (a_{ij})$ eine Matrix, dann versteht man unter der **transponierten** oder auch **adjungierten** Matrix A^T (manchmal auch als A' oder A^* notiert) von A , die Matrix $A^T := (a_{ji})$, ausführlicher geschrieben:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}.$$

(n, m) – Matrix
 (m, n) – Matrix

Eine $n \times m$ -Matrix wird also zu einer $m \times n$ -Matrix. Die tiefere Bedeutung der Transposition (Adjunktion) für die dargestellte lineare Abbildung ist auf den ersten Blick nicht ersichtlich und würde hier zu weit führen. Die Definition der Transponierten ist aber einfach und immer wieder nützlich.

Was wir bisher getan haben, war nicht mehr als ein Umschreiben von Vektoren mit mn Komponenten in rechteckige Schemata. Matrizen haben aber eine tiefere Bedeutung der wir uns nun zuwenden wollen.

1.3.2 Die Matrix einer linearen Abbildung bezüglich gegebener Basen

Inhalt in Kurzfassung: Ihre überragende Bedeutung in der Linearen Algebra erhalten Matrizen deshalb, weil sie sich in natürlicher Weise auch als Darstellungen linearer Abbildungen bezüglich gegebener Basen der zugrunde liegenden Vektorräume ergeben. Wie genau das zu verstehen ist, wird nun erklärt.

Seien V_1 und V_2 Vektorräume mit Basen $B_1 = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ bzw. $B_2 = (\mathbf{y}_1, \dots, \mathbf{y}_n)$ und $f : V_1 \rightarrow V_2$ linear. Aus Satz 1.2.4.1 wissen wir, dass f eindeutig bestimmt ist, wenn die Bilder $f(\mathbf{x}_1), \dots, f(\mathbf{x}_m)$ gegeben sind. Jedes dieser Bilder wiederum lässt sich in eindeutiger Weise als Linearkombination

$$f(\mathbf{x}_j) := \sum_{i=1}^n a_{i,j} \mathbf{y}_i$$

der \mathbf{y}_i schreiben. Also ist f (bezüglich der gegebenen Basen B_1 und B_2) eindeutig bestimmt durch die Gesamtheit der insgesamt $m \cdot n$ Koeffizienten $a_{i,j}$, $1 \leq i \leq n$ und $1 \leq j \leq m$. In diesem Sinn kann der linearen Abbildung f auf eindeutige Weise die $n \times m$ -Matrix

$$A = A_f = A_{f, B_1, B_2} := \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix}$$

zugeordnet werden. Diese Matrix A heißt auch die Matrix oder **Matrixdarstellung** von f bezüglich der Basen B_1 und B_2 . Ganz allgemein verwenden wir die Schreibweise $A = (a_{i,j})$, um festzuhalten, dass die Eintragung in der i -ten Zeile und j -ten Spalte einer Matrix A mit $a_{i,j}$ bezeichnet wird.

In analoger Vorgangsweise zu 1.3.1, wo wir einen Isomorphismus $\varphi : \mathcal{M}(n, m) \rightarrow \mathbb{R}^{n \times m}$ gefunden haben, legt der nun aufgedeckte Zusammenhang einen Isomorphismus zwischen dem Vektorraum $L(V_1, V_2)$ aller linearen Abbildungen $f : V_1 \rightarrow V_2$ und dem Vektorraum $\mathcal{M}(n, m)$ aller $n \times m$ -Matrizen nahe. Weil es notationell etwas angenehmer ist, kehren wir die Richtung der obigen Argumentation um und ordnen jeder $n \times m$ -Matrix $A = (a_{i,j})$ die lineare Abbildung $\psi(A) = \psi_{B_1, B_2}(A) = f_A = f_{A, B_1, B_2}$ zu, die laut Satz 1.2.4.1 eindeutig definiert ist durch

$$f_A : \mathbf{x}_j \mapsto \sum_{i=1}^n a_{i,j} \mathbf{y}_i, \quad j = 1, \dots, m.$$

Diese Abbildung ψ ist nicht nur bijektiv (wegen Satz 1.2.4.1). Man überzeugt sich unmittelbar, dass der Summe und dem Vielfachen von Matrizen die entsprechende Summe bzw. das entsprechende Vielfache der zugeordneten linearen Abbildungen entspricht:

$$f_{A_1 + A_2, B_1, B_2} = f_{A_1, B_1, B_2} + f_{A_2, B_1, B_2} \quad \text{und} \quad f_{rA, B_1, B_2} = r f_{A, B_1, B_2}.$$

Somit gilt $\psi(A + B) = \psi(A) + \psi(B)$ und $\psi(rA) = r\psi(A)$. Also ist

$$\psi : \mathcal{M}(n, m) \rightarrow L(V_1, V_2), \quad A \mapsto f_A$$

über die Bijektivität hinaus auch linear, folglich ein Isomorphismus zwischen den Vektorräumen $\mathcal{M}(n, m)$ und $L(V_1, V_2)$.

Übungsaufgabe 36. (P) Erklären Sie ausführlich die oben verwendeten Beziehungen

- $f_{A_1+A_2, B_1, B_2} = f_{A_1, B_1, B_2} + f_{A_2, B_1, B_2}$ und
- $f_{rA, B_1, B_2} = r f_{A, B_1, B_2}$.

Bezieht man sich auf die Standardvektorräume $V_1 = \mathbb{R}^m$ und $V_2 = \mathbb{R}^n$ und sind keine anderen Basen gegeben, so interpretiert man eine Matrix in Bezug auf die kanonischen Basen von V_1 und V_2 . In diesem Fall ist also $\mathbf{x}_j = \mathbf{e}_j^{(m)}$ für $j = 1, \dots, m$ und $\mathbf{y}_i = \mathbf{e}_i^{(n)}$ für $i = 1, \dots, n$ zu setzen. Dann entspricht die Matrix $A = (a_{i,j})$ jener linearen Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$, die den kanonischen Basisvektor $\mathbf{e}_j^{(m)} \in \mathbb{R}^m$ auf den j -ten Spaltenvektor von A abbildet, also

$$f = f_A : \mathbf{e}_j^{(m)} \mapsto \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix} = \sum_{i=1}^n a_{i,j} \mathbf{e}_i^{(n)} \in \mathbb{R}^n.$$

Die Spaltenvektoren von A sind also die Bilder der kanonischen Einheitsvektoren unter f . Damit ist aber auch klar, wie f auf einen beliebigen Vektor

$$\mathbf{x} = \sum_{j=1}^m x_j \mathbf{e}_j^{(m)}$$

wirkt, nämlich (Linearität!)

$$f_A(\mathbf{x}) = f_A \left(\sum_{j=1}^m x_j \mathbf{e}_j^{(m)} \right) = \sum_{j=1}^m x_j f_A(\mathbf{e}_j^{(m)}) = \sum_{j=1}^m x_j \sum_{i=1}^n a_{i,j} \mathbf{y}_i = \sum_{i=1}^n \left(\sum_{j=1}^m a_{i,j} x_j \right) \mathbf{y}_i.$$

Die Koordinaten y_i des Bildvektors $f_A(\mathbf{x}) = \mathbf{y} = (y_1, \dots, y_n)$ von $\mathbf{x} = (x_1, \dots, x_m)$ sind also gegeben durch $y_i = \sum_{j=1}^m a_{i,j} x_j$ für $i = 1, \dots, n$.

Haben wir es statt der kanonischen Basis mit irgendwelchen vorgegebenen Basen $B_1 = \{\mathbf{a}_1, \dots, \mathbf{a}_m\} \subseteq V_1$ und $B_2 = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subseteq V_2$ zu tun und beziehen wir die Matrix $A = (a_{i,j})$ darauf, so wird dadurch natürlich eine andere lineare Abbildung dargestellt. Dem m -tupel (x_1, \dots, x_m) entspricht dann der Vektor

$$\mathbf{x} = \sum_{j=1}^m x_j \mathbf{a}_j.$$

Die A bezüglich B_1 und B_2 zugeordnete lineare Abbildung f_{A, B_1, B_2} bildet diesen Vektor \mathbf{x} auf

$$f_{A, B_1, B_2}(\mathbf{x}) = f_A \left(\sum_{j=1}^m x_j \mathbf{a}_j \right) = \sum_{j=1}^m x_j f_A(\mathbf{a}_j) = \sum_{j=1}^m x_j \sum_{i=1}^n a_{i,j} \mathbf{b}_i = \sum_{i=1}^n \left(\sum_{j=1}^m a_{i,j} x_j \right) \mathbf{b}_i$$

ab. Die Koordinaten y_i dieses Bildvektors sind also gegeben durch $y_i = \sum_{j=1}^m a_{i,j}x_j$ für $i = 1, \dots, n$. Dieser Wert y_i lässt sich auch deuten als das Skalarprodukt des i -ten Zeilenvektors von A mit dem Vektor $(x_1, \dots, x_m) \in \mathbb{R}^m$, der \mathbf{x} bezüglich der Basis B_2 darstellt. Ist dieser Vektor $(x_1, \dots, x_m) = \mathbf{e}_j \in \mathbb{R}^m$ der j -te kanonische Einheitsvektor in \mathbb{R}^m , so ist $\mathbf{x} = \mathbf{a}_j$ und $y_i = a_{i,j}$, $i = 1, \dots, n$. Das bedeutet: Der j -te Spaltenvektor von A besteht aus den Komponenten des Bildes $f_A(\mathbf{a}_j)$ des j -ten Basisvektors $\mathbf{a}_j \in B_1$ bezüglich der Basis B_2 . Wenn wir $\mathbf{x} = (x_1, \dots, x_m)$ nicht als Zeilen-, sondern als Spaltenvektor schreiben, bedeutet das:

$$\mathbf{Ax} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m a_{1,j}x_j \\ \sum_{j=1}^m a_{2,j}x_j \\ \vdots \\ \sum_{j=1}^m a_{n,j}x_j \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \mathbf{y}$$

Im allgemeinen Fall sind die x_j Koordinaten bezüglich einer Basis B_1 von V_1 , die y_i bezüglich einer Basis B_2 von V_2 . Beziehen wir alles auf die Standardvektorräume $V_1 = \mathbb{R}^m$ und $V_2 = \mathbb{R}^n$ sowie auf die kanonischen Basen, so können wir A mit jener linearen Abbildung identifizieren, die $\mathbf{x} \in \mathbb{R}^m$ selbst (und nicht nur den durch \mathbf{x} bezüglich B_1 dargestellten Vektor) auf den Vektor $\mathbf{y} \in \mathbb{R}^n$ abbildet. Obiger Zusammenhang $\mathbf{Ax} = \mathbf{y}$ lässt sich also sowohl als Matrizenprodukt ($n \times m$ -Matrix mal $m \times 1$ -Matrix, siehe etwas später) wie auch als Abbildungsgleichung $A : \mathbf{x} \mapsto \mathbf{y}$ lesen. Von diesen Möglichkeiten wird im Matrizenkalkül permanent Gebrauch gemacht werden.

Der **Nullabbildung** $V_1 \rightarrow V_2$ mit $\mathbf{x} \mapsto \mathbf{o}_2$ für alle $\mathbf{x} \in V_1$ entspricht bezüglich beliebiger Basen die **Nullmatrix**, deren sämtliche Eintragungen $= 0$ sind.

Da jede Matrix $n \times m$ -Matrix bezüglich der kanonischen Basis einer linearen Abbildung $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ entspricht, lassen sich sämtliche Begriffe für lineare Abbildungen zwischen endlichdimensionalen Vektorräumen auf Matrizen übertragen. Zum Beispiel:

Definition 1.3.2.1. Sei A eine $n \times m$ -Matrix und $f = f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ die A bezüglich der kanonischen Basen von \mathbb{R}^m und \mathbb{R}^n zugeordnete lineare Abbildung. Dann definieren wir **Rang, Kern und Defekt** von A durch $\text{rg } A := \text{rg } f$, $\ker A := \ker f$ und $\text{def } A := \text{def } f$. Ist $m = n$, so heißt A **regulär**, wenn f regulär (d.h. bijektiv) ist, sonst **singulär**.

Weil jeder reelle Vektorraum der Dimension m bzw. n isomorph ist zu \mathbb{R}^m bzw. \mathbb{R}^n , überträgt sich der in 1.3.2.1 defintorisch festgeschriebene Zusammenhang von den Standardvektorräumen bezüglich der kanonischen Basen sehr leicht auf beliebige Vektorräume:

Proposition 1.3.2.2. Seien A eine $n \times m$ -Matrix, V_1 und V_2 Vektorräume mit $\dim V_1 = m$ und $\dim V_2 = n$, B_1, B_2 Basen von V_1 bzw. V_2 und $f : V_1 \rightarrow V_2$ die A bezüglich dieser Basen zugeordnete lineare Abbildung. Dann gilt $\text{rg } A = \text{rg } f$ und $\text{def } A = \text{def } f$. A ist genau dann regulär/singulär, wenn f regulär/singulär ist.

Übungsaufgabe 37. (E) Beweisen Sie Proposition 1.3.2.2. Hinweis: Verwenden Sie Isomorphismen zwischen \mathbb{R}^m und V_1 sowie zwischen \mathbb{R}^n und V_2 .

Übungsaufgabe 38. (T^*) Schreiben Sie die folgenden geometrisch definierten linearen Abbildungen im \mathbb{R}^2 in Matrixform um!

1. Spiegelung am Ursprung.
2. Spiegelung an der Gerade mit der Gleichung $ax + by = 0$.
3. Orthogonale Projektion auf die Gerade mit der Gleichung $ax + by = 0$.
4. Streckung um den Faktor λ .
5. Drehung um den Winkel ϕ .

Wie könnte jeweils eine Verallgemeinerung in den \mathbb{R}^3 bzw. den \mathbb{R}^n aussehen?

Übungsaufgabe 39. (T) Gibt es eine lineare Transformation, die den Vektor $(1, 0, 1)$ um den Faktor (-2) streckt, den Vektor $(0, 2, 0)$ um den Faktor $\frac{1}{2}$ staucht und den Vektor $(-1, 1, 1)$ am Ursprung spiegelt? Falls ja, geben Sie die zugehörige Abbildungsmatrix an und machen Sie die Probe.

Übungsaufgabe 40. (T) Wie Aufgabe 39 aber mit den Vektoren $(1, 1, 1)$, $(2, 1, 0)$ und $(0, 1, 2)$.

Übungsaufgabe 41. (E) Sei $a, b, c, d \in \mathbb{R}$. Betrachten Sie die Menge \mathcal{M} aller 2×2 Matrizen, sowie die Teilmengen \mathcal{S} bzw. \mathcal{A} , bestehend aus Matrizen der Form

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 0 & d \\ -d & 0 \end{pmatrix}.$$

1. Zeigen Sie, dass \mathcal{S} und \mathcal{A} Vektorräume sind und bestimmen Sie jeweils eine Basis.
2. Weisen Sie nach, dass man jede reelle 2×2 Matrix M eindeutig als Summe $M = S + A$ schreiben kann, wobei S aus dem Unterraum \mathcal{S} und A aus dem Unterraum \mathcal{A} ist.
3. Bestimmen Sie speziell diese Zerlegung für die Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

wobei a, b, c, d die letzten vier Ziffern Ihrer Matrikelnummer sind.

1.3.3 Das Matrizenprodukt

Inhalt in Kurzfassung: Das Produkt zweier Matrizen ist so definiert, dass es der Verkettung linearer Abbildungen entspricht. Zu seiner Berechnung hat man die Skalarprodukte der Zeilenvektoren der ersten Matrix mit den Spaltenvektoren der zweiten zu bilden.

Seien V_1, V_2 und V_3 Vektorräume mit Basen $B_1 = (\mathbf{a}_1, \dots, \mathbf{a}_m)$, $B_2 = (\mathbf{b}_1, \dots, \mathbf{b}_l)$ und $B_3 = (\mathbf{c}_1, \dots, \mathbf{c}_n)$, und $f : V_1 \rightarrow V_2$ und $g : V_2 \rightarrow V_3$ lineare Abbildungen sowie $h = g \circ f$ ihre Verkettung. Wir wollen das Produkt $C = AB$ von zwei Matrizen A und B so definieren, dass es der Verkettung $h = g \circ f$ entspricht. Um in der Notation die Reihenfolge beibehalten zu können sei also A die Matrixdarstellung von g bezüglich B_2 und B_3 , B die von f bezüglich B_1 und B_2 . Dann ist A eine $n \times l$ -Matrix, B eine $l \times m$ -Matrix. Das Produkt C zweier Matrizen kann also nur dann sinnvoll definiert werden, wenn die Spaltenzahl der ersten mit der Zeilenzahl der zweiten übereinstimmt, nämlich gleich der Dimension des Vektorraums V_2 ist.

Zur Berechnung der Einträge des Produktes $C = (c_{i,j})$ erinnern wir uns daran, dass die Spaltenvektoren die Bilder der Basisvektoren unter der durch die Matrix (bezüglich der vorgegebenen Basen) dargestellten Abbildung sind. In unserem Fall ist also $c_{i,j}$ die i -te Komponente des Vektors $h(\mathbf{a}_j) = g(f(\mathbf{a}_j))$. Die Matrixdarstellungen $B = (b_{k,j})$ von f und $A = (a_{i,k})$ von g bezüglich B_1, B_2 und B_3 besagen

$$f(\mathbf{a}_j) = \sum_{k=1}^l b_{k,j} \mathbf{b}_k \quad \text{und} \quad g(\mathbf{b}_k) = \sum_{i=1}^n a_{i,k} \mathbf{c}_i$$

für $j = 1, \dots, m$ bzw. $k = 1, \dots, l$. Ineinander eingesetzt liefert das unter Ausnutzung der Linearität von g sowie nach Vertauschung der Summation

$$h(\mathbf{a}_j) = g(f(\mathbf{a}_j)) = g\left(\sum_{k=1}^l b_{k,j} \mathbf{b}_k\right) = \sum_{k=1}^l b_{k,j} g(\mathbf{b}_k) = \sum_{i=1}^n \left(\sum_{k=1}^l a_{i,k} b_{k,j}\right) \mathbf{c}_i.$$

Hieraus sind die gesuchten Einträge $c_{i,j}$ von C als die Ausdrücke in der letzten großen Klammer abzulesen:

$$c_{i,j} = \sum_{k=1}^l a_{i,k} b_{k,j} = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,l} \end{pmatrix} \cdot \begin{pmatrix} b_{1,j} \\ b_{2,j} \\ \vdots \\ b_{n_2,j} \end{pmatrix},$$

das Skalarprodukt des i -ten Zeilenvektors von A und des j -ten Spaltenvektors von B .

Deshalb definiert man das Matrizenprodukt AB für Matrizen A und B , wo die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt wie folgt.

Definition 1.3.3.1. Sei $A = (a_{i,k})$ eine $n \times l$ -Matrix, $B = (b_{k,j})$ eine $l \times m$ -Matrix, so ist das **Produkt** AB definiert als $n \times m$ -Matrix $C = (c_{i,j})$ mit den Eintragungen $c_{i,j} = \sum_{k=1}^l a_{i,k} b_{k,j}$, $i = 1, \dots, n$ und $j = 1, \dots, m$.

Ein Beispiel:

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 4 & 5 \end{pmatrix}.$$

Der Eintrag 4 in der Matrix rechts steht in der zweiten Zeile und ersten Spalte, ist also das Skalarprodukt des zweiten Zeilenvektors $(2, 1, 4)$ des ersten Faktors mit dem ersten

Spaltenvektor $(1, 2, 0)^T$ des zweiten Faktors, also tatsächlich $2 \cdot 1 + 1 \cdot 2 + 4 \cdot 0 = 2 + 2 + 0 = 4$. Die anderen Einträge ergeben sich entsprechend.

Manchmal sind folgende Rechenregeln für transponierte Matrizen nützlich:

Proposition 1.3.3.2. *Die Transposition von Matrizen gehorcht folgenden Rechenregeln:*

- (i) $(A^T)^T = A$ für alle Matrizen A ,
- (ii) $(A + B)^T = A^T + B^T$ für alle Matrizen A, B vom selben Typ,
- (iii) $(AB)^T = B^T A^T$ für alle Matrizen A, B , für die man AB bilden kann.

Übungsaufgabe 42. (E) Überprüfen Sie Proposition 1.3.3.2

1.3.4 Matrizenkalkül

Inhalt in Kurzfassung: Lineare Abbildungen und ihre Verknüpfungen (Vektorraumoperationen, Verkettung) genügen Rechenregeln, die sich auf die dargestellten Matrizen übertragen. Daraus ergibt sich ein Rechenkalkül für Matrizen, wo die aus dem Rechnen mit Zahlen vertrauten Assoziativ- und Distributivgesetze weiterhin gelten. Auch Rechenregeln für neutrale Elemente (Null- und Einheitsmatrix) folgen gewohnten Mustern. Das Kommutativgesetz allerdings gilt nur bezüglich der Addition von Matrizen, nicht aber bezüglich der Multiplikation.

Weil die Matrizenmultiplikation der Komposition der zugehörigen linearen Abbildungen (z.B. zwischen den Standardvektorräumen bezüglich der kanonischen Basen) entspricht, vererben sich Assoziativgesetz $(h \circ g) \circ f = h \circ (g \circ f)$ sowie die beiden Distributivgesetze $f \circ (g_1 + g_2) = f \circ g_1 + f \circ g_2$ und $(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$ auf Matrizen:

$$(AB)C = A(BC), \quad A(B_1 + B_2) = AB_1 + AB_2 \quad \text{und} \quad (A_1 + A_2)B = A_1B + A_2B.$$

Übungsaufgabe 43. (P) Überprüfen Sie diese beiden Distributivgesetze für lineare Abbildungen und erklären Sie jeden Rechenschritt sorgfältig.

Wegen dieser formalen Ähnlichkeit mit der Multiplikation gewöhnlicher Zahlen spricht man auch von der **Multiplikation** linearer Abbildungen und, in weiterer Folge, der zugehörigen Matrizen. Besonders überzeugend wird das für $V_1 = V_2 = V_3 = V$. Dann bildet die Menge $L(V, V)$ aller linearen Abbildungen $f : V \rightarrow V$ bezüglich dieser Operationen (und analog die Menge $\mathcal{M}(n, n)$ aller $n \times n$ -Matrizen) einen sogenannten **Ring**², ähnlich den ganzen Zahlen mit Addition und Multiplikation. Es gibt auch ein Einselement, weil die identische Abbildung $\text{id}_V : V \rightarrow V, \mathbf{x} \mapsto \mathbf{x}$, linear ist. Bezüglich einer beliebigen Basis $B = B_1 = B_2 = B_3$ eines n -dimensionalen Vektorraums $V = V_1 = V_2 = V_3$ wird sie von der (n -dimensionalen) **Einheitsmatrix** $I = I_n = (\delta_{i,j})$ dargestellt. Dabei ist $\delta_{i,j}$

²Von einem Ring verlangt man Ähnliches wie von einem Körper, nämlich additiv eine abelsche Gruppe, außerdem das Distributivgesetz. Man begnügt sich aber multiplikativ mit dem Assoziativgesetz.

das sogenannte **Kronecker-Delta**³, welches für $i = j$ den Wert 1 und sonst den Wert 0 annimmt. Die n -dimensionale Einheitsmatrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

ist also jene quadratische $n \times n$ -Matrix, die in der sogenannten **Hauptdiagonale** lauter Einsen und sonst lauter Nullen hat.

Stellt eine quadratische Matrix A (bezüglich gewisser Basen) eine bijektive lineare Abbildung $f : V_1 \rightarrow V_2$ dar, dann gibt es ein wegen 1.2.1.3 gleichfalls lineare Umkehrabbildung $f^{(-1)}$. Diese hat (bezüglich derselben Basen) eine Matrixdarstellung, die man mit A^{-1} bezeichnet und **Inverse** von A nennt. Wegen $f^{(-1)} \circ f = \text{id}_{V_1}$ und $f \circ f^{(-1)} = \text{id}_{V_2}$ gilt erwartungsgemäß

$$A^{-1}A = AA^{-1} = I_n$$

mit $n = \dim V_1 = V_2$.

Die Analogie zwischen Zahlen und Matrizen hat allerdings auch ihre Grenzen. So ist die Multiplikation von Matrizen und somit auch von linearen Abbildungen in der Regel *nicht* kommutativ ist. Ein Beispiel dazu:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 3 \\ 12 & 7 \end{pmatrix} \neq \begin{pmatrix} 3 & 4 \\ 6 & 10 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Eine Folgerung daraus: Man kann zwar Matrizenpotenzen genauso zum Beispiel rekursiv definieren wie Potenzen gewöhnlicher Zahlen, nämlich $A^0 := I$ (Einheitsmatrix, neutrales Element bezüglich der Multiplikation) und $A^{n+1} := A^n A$, es gelten auch die Rechenregeln $A^m A^n = A^{m+n}$ und $(A^m)^n = A^{mn}$, es kann aber $(AB)^n \neq A^n B^n$ gelten.

Übungsaufgabe 44. (E) Finden Sie quadratische Matrizen A und B und ein $n \in \mathbb{N}$ mit $(AB)^n \neq A^n B^n$.

Übungsaufgabe 45. (E) Finden Sie ein Beispiel eines Vektorraums V und linearer Abbildungen $f, g : V \rightarrow V$ mit $f \circ g \neq g \circ f$. Hinweis: Die Dimension von V muss mindestens 2 sein, denn im Eindimensionalen entspricht alles dem Rechnen reeller Zahlen, die ja eine kommutative Multiplikation haben.

Als Beispiel und gleichzeitig als Wiederholung aus Mathematik 1 erinnern wir an die Hintereinanderausführung von Drehungen und ihre Matrixdarstellung: Die Matrizen

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix}$$

³benannt nach dem Mathematiker Leopold Kronecker (1823-1891)

1 Lineare Algebra

stellen die Drehungen der Ebene um die Winkel φ und ψ dar, die Matrix

$$C = \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix}$$

jene um den Winkel $\varphi + \psi$. Aus dem Additionstheorem für Sinus und Cosinus folgt

$$\begin{aligned} & \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi \cos \psi - \sin \varphi \sin \psi & -\cos \varphi \sin \psi - \sin \varphi \cos \psi \\ \sin \varphi \cos \psi + \cos \varphi \sin \psi & -\sin \varphi \sin \psi + \cos \varphi \cos \psi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix}. \end{aligned}$$

Zur Erinnerung: Diese Rechnung haben wir schon in Mathematik 1 beim Beweis der Additionstheoreme für Cosinus und Sinus durchgeführt.

Übungsaufgabe 46. (T) Berechnen Sie für die Matrizen

$$A = \begin{pmatrix} 0 & 5 & 1 \\ 4 & 0 & 2 \\ 1 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 5 \\ 1 & 8 & -1 \end{pmatrix}$$

die Ausdrücke $AB + BA$ und $AB - BA$.

Übungsaufgabe 47. (T) Gegeben sind die Matrizen

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 2 \\ 0 & -2 & 3 \\ 0 & 3 & -3 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Berechnen Sie

1. $A + B$ und $B + A$.
2. AB und BA .
3. C^2 und C^3 .

und geben Sie an, welche Rechenregeln, die für reelle Zahlen gelten, hier verletzt sind.

Übungsaufgabe 48. Berechnen Sie für die Matrizen A, B

$$A = \begin{pmatrix} 3 & 5 & 7 \\ 4 & 6 & 8 \\ 1 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 5 \\ 1 & 8 & -7 \end{pmatrix}$$

die Ausdrücke A^2, B^2, AB, BA und $(A + B)^2$. Finden Sie eine „binomische Formel“ für $(A + B)^2$, d.h. stellen Sie $(A + B)^2$ mit Hilfe von A^2, B^2, AB, BA dar.

Übungsaufgabe 49. (T) Berechnen Sie für die Matrizen

$$A = \begin{pmatrix} 3 & 5 \\ 4 & 6 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

alle Produkte, die aus zwei Faktoren gebildet werden können, d.h. etwa $A \cdot B, C \cdot C, \dots$

Übungsaufgabe 50. (T) Wie Aufgabe 49, aber mit Produkten, die aus drei Faktoren gebildet werden können.

Übungsaufgabe 51. (T) Berechnen Sie mittels vollständiger Induktion alle Matrixpotenzen $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$, $n \in \mathbb{N}$.

Übungsaufgabe 52. (T*) Wie Aufgabe 51, aber für die Matrix $\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ mit $\alpha \in \mathbb{R}$.

Übungsaufgabe 53. (T*) Wie Aufgabe 51, aber für die Matrix $\begin{pmatrix} \beta & 1 \\ 0 & \beta \end{pmatrix}$ mit $\beta \in \mathbb{R}$.

Übungsaufgabe 54. (E) Betrachten Sie die beiden Matrizen

$$B(\phi) = \begin{pmatrix} \cosh \phi & \sinh \phi \\ \sinh \phi & \cosh \phi \end{pmatrix}, \quad B(\psi) = \begin{pmatrix} \cosh \psi & \sinh \psi \\ \sinh \psi & \cosh \psi \end{pmatrix}.$$

Berechnen Sie $B(\phi)B(\psi)$ und zeigen Sie durch Einsetzen in die Definition (zur Erinnerung: $\cosh \phi = \frac{1}{2}(e^\phi + e^{-\phi})$ bzw. $\sinh \phi = \frac{1}{2}(e^\phi - e^{-\phi})$), dass $B(\phi)B(\psi) = B(\phi + \psi)$.

Übungsaufgabe 55. (E) Für eine beliebige $n \times n$ Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

bezeichnet $\text{Tr}(A) := a_{11} + a_{22} + a_{33} + \dots + a_{nn}$ die Spur (trace) von A .

1. Rechnen Sie nach, dass $\text{Tr}(\lambda A + \mu B) = \lambda \text{Tr}(A) + \mu \text{Tr}(B)$ gilt ($\mu, \lambda \in \mathbb{R}$).
2. Rechnen Sie nach, dass $\text{Tr}(AB) = \text{Tr}(BA)$ indem Sie in die Definition einsetzen.
3. Weisen Sie nach, dass i.A. $\text{Tr}(AB) \neq \text{Tr}(A) \text{Tr}(B)$, indem Sie ein Gegenbeispiel finden.

Übungsaufgabe 56. (T) Berechnen Sie für die Matrizen

$$A = \begin{pmatrix} 2 & 1 \\ 4 & 0 \\ -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix},$$

die Spur (siehe Übungsaufgabe 55) aller Produkte, die aus drei Faktoren gebildet werden können (und sofern die Spur für dieses Produkt definiert ist).

Hinweis: Überlegen sich zuerst, für welche Ausdrücke in A, B, C die Spur überhaupt definiert ist. Wenn Sie dann noch geeignete Rechenregeln für die Spur verwenden, können Sie sich das Berechnen einiger Matrizenprodukte ersparen.

Übungsaufgabe 57. (E) Es sei Λ eine Diagonalmatrix und A eine allgemeine Matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \\ 0 & & \ddots & \vdots \\ 0 & & \cdots & \lambda_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Berechnen Sie ΛA und $A\Lambda$. Wie könnte eine Bedingung an Λ lauten, damit $A\Lambda = \Lambda A$ sichergestellt ist?

Übungsaufgabe 58. (E) Zeigen Sie: Für eine feste $n \times n$ Matrix B gilt $AB = BA$ für alle $n \times n$ Matrizen A dann und nur dann, wenn $B = \lambda I_n$ (dabei ist $\lambda \in \mathbb{R}$ und I_n die $n \times n$ Einheitsmatrix).

Mit anderen Worten: wenn B kein Vielfaches der Einheitsmatrix ist, gibt es immer ein A sodass $AB \neq BA$.

Anleitung: Berechnen Sie die Produkte von B mit Matrizen A der Form: a_{ij} ist immer 0 außer an einer Stelle.

1.3.5 Transformation bei Basiswechsel

Inhalt in Kurzfassung: Die Darstellung der Koordinaten von Vektoren bezüglich einer Basis erfährt beim Übergang zu einer anderen Basis eine Transformation, die sich als linear erweist. Folglich lässt sich dieser Übergang durch eine (reguläre quadratische) Matrix beschreiben. Analoges wie für die Koordinaten von Vektoren gilt auch für die Einträge von Matrizen, die ja lineare Abbildungen bezüglich bestimmter Basen darstellen.

Es sei V ein Vektorraum mit zwei Basen $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ und $B' = (\mathbf{b}'_1, \dots, \mathbf{b}'_n)$. Dann hat jeder Vektor $\mathbf{a} \in V$ Darstellungen

$$\mathbf{a} = \sum_{i=1}^n x_i \mathbf{b}_i = \sum_{j=1}^n x'_j \mathbf{b}'_j$$

mit eindeutigen x_i und x'_j . Wir bezeichnen mit $\mathbf{x} = (x_1, \dots, x_n)$ und $\mathbf{x}' = (x'_1, \dots, x'_n)$ die resultierenden Koordinaten- n -Tupel und fassen diese als Elemente des Standardvektorraums \mathbb{R}^n auf. Die Zuordnung $\mathbf{x} \mapsto \mathbf{a}$ ist bijektiv und linear ebenso wie die Zuordnung $\mathbf{a} \mapsto \mathbf{x}'$, folglich auch deren Verkettung $\mathbf{x} \mapsto \mathbf{x}'$. Folglich gibt es quadratische und zueinander inverse Matrizen S, S^{-1} , die diesen Übergang in beide Richtungen beschreiben:

$$\mathbf{x}' = S\mathbf{x} \quad \text{bzw.} \quad \mathbf{x} = S^{-1}\mathbf{x}'$$

für beliebige $\mathbf{x} \in V$. Sei nun A die (quadratische) Matrix für die lineare Abbildung $f : V \rightarrow V$ bezüglich B und A' die Matrix für f bezüglich B' . Weiters seien \mathbf{x} und \mathbf{x}' wie oben die Koordinaten des Vektors \mathbf{a} bezüglich B bzw. B' , analog \mathbf{y} und \mathbf{y}' die Koordinaten des Vektors $f(\mathbf{a})$ bezüglich B bzw. B' . Dann folgt aus dem Bisherigen

$$A'\mathbf{x}' = \mathbf{y}' = S\mathbf{y} = SA\mathbf{x} = SAS^{-1}\mathbf{x}'.$$

Da dies für alle $\mathbf{x} \in V$ gilt und die Matrixdarstellung einer linearen Abbildung eindeutig ist, folgt

$$A' = SAS^{-1} \quad \text{bzw.} \quad A = S^{-1}A'S.$$

Die Matrizen S und S^{-1} ermöglichen also auch die Umrechnung von quadratischen Matrizen, aufgefasst als Darstellungen linearer Abbildungen, zwischen den Basen B und B' .

Analoge Überlegungen lassen sich allgemeiner anstellen, wenn es nämlich um lineare Abbildungen $f : V_1 \rightarrow V_2$ zwischen zwei verschiedenen Vektorräumen V_1 und V_2 geht mit jeweils zwei verschiedenen Basen B_1 und B'_1 bzw. B_2 und B'_2 . Vermittle S_1 den Übergang zwischen B_1 und B'_1 , S_2 jenen zwischen B_2 und B'_2 . Außerdem sei A die Matrixdarstellung von f bezüglich B_1 und B_2 , A' jene bezüglich B'_1 und B'_2 , überdies $\mathbf{y} = A\mathbf{x}$ und $\mathbf{y}' = A'\mathbf{x}'$. Dann zeigen analoge Überlegungen wie oben

$$A'\mathbf{x}' = \mathbf{y}' = S_2\mathbf{y} = S_2A\mathbf{x} = S_2AS_1^{-1}\mathbf{x}',$$

also

$$A' = S_2AS_1^{-1} \quad \text{bzw.} \quad A = S_2^{-1}A'S_1.$$

1.4 Multilinearität

Lineare Abbildungen, wie wir sie in Abschnitt 1.2 kennengelernt haben, hängen von nur einer Variablen ab. Kaum weniger wichtig sind Abbildungen, die von zwei oder mehreren Komponenten abhängen, und in jeder Komponente linear sind. Man spricht von multilinearen Abbildungen. Ähnlich wie die linearen Abbildungen von einem festen Vektorraum in einen anderen, bilden für gegebene Vektorräume V_i auch die multilinearen Abbildungen $m : V_1 \times \dots \times V_n \rightarrow V_0$ in natürlicher Weise wieder einen Vektorraum. Als besonders interessant werden sich später gewisse Multilinearformen (skalarwertige multilineare Abbildungen, d.h. mit $\dim V_0 = 1$) erweisen, die symmetrisch oder schief-symmetrisch sind. Die Grundlegenden Definitionen sind Inhalt von 1.4.1, in 1.4.2 werden

überblicksartig einige Beispiele vorgestellt. Vertieft wird in 1.4.3 der Begriff des Skalarproduktes, der von den Standardvektorräumen \mathbb{R}^n schon aus Mathematik 1 bekannt ist. Abstrakt gesprochen handelt es sich bei einem Skalarprodukt um eine Bilinearform (zweistellige Multilinearform), die zusätzlich symmetrisch und positiv definit ist. Studiert man Skalarprodukte, so werden automatisch auch orthogonale Transformationen und Matrizen interessant, denen 1.4.4 gewidmet ist. Im letzten Unterabschnitt 1.4.5 geht es um Determinanten. Auf Vektoren des \mathbb{R}^n angewendet, ist die Determinante jene (eindeutig bestimmte) n -stellige schiefsymmetrische Multilinearformen mit $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$ (kanonische Basis). In natürlicher Weise kann damit auch $\det(A)$ für eine $n \times n$ -Matrix A und $\det(f)$ für eine Lineartransformation $f : V \rightarrow V$ eines endlichdimensionalen Vektorraumes V definiert werden.

1.4.1 Multilineare Abbildungen

Inhalt in Kurzfassung: Nach einer kurzen Motivation (Vergleich Summe und Produkt mehrerer reeller Zahlen) definieren wir den Begriff der multilinearen Abbildung. Eine solche hängt von n Vektoren (eventuell aus verschiedenen Vektorräumen V_i) ab, wobei diese Abhängigkeit in jeder einzelnen Komponente linear ist. In ähnlicher Weise, wie sich lineare Abbildungen durch (zweidimensional anschreibbare) Matrizen darstellen lassen, kann man auch multilineare Abbildungen durch höherdimensionale (also im Allgemeinen nicht so angenehm anschreibbare) Zahlenschemata beschreiben, sofern Basen der involvierten Vektorräume vorliegen. Von besonderem Interesse sind (für $V_1 = \dots = V_n$) Multilinearformen (deren Werte eindimensional, also Zahlen sind) sowie symmetrische und anti- oder schiefsymmetrische multilineare Abbildungen.

Um den Unterschied zwischen linearen und multilinearen Abbildungen zu verdeutlichen, wählen wir eingangs das sehr einfache Beispiel der Summe s bzw. des Produktes p von n Zahlen $x_1, \dots, x_n \in \mathbb{R}$, beide aufgefasst als Abbildungen $s, p : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$s : (x_1, \dots, x_n) \mapsto x_1 + \dots + x_n, \quad p : (x_1, \dots, x_n) \mapsto x_1 \cdot \dots \cdot x_n$$

Fassen wir $\mathbf{x} = (x_1, \dots, x_n)$ bzw. $\mathbf{y} = (y_1, \dots, y_n)$ als Vektoren im Standardvektorraum \mathbb{R}^n auf, so gilt

$$s(\mathbf{x} + \mathbf{y}) = s(\mathbf{x}) + s(\mathbf{y}) \quad \text{und} \quad s(r\mathbf{x}) = rs(\mathbf{x})$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und $r \in \mathbb{R}$. Folglich ist s linear. Das gilt *nicht* für p . Zum Beispiel ist für $n = 2$ und $\mathbf{x} = (1, 1)$

$$p(\mathbf{x} + \mathbf{x}) = p(2\mathbf{x}) = p(2, 2) = 2 \cdot 2 = 4 \neq 2 = p(1, 1) + p(1, 1) = p(\mathbf{x}) + p(\mathbf{x}).$$

Trotzdem hat p eine Linearitätseigenschaft, nämlich bezogen auf jede der Komponenten, wenn die anderen fest bleiben. Für die erste Komponente:

$$\begin{aligned} p(x_1 + x'_1, x_2, \dots, x_n) &= (x_1 + x'_1)x_2 \dots x_n = x_1x_2 \dots x_n + x'_1x_2 \dots x_n = \\ &= p(x_1, x_2, \dots, x_n) + p(x'_1, x_2, \dots, x_n), \end{aligned}$$

und

$$p(rx_1, x_2, \dots, x_n) = (rx_1) \cdot x_2 \cdot \dots \cdot x_n = r(x_1 \cdot x_2 \cdot \dots \cdot x_n) = rp(x_1, x_2, \dots, x_n),$$

analog für die anderen. Man beachte auch, dass s diese Eigenschaft nicht hat, weil zum Beispiel ($n = r = 2$, $x_1 = x'_1 = x_2 = 1$ setzen) $s(2, 1) = 3 \neq 4 = 2s(1, 1)$. Lediglich für $n = 1$ fallen Linearität und Multilinearität zusammen, und es gilt sogar $s = p$.

Die beschriebene Eigenschaft von p verallgemeinernd – und zwar übernehmen nun beliebige (reelle) Vektorräume die diversen Rollen von \mathbb{R} – sprechen wir folgende Definition aus:

Definition 1.4.1.1. Für $n \in \mathbb{N}$ seien $V_1, V_2, \dots, V_n, V_0$ Vektorräume. Dann heißt eine Abbildung

$$m : V_1 \times V_2 \times \dots \times V_n \rightarrow V_0$$

multilinear, wenn m in jeder Komponente linear ist, d.h. wenn für alle $j = 1, \dots, n$ und für jede feste Wahl von Vektoren $\mathbf{a}_{j'} \in V_{j'}$, $1 \leq j' \leq n$, $j' \neq j$, die Abbildung $m_j : V_j \rightarrow V_0$ (eigentlich $m_j = m_{j, \mathbf{a}_1, \dots, \mathbf{a}_{j-1}, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n}$), gegeben durch

$$m_j : \mathbf{x} \mapsto m(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{j-1}, \mathbf{x}, \mathbf{a}_{j+1}, \mathbf{a}_{j+2}, \dots, \mathbf{a}_n)$$

linear von V_j nach V_0 ist. Die Zahl n heißt auch die **Stelligkeit** der multilinearen Abbildung m . Entsprechend heißt m auch **n -stellig**.

Ist $V_0 = \mathbb{R}$ (eindimensional, bei komplexen Vektorräumen ist $V_0 = \mathbb{C}$ zu setzen, für Vektorräume über einem anderen Körper entsprechend), so heißt m auch eine (n -stellige)

Multilinearform auf V . Eine **Bilinearform** ist eine 2-stellige Multilinearform.

Wir wissen: Lineare Abbildungen können mit Hilfe von Basen durch Matrizen dargestellt werden und bilden selbst wieder Vektorräume. Das beruht auf ganz ähnlichen Überlegungen, wie wir sie nun über multilineare Abbildungen anstellen wollen.

Seien dazu $d_j := \dim V_j$ die Dimensionen der V_j und $B_j = (\mathbf{b}_{j,1}, \dots, \mathbf{b}_{j,d_j})$ für $j = 0, 1, \dots, n$, Basen der V_j . Dann hat jeder Vektor $\mathbf{a}_j \in V_j$ bezüglich B_j eine eindeutige Darstellung $\mathbf{a}_j = \sum_{i_j=1}^{d_j} a_{j,i_j} \mathbf{b}_{j,i_j}$ als LK. Ist $m : V_1 \times V_2 \times \dots \times V_n \rightarrow \mathbb{R}$ multilinear, dann ergibt das analoge Argument wie bei linearen Abbildungen, diesmal jedoch iteriert auf jede der n Komponenten angewendet:

Proposition 1.4.1.2. Mit obigen Bezeichnungsweisen gilt für eine multilineare Abbildung m :

$$m(\mathbf{a}_1, \dots, \mathbf{a}_n) = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} \dots \sum_{i_n=1}^{d_n} a_{1,i_1} a_{2,i_2} \dots a_{n,i_n} m(\mathbf{b}_{1,i_1}, \mathbf{b}_{2,i_2}, \dots, \mathbf{b}_{n,i_n}).$$

Zwecks einfacherer Notation verwenden wir das Symbol t für Tupel $t = (i_1, \dots, i_n)$ mit $1 \leq i_j \leq d_j$ für $j = 1, \dots, n$ und fassen alle solchen t zur Menge T zusammen, die $d := |T| = d_1 \cdot d_2 \cdot \dots \cdot d_n$ Elemente hat. Noch etwas verkürzend schreiben wir

1 Lineare Algebra

$m_t := m(\mathbf{b}_{1,i_1}, \mathbf{b}_{2,i_2}, \dots, \mathbf{b}_{n,i_n})$. Dann ist die multilineare Abbildung m gemäß obiger Formel durch die Werte m_t , $t \in T$, eindeutig festgelegt. Jeder dieser Werte hat eine Darstellung

$$m_t = \sum_{i_0=1}^{d_0} r_{t,i_0} \mathbf{b}_{0,i_0} \in V_0$$

mit eindeutigen $r_{t,i_0} \in \mathbb{R}$.

Umgekehrt gibt es zu jeder Wahl von Zahlen $r_{t,i_0} \in \mathbb{R}$ mit $t \in T$ und $1 \leq i_0 \leq d_0$ genau eine multilineare Abbildung $m : V_1 \times \dots \times V_n \rightarrow V_0$, die diese Gleichung erfüllt, nämlich die durch

$$m(\mathbf{a}_1, \dots, \mathbf{a}_n) = m \left(\sum_{i_1=1}^{d_1} a_{1,i_1} \mathbf{b}_{1,i_1}, \dots, \sum_{i_n=1}^{d_n} a_{n,i_n} \mathbf{b}_{n,i_n} \right) := \sum_{i_0=1}^{d_0} \left(\sum_{t \in T} r_{t,i_0} \right) \mathbf{b}_{0,i_0}$$

definierte.

Übungsaufgabe 59. (E) Begründen Sie, warum die so definierte Abbildung

$$m : V_1 \times \dots \times V_n \rightarrow V_0$$

tatsächlich multilinear ist.

Bezeichne $M(V_1, \dots, V_n; V_0)$ die Menge aller solchen multilinearen Abbildungen $m : V_1 \times \dots \times V_n \rightarrow V_0$. In natürlicher Weise können zwei solche Abbildungen $m_1, m_2 \in M(V_1, \dots, V_n; V_0)$ addiert und mit Skalaren so multipliziert werden, dass $M(V_1, \dots, V_n; V_0)$ selbst zu einem Vektorraum wird. Denn für $m, m_1, m_2 \in M(V_1, \dots, V_n; V_0)$ und $r \in \mathbb{R}$ liefern die Definitionen

$$\begin{aligned} (m_1 + m_2)(\mathbf{x}_1, \dots, \mathbf{x}_n) &:= m_1(\mathbf{x}_1, \dots, \mathbf{x}_n) + m_2(\mathbf{x}_1, \dots, \mathbf{x}_n), \\ (rm)(\mathbf{x}_1, \dots, \mathbf{x}_n) &:= r(m(\mathbf{x}_1, \dots, \mathbf{x}_n)) \end{aligned}$$

wieder multilineare Abbildungen $m_1 + m_2$ und rm derart, dass alle Gesetze eines Vektorraums erfüllt sind. Das nachzurechnen ist eine einfache Übungsaufgabe:

Übungsaufgabe 60. (E) Zeigen Sie, dass es sich bei $m_1 + m_2$ und rm tatsächlich wieder um multilineare Abbildungen handelt und zeigen Sie durch Nachrechnen der Vektorraumgesetze für diese Operationen, dass $M(V_1, \dots, V_n; V_0)$ auf diese Weise wirklich zu einem Vektorraum wird.

Wir bezeichnen mit m_{t,i_0} jene eindeutig bestimmte multilineare Abbildung mit $m_{t,i_0}(t') = \mathbf{b}_{i_0}$ für $t' = t$ und $m_{t,i_0}(t') = \mathbf{o}_0$ für $t' \neq t$ ($\mathbf{o}_0 = \text{Nullvektor in } V_0$). Dann besagen unsere bisherigen Überlegungen, dass jedes $m \in M(V_1, \dots, V_n; V_0)$ eine eindeutige Darstellung als LK der m_{t,i_0} ($t \in T$ und $i_0 \in \{1, 2, \dots, d_0\}$) hat, diese also eine Basis B von $M(V_1, \dots, V_n; V_0)$ bilden. Ihre Kardinalität ist $|B| = |T| \cdot d_0 = \prod_{j=0}^n d_j$. So wie jede lineare (= 1-stellige multilineare) Abbildung zwischen zwei Vektorräumen, von denen je eine Basis vorgegeben ist, durch eine Matrix ($1 + 1 = 2$ -dimensionales Zahlenschema) dargestellt werden kann, wären für n -stellige multilineare Abbildungen entsprechend $n + 1$ -dimensionale Zahlenschemata erforderlich, bestehend aus den Komponenten r_{t,i_0} (der Index t ist selbst n -dimensional). Zusammenfassend gilt:

Satz 1.4.1.3. Sind V_j , $j = 0, 1, \dots, n$, Vektorräume, so bildet die Menge

$$M(V_1, \dots, V_n; V_0)$$

aller multilinearen Abbildungen $m : V_1 \times \dots \times V_n \rightarrow V_0$ mit den oben definierten Operationen einen Vektorraum. Sind die V_j endlichdimensional, so auch $M(V_1, \dots, V_n; V_0)$ mit

$$\dim M(V_1, \dots, V_n; V_0) = \prod_{j=0}^n d_j.$$

Sind Basen $B_j = (\mathbf{b}_{j,1}, \dots, \mathbf{b}_{j,d_j})$ der V_j ($j = 0, 1, \dots, n$) gegeben, so sei T die Menge aller n -Tupel $t = (i_1, \dots, i_n)$ mit $1 \leq i_j \leq d_j$ und $1 \leq j \leq n$. Für jedes $t \in T$ und $i_0 \in \{1, 2, \dots, d_0\}$ gibt es genau eine Multilinearform $m_{t,i_0} \in M(V_1, \dots, V_n; V_0)$ mit

$$\begin{aligned} m_{t,i_0}(t') &= \mathbf{b}_{i_0} \text{ für } t' = t \text{ und} \\ m_{t,i_0}(t') &= \mathbf{o}_0 \text{ für } t' \neq t \end{aligned}$$

mit dem Nullvektor $\mathbf{o}_0 \in V_0$. Die Familie

$$B = (m_{t,i_0})_{t \in T, 1 \leq i_0 \leq d_0}$$

dieser m_{t,i_0} ist eine Basis B von $M(V_1, \dots, V_n; V_0)$.

Übungsaufgabe 61. (E) Begründen Sie nochmals Satz 1.4.1.3. Hinweis: Ein sauberer Beweis gelingt am besten mit Induktion nach n .

In der weiteren Folge werden wir uns vor allem mit multilinearen Abbildungen befassen, die eine der folgenden beiden Eigenschaften haben:

Definition 1.4.1.4. Eine n -stellige multilineare Abbildung $m \in M(V, V, \dots, V; V_0)$ (wo also $V_1 = V_2 = \dots = V_n = V$ alle gleich sind) heißt **symmetrisch**, wenn sich ihr Wert bei Vertauschung zweier Komponenten nicht ändert, wenn also für alle $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ und alle Indizes i, j mit $1 \leq i < j \leq n$ gilt:

$$\begin{aligned} m(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) &= \\ m(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_j, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) & \end{aligned}$$

Ändert sich dabei hingegen das Vorzeichen, gilt also

$$\begin{aligned} m(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) &= \\ -m(\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_j, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{j-1}, \mathbf{x}_i, \mathbf{x}_{j+1}, \dots, \mathbf{x}_n) & \end{aligned}$$

für alle $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ und alle $1 \leq i < j \leq n$, so heißt m **schief-symmetrisch**.

Symmetrie und Schief-symmetrie sind recht starke Bedingungen. Ist $B = B_1 = \dots = B_n = (\mathbf{b}_1, \dots, \mathbf{b}_m)$ eine Basis von $V = V_1 = \dots = V_n$, so genügen zur eindeutigen Festlegung einer symmetrischen multilinearen Abbildung m die Werte $m(\mathbf{b}_{j_1}, \mathbf{b}_{j_2}, \dots, \mathbf{b}_{j_n})$ für jene Index- n -Tupel $t = (i_1, \dots, i_n)$ mit $i_1 \leq \dots \leq i_n$. Für schief-symmetrisches m gilt

noch mehr. Denn wenn zwei Komponenten mit derselben Eintragung vertauscht werden, muss sich einerseits das Vorzeichen des Wertes von m ändern, andererseits muss dieser Wert gleich bleiben, was nur für den Nullvektor \mathbf{o}_0 in V_0 möglich ist. Also sind alle Index- n -Tupel redundant, wo zwei Indizes übereinstimmen, und es bleiben lediglich jene mit strikter Ungleichung $i_1 < \dots < i_n$ übrig. Vor allem bei Determinanten wird dieser Umstand bemerkenswerte Vereinfachungen nach sich ziehen.

1.4.2 Interessante Beispiele multilinearer Abbildungen

Inhalt in Kurzfassung: Bereits bekannt ist das Vektorprodukt in \mathbb{R}^3 , die anderen nun zu erwähnenden Beispiele sind durchwegs sogar Linearformen, d.h. skalarwertig: Tensoren, etwas spezieller Bilinearformen und schließlich – mit Verweis auf die restlichen Unterabschnitte, wo sie ausführlicher besprochen werden – Skalarprodukte und Determinanten.

In den folgenden Beispielen sei weiterhin die Notation $m : V_1 \times \dots \times V_n \rightarrow V_0$ für eine multilineare Abbildungen festgehalten.

Für $n = 1$ besteht $M(V_1, \dots, V_n; V_0) = M(V_1; V_0)$ genau aus den linearen Abbildungen von V_1 nach V_0 , die wir schon in 1.2 ausführlichst behandelt haben. Darauf müssen wir nicht nochmals eingehen.

Für $n = 2$ kennen wir bereits wichtige Beispiele, etwa Vektorprodukt und Skalarprodukt. Auf Skalarprodukten ist mit 1.4.3 ein eigener Unterabschnitt gewidmet. Daher zunächst zu den Vektorprodukten.

Das **Vektorprodukt**: Hier ist $V_1 = V_2 = V_0 = \mathbb{R}^3$. Das Vektorprodukt $\mathbf{x} \times \mathbf{y} \in \mathbb{R}^3$ ist, wie aus Mathematik 1 bekannt, für zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ wie folgt definiert:

$$\mathbf{x} \times \mathbf{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

Als Gedächtnisstütze: Die Eintragungen des resultierenden Vektors lassen sich auch als Determinanten

$$\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = x_1 y_3 - x_3 y_1, \quad \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} = x_3 y_1 - x_1 y_3 \quad \text{und} \quad \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1,$$

siehe auch 1.4.5, schreiben.

Man überzeugt sich unmittelbar, dass das Vektorprodukt linear in beiden Komponenten und schiefsymmetrisch ist. Mit diesen Eigenschaften ist es bereits eindeutig bestimmt, wenn die drei Werte $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ und $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$ gegeben sind.

Übungsaufgabe 62. (E) Begründen Sie nochmals in aller Ausführlichkeit die hier behaupteten Aussagen.

Tensoren: Dabei handelt es sich um Multilinearformen $m : V_1 \times \dots \times V_n \rightarrow \mathbb{R}$, wobei es einen gemeinsamen Vektorraum V gibt derart, dass für alle $i = 1, \dots, n$ entweder $V_i = V$ oder $V_i = V^* := L(V, \mathbb{R})$ ist. V^* heißt auch der **Dualraum** von V . So wie Multilinearformen generell, sind Tensoren durch ihre Werte auf einer Basis (die sogenannten **Komponenten** der Tensors) eindeutig bestimmt. Dabei genügt es eine Basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ für V vorzugeben, weil damit automatisch auch die zu B **duale Basis** $B^* = (\mathbf{b}_1^*, \dots, \mathbf{b}_n^*)$ vorliegt. Dabei sind die Linearformen $\mathbf{b}_i^* : V \rightarrow \mathbb{R}$ (eindeutig) definiert durch $\mathbf{b}_i^*(\mathbf{b}_j) := \delta_{i,j}$ (Kronecker- δ , = 1 für $i = j$ und = 0 sonst).

Anwendungen: Ein in Verbindung mit partiellen Ableitungen, siehe Abschnitt 2.2, im Bauingenieurwesen (Festigkeitslehre) wichtiges Beispiel mit $V = \mathbb{R}^3$ und $n = 2$ ist der Spannungstensor, der sich bezüglich einer gegebenen Basis (zum Beispiel der kanonischen) durch eine quadratische 3×3 -Matrix darstellen lässt.

Bilinearformen: Bei einer Bilinearform auf einem Vektorraum V handelt es sich um eine Multilinearform mit $n = 2$, $V_1 = V_2 = V$ und $V_0 = \mathbb{R}$, also um eine **bilineare** Abbildung $\beta : V \times V \rightarrow \mathbb{R}$. Gemäß der allgemeinen Theorie ist so ein β eindeutig festgelegt durch die Werte $\beta_{i,j} := \beta(\mathbf{b}_i, \mathbf{b}_j)$ für $1 \leq i, j \leq d$, sofern $B = (\mathbf{b}_1, \dots, \mathbf{b}_d)$ eine Basis des d -dimensionalen Vektorraums V ist. Diese Information lässt sich in einer quadratischen Matrix

$$Q = \begin{vmatrix} \beta_{1,1} & \beta_{1,2} & \dots & \beta_{1,d} \\ \beta_{2,1} & \beta_{2,2} & \dots & \beta_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{d,1} & \beta_{d,2} & \dots & \beta_{d,d} \end{vmatrix}$$

anschreiben. Die Doppelstriche (statt abgerundeter Klammern) sollen symbolisieren, dass durch Q keine lineare Abbildung zwischen zwei d -dimensionalen Räumen, sondern eben eine bilineare Abbildung auf einem d -dimensionalen Raum dargestellt wird. Für zwei beliebige Vektoren $\mathbf{x}, \mathbf{y} \in V$ mit den Darstellungen $\mathbf{x} = \sum_{i=1}^d x_i \mathbf{b}_i$ und $\mathbf{y} = \sum_{j=1}^d y_j \mathbf{b}_j$ ergibt sich

$$\beta(\mathbf{x}, \mathbf{y}) = \beta \left(\sum_{i=1}^d x_i \mathbf{b}_i, \sum_{j=1}^d y_j \mathbf{b}_j \right) = \sum_{i=1}^d \sum_{j=1}^d x_i y_j \beta_{i,j} = \sum_{i=1}^d \sum_{j=1}^d x_i \beta_{i,j} y_j = \mathbf{x}_Z Q \mathbf{y}_S$$

mit dem Matrizenprodukt aus dem Zeilenvektor $\mathbf{x}_Z = (x_1, \dots, x_d)$, der Matrix Q und

dem Spaltenvektor $\mathbf{y}_S = \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix}$. Die Abbildung $q = q_\beta : V \rightarrow \mathbb{R}$, $q_\beta(\mathbf{x}) := \beta(\mathbf{x}, \mathbf{x})$ heißt

die der Bilinearform β zugeordnete **quadratische Form**.

Skalarprodukte: Dabei handelt es sich um symmetrische Bilinearformen auf einem Vektorraum V , die darüber hinaus positiv definit sind. Ein Beispiel, das Standardskalarprodukt, ist bereits aus Mathematik 1 bekannt. Weil Skalarprodukten mit 1.4.3 ein eigener Unterabschnitt gewidmet sein wird, begnügen wir uns an dieser Stelle mit ihrer

Erwähnung.

Determinanten: Auch Determinanten werden noch ausführlich behandelt werden (siehe 1.4.5), weshalb es an dieser Stelle nur um die Einordnung in den Kontext allgemeiner multilinearer Abbildungen geht. Und zwar ist die Determinante eine n -stellige schiefsymmetrische Multilinearform auf einem Vektorraum V der Dimension $\dim V = n$.

1.4.3 Skalarprodukt

Inhalt in Kurzfassung: Das bereits aus Mathematik 1 bekannte Standardskalarprodukt auf den Standardvektorräumen (den sogenannten euklidischen Räumen) verallgemeinert versteht man unter einem Skalarprodukt auf einem beliebigen Vektorraum eine positiv definite symmetrische Bilinearform. Alleine aus diesen Eigenschaften folgen zum Beispiel die Cauchy-Schwarzsche und die Dreiecksungleichung. Das Orthogonalisierungsverfahren von Schmidt erlaubt es, aus beliebig vorgegebenen Vektoren orthogonale mit gleichbleibender linearer Hülle zu generieren. Insbesondere lässt sich aus einer beliebigen Basis eine Orthonormalbasis gewinnen. Mit einer solchen wird offensichtlich, dass jeder endlichdimensionale Vektorraum mit Skalarprodukt strukturgleich mit einem euklidischen ist.

Definitionsgemäß ist ein Skalarprodukt auf einem Vektorraum V eine symmetrische, positiv definite Bilinearform $V \times V \rightarrow \mathbb{R}$, die wir mit $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{xy}$ anschreiben, manchmal auch als $\mathbf{xy} = \mathbf{x} \cdot \mathbf{y}$ oder als $\mathbf{xy} = \langle \mathbf{x}, \mathbf{y} \rangle$. Explizit bedeuten die Linearitätsbedingungen für die erste Komponente

$$(\mathbf{x}_1 + \mathbf{x}_2)\mathbf{y} = \mathbf{x}_1\mathbf{y} + \mathbf{x}_2\mathbf{y} \quad \text{und} \quad (r\mathbf{x})\mathbf{y} = r(\mathbf{xy}),$$

für die zweite

$$\mathbf{x}(\mathbf{y}_1 + \mathbf{y}_2) = \mathbf{xy}_1 + \mathbf{xy}_2 \quad \text{und} \quad \mathbf{x}(r\mathbf{y}) = r(\mathbf{xy}).$$

Symmetrie bedeutet

$$\mathbf{xy} = \mathbf{yx} \quad \text{für alle } \mathbf{x}, \mathbf{y} \in V,$$

positive Definitheit

$$\mathbf{xx} > 0 \quad \text{für alle } \mathbf{x} \in V \setminus \{\mathbf{o}\}.$$

Wegen der positiven Definitheit ist auch die **Länge** oder **Norm**

$$\|\mathbf{x}\| := \sqrt{\mathbf{xx}}$$

für alle $\mathbf{x} \in V$ definiert. Damit definiert man auch den **Abstand** oder **Distanz**

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$$

zweier Vektoren $\mathbf{x}, \mathbf{y} \in V$.

Bereits aus Mathematik 1 ist das Beispiel des n -dimensionalen **euklidischen Standardvektorraums** $V = \mathbb{R}^n$ mit dem Skalarprodukt

$$\mathbf{xy} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \sum_{i=1}^n x_i y_i$$

bekannt. Als Multilinearform ist dieses Skalarprodukt eindeutig bestimmt durch seine Werte auf einer Basis, in diesem Fall durch die einfachst mögliche Festlegung $\mathbf{e}_i \mathbf{e}_j = 0$ für $i \neq j$ und $\mathbf{e}_i \mathbf{e}_i = 1$, $1 \leq i, j \leq n$, für die kanonischen Basisvektoren \mathbf{e}_i . Man rufe sich die geometrische Interpretation in Erinnerung: $\mathbf{xy} = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cos \alpha$, wobei $\|\mathbf{x}\|, \|\mathbf{y}\|$ die Längen der beiden Vektoren \mathbf{x} und \mathbf{y} sind und α der Winkel, den sie einschließen. Einfache geometrische Überlegungen zeigen sowohl die Multilinearität des Skalarproduktes wie auch die angegebenen Werte für die \mathbf{e}_i .

Für die **euklidischen Räume** (damit sind definitionsgemäß die Vektorräume \mathbb{R}^n zusammen mit dem Standardskalarprodukt gemeint) haben wir wichtige Ungleichungen bewiesen und dabei nur jene Eigenschaften verwendet, die Teil der Definition eines Skalarproduktes sind. Folglich gelten diese Ungleichungen auch in beliebigen Räumen mit Skalarprodukt. Wir fassen diese Eigenschaften und einige weitere zusammen:

Satz 1.4.3.1. *In einem Vektorraum V mit Skalarprodukt gilt für alle $\mathbf{x}, \mathbf{y} \in V$ die Cauchy-Schwarzsche Ungleichung:*

$$|\mathbf{xy}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|,$$

wobei Gleichheit $|\mathbf{xy}| = \|\mathbf{x}\| \cdot \|\mathbf{y}\|$ genau dann gilt, wenn die beiden Vektoren linear abhängig sind.

Außerdem gilt für alle $\mathbf{x}, \mathbf{y} \in V$ die **Dreiecksungleichung**

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,$$

wobei Gleichheit $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ genau dann gilt, wenn einer der beiden Vektoren ein nichtnegatives Vielfaches des anderen ist.

Darüber hinaus ist $\|r\mathbf{x}\| = |r| \cdot \|\mathbf{x}\|$ für alle $r \in \mathbb{R}$ und $\mathbf{x} \in V$.

Die Abstandsfunktion $d : V \times V \rightarrow \mathbb{R}$, $(\mathbf{x}, \mathbf{y}) \mapsto d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$ hat alle Eigenschaften einer **Metrik**:

- *Positivität:* $d(\mathbf{x}, \mathbf{y}) \geq 0$ und $d(\mathbf{x}, \mathbf{y}) = 0$ nur für $\mathbf{x} = \mathbf{y}$,
- *Symmetrie:* $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ und
- *Dreiecksungleichung:* $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$.

Übungsaufgabe 63. (E) Beweisen Sie Satz 1.4.3.1. Hinweis: Für die meisten Teilaussagen können Sie die entsprechenden Beweise fast unmittelbar aus Mathematik 1 übertragen.

Ein unendlichdimensionales Beispiel ist für $a < b$ der Vektorraum $V = C([a, b])$ aller stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$ mit dem **Standardskalarprodukt**

$$\langle f, g \rangle := \int_a^b f(x)g(x)dx$$

auf $C([a, b])$ oder

$$\langle f, g \rangle_\varphi := \int_a^b f(x)g(x)\varphi(x)dx$$

für $f, g \in C[a, b]$ mit einer festen Funktion $\varphi \in C[a, b]$, $\varphi > 0$, genannt auch **Gewichtsfunktion**.

Übungsaufgabe 64. (E)

1. Begründen Sie, warum es sich bei $C([a, b])$ wie oben tatsächlich um einen Raum mit Skalarprodukt $\langle f, g \rangle$ bzw. $\langle f, g \rangle_\varphi$ handelt.
2. Welche Gestalt nehmen für diesen Raum die Aussagen aus Satz 1.4.3.1 an?

In Anlehnung an die Situation bei den euklidischen Standardvektorräumen nennt man auch im allgemeinen Fall eines Vektorraumes V mit Skalarprodukt zwei Vektoren $\mathbf{x}, \mathbf{y} \in V$ **orthogonal** oder **normal** (zueinander oder aufeinander), wenn $\mathbf{x}\mathbf{y} = 0$ gilt. Die Wegen der Ungleichung von Cauchy-Schwarz aus 1.4.3.1 gibt es für $\mathbf{x} \neq \mathbf{o} \neq \mathbf{y}$ eine eindeutig bestimmte Zahl $\alpha \in [0, \pi]$ mit

$$\cos \alpha = \frac{\mathbf{x}\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|},$$

genannt der **Winkel** zwischen den Vektoren $\mathbf{x}, \mathbf{y} \neq \mathbf{o}$. Vektoren $\mathbf{x} \in V$ der Länge $\|\mathbf{x}\| = 1$ heißen **normiert**.

Eine Menge $B \subseteq V$, in der je zwei $\mathbf{x} \neq \mathbf{y} \in B$ orthogonal sind, heißt **Orthogonalsystem** (OS). Sind zusätzlich alle $\mathbf{x} \in B$ normiert so heißt B ein **Orthonormalsystem** (ONS). Eine Basis B von V , die zusätzlich ein Orthogonal- oder sogar ein Orthonormalsystem ist, heißt eine **Orthogonalbasis** (OB) beziehungsweise, sofern $\|\mathbf{x}\| = 1$ für alle $\mathbf{x} \in B$, eine **Orthonormalbasis** (ONB) von V .

Ist in einem n -dimensionalen Vektorraum V mit Skalarprodukt ein n -elementiges ONS $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ mit $\mathbf{b}_i \neq \mathbf{o}$ für $i = 1, \dots, n$ gegeben, so handelt es sich automatisch um eine Basis von V . Das folgt aus:

Proposition 1.4.3.2. In einem Vektorraum V mit Skalarprodukt ist jedes Orthogonalsystem aus Vektoren $\neq \mathbf{o}$ linear unabhängig.

Beweis. Seien die $\mathbf{b}_i \in V$ paarweise orthogonal und $\neq \mathbf{o}$ und

$$\mathbf{x} := \sum_{i=1}^n r_i \mathbf{b}_i$$

mit gewissen $r_i \in \mathbb{R}$. Aus der Annahme $\mathbf{x} = \mathbf{o}$ haben wir $r_i = 0$ für alle $i = 1, \dots, n$ zu folgern. Sei dazu $s_i := \mathbf{b}_i \mathbf{b}_i = \|\mathbf{b}_i\|^2 > 0$. Dann gilt tatsächlich

$$r_i = s_i^{-1} r_i \mathbf{b}_i \mathbf{b}_i = s_i^{-1} \mathbf{b}_i \left(\sum_{j=1}^n r_j \mathbf{b}_j \right) = s_i^{-1} \mathbf{b}_i \mathbf{x} = s_i^{-1} \mathbf{b}_i \mathbf{o} = 0.$$

Dabei wurde in der zweiten Umformung verwendet, dass \mathbf{b}_i auf alle \mathbf{b}_j mit $j \neq i$ orthogonal steht. \square

Ausgehend von einer beliebigen Basis von V lässt sich sehr leicht eine orthonormale konstruieren. Die übliche Methode, die man dazu benützt, ist das **Orthogonalisierungsverfahren von Schmidt**. Man geht aus von einer endlichen oder auch unendlichen Folge linear unabhängiger Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n, \dots$ und definiert rekursiv

$$\mathbf{y}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \quad \text{und} \quad \mathbf{y}_k = \frac{\tilde{\mathbf{y}}_k}{\|\tilde{\mathbf{y}}_k\|}, \quad \text{mit} \quad \tilde{\mathbf{y}}_k = \mathbf{x}_k - \sum_{i=1}^{k-1} (\mathbf{x}_k \cdot \mathbf{y}_i) \mathbf{y}_i.$$

Die so definierten \mathbf{y}_k bilden, wie man leicht nachprüft, ein ONS derart, dass für alle n die lineare Hülle der Vektoren $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ mit jener der Vektoren $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ übereinstimmt. Also ist $\{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ eine ONB von $\text{LH}(\mathbf{x}_1, \dots, \mathbf{x}_n)$.

Übungsaufgabe 65. (E) Prüfen Sie diese Behauptung nach.

Ist ein ONS B sogar eine Basis von V , so erfüllen die Vektoren aus B dieselben Orthogonalitätsbeziehungen wie die die kanonischen Basisvektoren im Standardvektorraum \mathbb{R}^n . Konkret bedeutet das für eine Orthonormalbasis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$: Je zwei Vektoren $\mathbf{x}, \mathbf{y} \in V$ haben Darstellungen

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{b}_i \quad \text{und} \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{b}_i$$

mit eindeutig bestimmten $x_i, y_i \in \mathbb{R}$. Ihr Skalarprodukt ergibt sich wegen der Multilinearität daraus nach bewährtem Muster als

$$\mathbf{x} \mathbf{y} = \left(\sum_{i=1}^n x_i \mathbf{b}_i \right) \left(\sum_{i=1}^n y_i \mathbf{b}_i \right) = \sum_{i=1}^n x_i y_i,$$

genauso – so wie für die Vektoren $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ bezüglich des Standardskalarproduktes. In diesem Sinn hat daher jeder beliebige n -dimensionale Vektorraum mit Skalarprodukt die gleiche Struktur wie der n -dimensionale euklidische Standardvektorraum, und zwar nicht bloß als Vektorraum, sondern eben auch bezüglich des Skalarproduktes. In diesem Sinn repräsentieren die euklidischen Standardvektorräume sämtliche endlichdimensionalen reellen Vektorräume mit Skalarprodukt.

Übungsaufgabe 66. (T) Wenden Sie das Orthogonalisierungsverfahren von Schmidt auf $V = \mathbb{R}^n$ und die Vektoren $\mathbf{x}_1 = (1, 0, 0, 0)$, $\mathbf{x}_2 = (1, 1, 0, 0)$, $\mathbf{x}_3 = (1, 1, 1, 0)$ und $\mathbf{x}_4 = (1, 1, 1, 1)$ an.

Übungsaufgabe 67. (T*) Man kann das Orthogonalisierungsverfahren von Schmidt auch im unendlichdimensionalen Raum $C([-1, 1])$ mit dem oben eingeführten Skalarprodukt

$$\langle f, g \rangle := \int_a^b f(x)g(x)dx$$

einsetzen und als \mathbf{x}_n die Funktionen $f_n(x) := x^n$, beginnend mit $n = 0$, wählen. Berechnen Sie die daraus resultierenden \mathbf{y}_n für $n = 0, 1, 2, 3$.

Bemerkung zur Übungsaufgabe 67: Man erhält auf diese Weise im Wesentlichen die sogenannten Legendre-Polynome. Diese bilden sogar ein **vollständiges Orthonormalsystem**. In einem unendlichdimensionalen Raum mit Skalarprodukt entspricht dieser Begriff ungefähr dem einer Orthonormalbasis in endlichdimensionalen Fall. Allerdings muss man endliche Linearkombination durch unendliche konvergente Reihen ersetzen.

Bemerkung zu Vektorräumen über anderen Körpern als \mathbb{R} : Bilinearformen und weitgehend auch Skalarprodukte lassen sich genauso für Vektorräume über einem anderen Körper als \mathbb{R} definieren, etwa für komplexe Vektorräume. Die Positivitätsbedingung ist dann i.a. aber nicht mehr sinnvoll. Eine wichtige Variante ist der Begriff des **unitären** komplexen Vektorraums. Ihm liegt statt einer Bilinearform eine sogenannte **Hermiteische Form** zugrunde, die nicht symmetrisch ist, sondern $\mathbf{y}\mathbf{x} = \overline{\mathbf{x}\mathbf{y}}$ (konjugiert komplexe Zahl) erfüllt, außerdem $\mathbf{x}(c\mathbf{y}) = \overline{c}\mathbf{x}\mathbf{y}$ statt $\mathbf{x}(c\mathbf{y}) = c\mathbf{x}\mathbf{y}$ für $c \in \mathbb{C}$. Wegen

$$\mathbf{x}(c\mathbf{y}) = \overline{(c\mathbf{y})\mathbf{x}} = \overline{c(\mathbf{y}\mathbf{x})} = \overline{c}\overline{\mathbf{y}\mathbf{x}} = \overline{c}\mathbf{y}\mathbf{x} \neq c\mathbf{x}\mathbf{y}$$

für $c \in \mathbb{C} \setminus \mathbb{R}$ und $\mathbf{x}\mathbf{y} \neq 0$ sind Hermiteische Formen in der zweiten Komponente nicht linear, also auch keine Multilinearformen in unserem Sinn. Trotzdem ist wegen $\mathbf{x}\mathbf{x} = \overline{\mathbf{x}\mathbf{x}}$ das Produkt eines Vektors mit sich selbst reell. Wie im euklidischen Fall fordert man Positivität für $\mathbf{x} \neq \mathbf{o}$ und kann eine sehr ähnliche Theorie entwickeln. Hier werden wir diese Richtung aber nicht weiter verfolgen und begnügen uns mit dieser Andeutung.

Übungsaufgabe 68. (E) Betrachten Sie die Menge \mathcal{M} aller 2×2 Matrizen als Vektorraum. Durch $\langle A, B \rangle := \text{Tr}(AB)$ wird „fast“ ein Skalarprodukt auf \mathcal{M} definiert. Berechnen Sie $\text{Tr}(B_i B_j)$ mit $1 \leq i, j \leq d$ für eine selbstgewählte Basis von \mathcal{M} .

Übungsaufgabe 69. (E) Wie Aufgabe 68 aber mit 3×3 Matrizen.

Übungsaufgabe 70. (T) Entscheiden Sie jeweils, ob ein Skalarprodukt auf $C[-1, 1]$ vorliegt:

$$\langle f, g \rangle_0 = \int_{-1}^1 f(x)g(x) dx \quad \langle f, g \rangle_1 = \int_{-1}^1 x f(x)g(x) dx \quad \langle f, g \rangle_2 = \int_{-1}^1 x^2 f(x)g(x) dx.$$

Übungsaufgabe 71. (T^*) Finden Sie Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^2$, sodass

$$\vec{a} \cdot \vec{a} = \langle 1, 1 \rangle, \quad \vec{b} \cdot \vec{b} = \langle x, x \rangle, \quad \vec{a} \cdot \vec{b} = \langle 1, x \rangle,$$

wobei \cdot das gewöhnliche Skalarprodukt auf \mathbb{R}^2 und $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf $C[0, 1]$ ist. 1 und x sind als Elemente von $C[0, 1]$ zu verstehen.

Hinweis: $\cos 60^\circ = \frac{1}{2}$.

Übungsaufgabe 72. (T^*)

Wie Aufgabe 71 aber mit Vektoren $\vec{a}, \vec{b} \in C[0, 1]$.

Übungsaufgabe 73. (T) Wenden Sie das Orthogonalisierungsverfahren von Schmidt an, um ausgehend von der Basis

$$\{1, 1 + x, 1 + x + x^2\}$$

eine Orthonormalbasis des Vektorraumes $P_2(x)$ der quadratischen Polynome auf $[-1, 1]$ zu berechnen. Dabei ist das innere Produkt $\langle p, q \rangle$ wie üblich definiert durch

$$\langle p, q \rangle = \int_{-1}^1 p(x)q(x)dx.$$

1.4.4 Orthogonale Matrizen

Inhalt in Kurzfassung: Ähnlich, wie der Begriff der linearen Abbildung fast zwangsläufig mit der Vektorraums einhergeht, stößt man, wenn ein Skalarprodukt gegeben ist, auf Orthogonalität als interessante stärkere Eigenschaft von linearen Abbildungen. Denn mit jedem Skalarprodukt sind auch geometrische Größen wie Längen und Winkel definiert. Also sind jene linearen Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ von besonderer Bedeutung, die diese Größen nicht verändern, wie zum Beispiel Drehungen. Man nennt sie orthogonal. Quadratische Matrizen nennt man orthogonal, wenn sie bezüglich der kanonischen Basis orthogonale Transformationen des euklidischen Standardvektorraums darstellen.

Allgemein definieren wir:

Definition 1.4.4.1. Sind V_1 und V_2 Vektorräume mit Skalarprodukt, so heißt eine lineare Abbildung $f : V_1 \rightarrow V_2$ **orthogonal**, wenn für alle $\mathbf{x}, \mathbf{y} \in V$

$$f(\mathbf{x})f(\mathbf{y}) = \mathbf{xy}$$

gilt. Insbesondere bleibt dann auch die Norm unverändert:

$$\|f(\mathbf{x})\| = \sqrt{f(\mathbf{x})f(\mathbf{x})} = \sqrt{\mathbf{xx}} = \|\mathbf{x}\|$$

Eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt **orthogonal**, wenn sie bezüglich des Standardskalarproduktes orthogonal ist. Eine $n \times n$ -Matrix A heißt **orthogonal**, wenn A bezüglich der kanonischen Basis eine orthogonale Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ darstellt.

Man beachte, dass eine orthogonale Abbildungen $f : V_1 \rightarrow V_2$ stets injektiv ist: Aus $f(\mathbf{x}) = \mathbf{o}_2$ folgt $\|\mathbf{x}\| = \|f(\mathbf{x})\| = 0$, also $\mathbf{x} = \mathbf{o}_1$. Ist $\dim V_1 = \dim V_2 = n$ endlich, dann muss so ein f sogar bijektiv sein. Insbesondere ist das für $V_1 = V_2 = \mathbb{R}^n$ der Fall. Folglich ist jede orthogonale Matrix regulär.

Leicht lässt sich einsehen, worauf es bei orthogonalen Matrizen A ankommt. Seien $\mathbf{a}_1, \dots, \mathbf{a}_n$ die Spaltenvektoren von A . Weil diese ja auch die Bildvektoren der kanonischen Basisvektoren \mathbf{e}_i sind und das Skalarprodukt erhalten bleiben soll, muss $\mathbf{a}_i \mathbf{a}_j = \mathbf{e}_i \mathbf{e}_j = \delta(i, j)$ (Kronecker- δ) gelten, also müssen die \mathbf{a}_i eine Orthonormalbasis bilden. Weil die Spaltenvektoren von A gleichzeitig die Zeilenvektoren der Transponierten A^T sind, sind die $\mathbf{a}_i \mathbf{a}_j = \delta(i, j)$ gleichzeitig die Eintragungen der Matrix $A^T A$, in Matrixschreibweise: $A^T A = I_n$, folglich $A^T = A^{-1}$. Dann gilt aber auch $(A^T)^T A^T = A A^T = (A^T A)^T = I_n^T = I_n$, folglich ist auch A^T orthogonal. Wir können wie folgt zusammenfassen:

Satz 1.4.4.2. Für eine reelle (n, n) -Matrix A sind folgende Aussagen äquivalent:

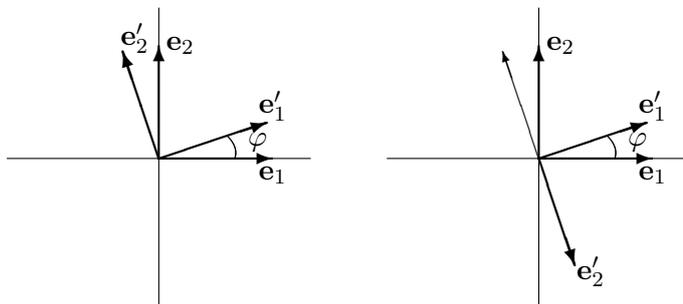
1. A ist orthogonal, stellt also definitionsgemäß eine orthogonale Transformation des Standardvektorraums dar.
2. Die Spaltenvektoren von A , aufgefasst als Vektoren des \mathbb{R}^n , bilden eine orthonormale Basis.
3. Die Zeilenvektoren von A , aufgefasst als Vektoren des \mathbb{R}^n , bilden eine orthonormale Basis.

Wichtige Beispiele orthogonaler Matrizen sind Drehungen und Drehspiegelungen: Die Matrizen

$$\begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}, \begin{pmatrix} \cos \varphi & \sin \varphi \\ \sin \varphi & -\cos \varphi \end{pmatrix} \quad (\varphi \in \mathbb{R} \text{ fest})$$

sind orthogonal. Die erste entspricht der Drehung um den Winkel φ , die zweite einer Drehspiegelung, d.h. der Verkettung einer Drehung und einer anschließenden Spiegelung. Man kann sogar leicht zeigen, dass alle orthogonalen $(2, 2)$ -Matrizen so aussehen:

Übungsaufgabe 74. (E) Sei A eine orthogonale 2×2 -Matrix. Zeigen Sie, dass es ein φ gibt, so dass A von einem der beiden angegebenen Typen, d.h. eine Drehung oder eine Drehspiegelung ist. Hinweis: $\det(A) = 1$ entspricht dem ersten Typ (Drehung), $\det(A) = -1$ dem zweiten (Drehspiegelung). Orientieren Sie sich an folgender Abbildung.



Übungsaufgabe 75. (T) Auch die (3,3)-Matrix

$$\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

ist orthogonal. Prüfen Sie das nach und beschreiben Sie die geometrische Wirkung der (bezüglich der kanonischen Basis) zugehörigen linearen Abbildung.

Übungsaufgabe 76. (T) Bestimmen Sie alle symmetrischen und orthogonalen Matrizen der Gestalt

$$\frac{1}{3} \begin{pmatrix} 1 & * & * \\ 2 & * & * \\ 2 & * & * \end{pmatrix}.$$

Übungsaufgabe 77. (P) Begründen Sie: Sind A und B orthogonale (n, n) -Matrizen, so auch AB , A^{-1} und A^T . Außerdem ist die Einheitsmatrix orthogonal.

Verzichtet man auf Linearität und begnügt sich damit, dass Abstände erhalten bleiben, so gelangt man zur Klasse der orthogonalen **Bewegungen**: Es sei A eine orthogonale (n, n) -Matrix und $\mathbf{a} \in \mathbb{R}^n$. Die Abbildung

$$(1) \quad \mathbf{f}: \mathbf{x} \rightarrow A\mathbf{x} + \mathbf{a} \quad \text{für } \mathbf{x} \in \mathbb{R}^n$$

hat dann, wie man leicht sieht, folgende Eigenschaft:

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n,$$

sie ist also **abstandstreu**. Solche Abbildungen nennt man **Bewegungen** oder **Isometrien**. Man kann zeigen, dass es außer den Bewegungen der Form (1) keine weiteren gibt. Bleibt bei einer Bewegung eine Gerade punktweise fest, so spricht man von einer **Drehung** oder **Rotation** mit dieser Geraden als **Rotationsachse**. Bleibt eine (Hyper-)Ebene punktweise fest, spricht man von einer **Spiegelung** an dieser Ebene.

Übungsaufgabe 78. (T) Weisen Sie nach, dass die folgenden Matrizen Drehungen entsprechen

$$T_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad T_2 = \frac{1}{15} \begin{pmatrix} 5 & 10 & 10 \\ -2 & 11 & -10 \\ -14 & 2 & 5 \end{pmatrix}$$

Übungsaufgabe 79. (T) Überprüfen Sie für welche Parameterwerte α die folgende Matrix eine Drehung im \mathbb{R}^3 beschreibt:

$$\frac{1}{9} \begin{pmatrix} 8 & 1 & \alpha \\ -\alpha & -4 & 7 \\ -1 & -8 & -4 \end{pmatrix}.$$

Projektionen, wo Bild und Kern orthogonal aufeinander stehen heißen **Orthogonalprojektionen**. Dazu noch einige Übungsaufgaben. Man beachte aber, dass Projektionen (außer im Trivialfall der Identität, die auf den gesamten Raum projiziert) selbst nicht orthogonal sind.

Übungsaufgabe 80. (E) Es sei $\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$ ein Einheitsvektor, also

$$\mathbf{n}^T \mathbf{n} = (n_1, n_2, n_3) \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = n_1^2 + n_2^2 + n_3^2 = \|\mathbf{n}\|^2 = 1.$$

Ferner sei E die Ursprungsebene mit Normaleneinheitsvektor \mathbf{n} , also die Menge der Punkte \mathbf{x} mit

$$\mathbf{n}^T \mathbf{x} = (n_1, n_2, n_3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = n_1 x_1 + n_2 x_2 + n_3 x_3 = 0.$$

Geben Sie die Matrix P der Orthogonalprojektion auf E sowie die Matrix S der Spiegelung an E an.

Hinweis zur Kontrolle: Es ergibt sich $P = I - \mathbf{nn}^T$, wobei

$$\mathbf{nn}^T = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} (n_1 n_2 n_3) = \begin{pmatrix} n_1 n_1 & n_1 n_2 & n_1 n_3 \\ n_2 n_1 & n_2 n_2 & n_2 n_3 \\ n_3 n_1 & n_3 n_2 & n_3 n_3 \end{pmatrix},$$

sowie $S = I - 2\mathbf{nn}^T$.

Anleitung: Man verwende den Ansatz $\mathbf{x}' - \mathbf{x} = \mathbf{n}\lambda$ und bestimme λ durch Linksmultiplikation mit \mathbf{n}^T (Man beachte: $\mathbf{x}' \in E!$). Wegen $\mathbf{x}'' = \mathbf{x}' + (\mathbf{x}' - \mathbf{x})$ erhält man daraus auch die Darstellung für S .

Übungsaufgabe 81. (E) Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ zwei linear unabhängige Vektoren. Bestimmen Sie die Orthogonalprojektion $P_{a,b} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ auf $\text{span}\{\mathbf{a}, \mathbf{b}\}$. Wie sieht die Matrixdarstellung von $P_{a,b}$ aus?

Übungsaufgabe 82. (T*) Die Orthogonal-Projektion auf die Ebene $z = 0$ lässt sich durch die Matrix

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

beschreiben. Bestimmen Sie die Matrix der Orthogonalprojektion auf die Ebene

$$E := \left\{ \lambda \begin{pmatrix} 7 \\ -4 \\ -4 \end{pmatrix} + \mu \begin{pmatrix} -4 \\ 1 \\ -8 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\},$$

indem Sie einen Basiswechsel $\mathbf{x}' = S^{-1}\mathbf{x}$ durchführen, sodass bzgl. der neuen Basis E die Gleichung $z' = 0$ hat.

Übungsaufgabe 83. (E) Sei $\mathbf{a} \in \mathbb{R}^3$ ein fester Vektor. Bestimmen Sie die Orthogonalprojektion $P_{\mathbf{a}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ auf den von \mathbf{a} erzeugten Unterraum. Wie sieht die Matrixdarstellung von $P_{\mathbf{a}}$ aus?

Übungsaufgabe 84. (T) Finden Sie die Orthogonal-Projektion von $v = (0, 3, 3) \in \mathbb{R}^3$ auf den Unterraum, der von den beiden (nicht orthonormalen!) Vektoren $\phi_1 = (1, 1, 0)$ und $\phi_2 = (1, 0, 2)$ aufgespannt wird.

Übungsaufgabe 85. (T) Finden Sie die Orthogonal-Projektion von $f(x) = 1 + e^x \in C[0, 1]$ auf den Unterraum, der von den beiden (nicht orthonormalen!) Vektoren $\phi_1(x) = 1$ und $\phi_2(x) = x$ aufgespannt wird.

Übungsaufgabe 86. (E) Rechnen Sie für $\vec{n} = e_1, e_2$ bzw. e_3 nach, dass eine Drehung um die Achse \vec{n} (mit $\vec{n} \cdot \vec{n} = 1$) um den Winkel ϕ durch die Abbildungsvorschrift

$$R_{\vec{n}, \phi}(x) = \vec{n}(\vec{n} \cdot x) + (\vec{n} \times x) \times \vec{n} \cos \phi + (\vec{n} \times x) \sin \phi$$

beschrieben wird. Skizze!

1.4.5 Determinanten

Inhalt in Kurzfassung: Abstrakt gesprochen ist die Determinante \det jene schiefsymmetrische Multilinearform auf \mathbb{R}^n in n Variablen, die für die kanonische Basis den Wert $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$ liefert. Dadurch ist \det eindeutig bestimmt. Für $n = 2$ lässt sich $\det(\mathbf{x}, \mathbf{y})$ auch als vorzeichenbehaftete Fläche des von \mathbf{x} und \mathbf{y} aufgespannten Parallelogramms deuten. Analog ist $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ für $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ das (ebenfalls vorzeichenbehaftete) Volumen des von diesen drei Vektoren aufgespannten Parallelepipeds. Die Formeln, die sich in den Dimensionen $n = 2, 3$ für die Determinante ergeben, lassen sich in natürlicher Weise auf beliebiges $n \in \mathbb{N}$ verallgemeinern. Für eine quadratische Matrix A definiert man $\det(A) = \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$ für die Spaltenvektoren \mathbf{a}_i von A . Die Zeilenvektoren liefern denselben Wert. Es gilt der Multiplikationssatz $\det(AB) = \det(A)\det(B)$. Aus ihm lässt sich folgern, dass für eine lineare Transformation $f : V \rightarrow V$ eines n -dimensionalen Vektorraums V und eine Matrixdarstellung A von f bezüglich einer Basis B von V der Wert $\det(f) := \det A$ nicht von der speziellen Wahl der Basis B abhängt. Der Betrag von $\det(f)$ ist, geometrisch gesprochen, jener multiplikative Faktor, mit dem n -dimensionale Volumina durch f vergrößert werden. Die am häufigsten verwendete Eigenschaft der Determinante besteht darin, dass sie genau dann den Wert $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$ liefert, wenn die Vektoren \mathbf{a}_i linear abhängig sind. Folglich ist $\det(A) = 0$ bzw. $\det(f) = 0$ genau dann, wenn A bzw. f singulär ist.

Die bereits in 1.4.2 erwähnte Determinante \det ist, ähnlich dem Skalarprodukt, eine Größe mit einer sehr attraktiven geometrischen Interpretation. Für $n = 2$ und $V = \mathbb{R}^2$ ist $\det(\mathbf{x}, \mathbf{y})$ die vorzeichenbehaftete Fläche des Parallelogramms, das die Vektoren \mathbf{x}

und \mathbf{y} aufspannen. Und zwar vereinbart man $\det(\mathbf{x}, \mathbf{y}) > 0$, wenn der Vektor \mathbf{x} auf kürzerem Wege im positiven Sinn in die Richtung von \mathbf{y} gedreht werden kann – so wie bei üblicher Bezeichnungsweise die x -Achse durch eine Drehung um den Winkel $\frac{\pi}{2}$ gegen den Uhrzeigersinn in die y -Achse übergeht. Andernfalls ist $\det(\mathbf{x}, \mathbf{y}) \leq 0$, wobei $\det(\mathbf{x}, \mathbf{y}) = 0$ genau dann gilt, wenn \mathbf{x} und \mathbf{y} kollinear sind. Insbesondere gilt $\det(\mathbf{x}, \mathbf{x}) = 0$ für alle $\mathbf{x} \in \mathbb{R}^2$. Eine einfache Zeichnung (Übung) bestätigt sofort die Linearitätsbeziehungen

$$\det(\mathbf{x}, \mathbf{y}_1 + \mathbf{y}_2) = \det(\mathbf{x}, \mathbf{y}_1) + \det(\mathbf{x}, \mathbf{y}_2) \quad \text{und} \quad \det(\mathbf{x}, r\mathbf{y}) = r \det(\mathbf{x}, \mathbf{y})$$

in der zweiten Komponente. Wegen der Vorzeichenvereinbarung gilt

$$\det(\mathbf{x}, \mathbf{y}) = -\det(\mathbf{y}, \mathbf{x}),$$

weshalb sich die Linearität von der zweiten auch auf die erste Komponente überträgt.

Folglich ist \det eine Bilinearform und als solche durch ihre Werte auf einer Basis B , zum Beispiel $B = (\mathbf{e}_1, \mathbf{e}_2)$ eindeutig bestimmt. Wegen der Schiefsymmetrie gilt $\det(\mathbf{e}_1, \mathbf{e}_1) = \det(\mathbf{e}_2, \mathbf{e}_2) = 0$ und $\det(\mathbf{e}_2, \mathbf{e}_1) = -\det(\mathbf{e}_1, \mathbf{e}_2)$, folglich hängt alles vom Wert $\det(\mathbf{e}_1, \mathbf{e}_2)$ ab. Das von \mathbf{e}_1 und \mathbf{e}_2 aufgespannte Parallelogramm ist ein Einheitsquadrat, und die Drehung von \mathbf{e}_1 um einen rechten Winkel auf \mathbf{e}_2 erfolgt im positiven Sinn. Also ist $\det(\mathbf{e}_1, \mathbf{e}_2) = 1$ zu setzen, und wir erhalten für zwei beliebige Vektoren $\mathbf{x} = (x_1, x_2), \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$:

$$\det(\mathbf{x}, \mathbf{y}) = x_1 y_1 \det(\mathbf{e}_1, \mathbf{e}_1) + x_1 y_2 \det(\mathbf{e}_1, \mathbf{e}_2) + x_2 y_1 \det(\mathbf{e}_2, \mathbf{e}_1) + x_2 y_2 \det(\mathbf{e}_2, \mathbf{e}_2),$$

also

$$\det(\mathbf{x}, \mathbf{y}) = \det \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) = x_1 y_2 - x_2 y_1.$$

Übungsaufgabe 87. (E) *Rekapitulieren Sie ausführlich, warum die angegebene geometrische Definition der Determinante für Vektoren im \mathbb{R}^2 genau zu dieser Formel führt.*

Ganz analog interpretiert man für $n = 3$ und $V = \mathbb{R}^3$ die Determinante $\det(\mathbf{x}, \mathbf{y}, \mathbf{z})$ dreier Vektoren $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ als vorzeichenbehaftetes Volumen des von \mathbf{x}, \mathbf{y} und \mathbf{z} aufgespannten Parallelepipeds. Das Vorzeichen wird positiv, wenn die drei Vektoren im Sinne einer Rechtsschraubenregel zueinander stehen, genauer: Denkt man sich die drei Vektoren orthogonal aufeinander, eine Schraube in Richtung von \mathbf{z} , so bewirkt eine Drehung von \mathbf{x} auf kürzestem Wege hin zu \mathbf{y} eine Bewegung der Schraube in Richtung \mathbf{z} . Alternativ denke man an eine rechte Hand, wo Daumen, Zeige und Mittelfinger orthogonal aufeinander stehen. Repräsentiert der Mittelfinger, so deutet er in jene Richtung, in die sich eine Schraube bewegen würde, wenn man sie so dreht, wie der Daumen hin zum Zeigefinger.

Leichter handhabbar wird die Vorzeichendefinition, wenn man ohne Bezugnahme auf geometrische Verhältnisse einfach nur formal $\det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = 1$ festlegt. Für die Vertauschungen ergibt sich über die Schiefsymmetrie (Veränderung des Vorzeichens bei der Vertauschung von zwei Komponenten):

$$\det(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3) = \det(\mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_1) = \det(\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2) = -\det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = -1$$

und bei nochmaliger geeigneter Vertauschung

$$\det(\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1) = \det(\mathbf{e}_3, \mathbf{e}_1, \mathbf{e}_2) = -\det(\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3) = -(-1) = 1.$$

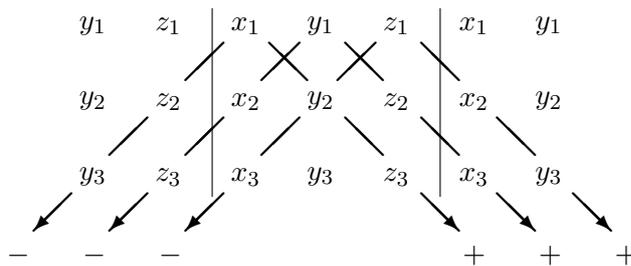
Bedenkt man $\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$, sofern von den drei Vektoren $\mathbf{x}, \mathbf{y}, \mathbf{z}$ zwei gleich sind, so haben wir alle Werte für die Basis beisammen, weshalb auch die dreidimensionale Determinante eindeutig bestimmt ist. Und zwar erhalten wir für die Determinante dreier beliebiger Vektoren $\mathbf{x} = (x_1, x_2, x_3), \mathbf{y} = (y_1, y_2, y_3), \mathbf{z} = (z_1, z_2, z_3) \in \mathbb{R}^3$ wieder unter Ausnutzung der Multilinearität:

$$\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \det\left(\sum_{i=1}^3 x_i \mathbf{e}_i, \sum_{j=1}^3 y_j \mathbf{e}_j, \sum_{k=1}^3 z_k \mathbf{e}_k\right) = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 x_i y_j z_k \det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$$

Nur jene drei Faktoren $\det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$ mit $\{i, j, k\} = \{1, 2, 3\}$, wo also kein Index doppelt vorkommt, sind von 0 verschieden, und zwar $= \pm 1$ (siehe oben). Schlussendlich erhalten wir also

$$\det(\mathbf{x}, \mathbf{y}, \mathbf{z}) = x_1 y_2 z_3 + x_2 y_3 z_1 + x_3 y_1 z_2 - x_3 y_2 z_1 - x_2 y_1 z_3 - x_1 y_3 z_2.$$

Eine einprägsame graphische Darstellung ist als **Merkregel von Sarrus** **Merkregel von Sarrus** bekannt:



Die Elemente längs eines Pfeils werden miteinander und außerdem mit 1 oder -1 multipliziert, je nach Pfeilrichtung. Dann addiert man die so erhaltenen Produkte.

Achtung: Die Regel von Sarrus ist nur für $n = 3$ definiert, hin und wieder auftauchende scheinbare Verallgemeinerungen für $n \geq 4$ mit Summen von 8 statt 6 Produkten sind falsch. Wie wir gleich sehen werden, sind nämlich $n!$, für $n = 4$ also $n! = 24$ Summanden erforderlich.

Für größere n lässt sich die Determinante von n Vektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathbb{R}^n$ mit $\mathbf{a}_j = (a_{1,j}, \dots, a_{n,j}), j = 1, \dots, n$, ganz analog motivieren bzw. definieren.⁴ Doch was heißt hier analog? In der Formel für $n = 3$ treten als Summanden Ausdrücke der Form $x_i y_j z_k \det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$ auf, wobei der Faktor $\det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$ lediglich für das Vorzeichen zuständig ist, und unter allen möglichen Produkten $x_i y_j z_k$ nur jene mit $\{i, j, k\} = \{1, 2, 3\}$

⁴Für $n = 1$ ist $V = \mathbb{R}$, die Vektoren \mathbf{x} sind reelle Zahlen und die Determinante ist die Identität: $\det(\mathbf{x}) = \mathbf{x}$.

wirklich auftreten. Bei unserer Notation für den allgemeinen Fall ist x_i durch $a_{i,1}$ zu ersetzen, y_j durch $a_{j,2}$ und z_k durch $a_{k,3}$. Die Zuordnung $1 \mapsto i$, $2 \mapsto j$ und $3 \mapsto k$ ist eine Permutation der Menge $\{1, 2, 3\}$ der verfügbaren Indizes. Das Vorzeichen ergibt sich daraus, wie viele Vertauschungen nötig sind, um von der natürlichen Reihenfolge 1-2-3 zur Reihenfolge $\pi(1)$ - $\pi(2)$ - $\pi(3)$ zu gelangen; ist eine einzige Vertauschung oder, allgemeiner, eine ungerade Anzahl von Vertauschungen erforderlich, so wird $\det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = -1$, andernfalls $\det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) = 1$. Man nennt deshalb π im ersten Fall eine **ungerade Permutation**, im zweiten eine **gerade Permutation** und schreibt auch $\operatorname{sgn}(\pi) = -1$ bzw. $\operatorname{sgn}(\pi) = 1$. Man nennt $\operatorname{sgn}(\pi)$ auch **Signum** von π . Wie üblich bezeichnen wir mit $\pi_2 \circ \pi_1$ die Verkettung der Permutationen π_1 (zuerst) und π_2 (anschließend), die durch $\pi_2 \circ \pi_1(i) := \pi_2(\pi_1(i))$ definiert ist. Offenbar gilt $\operatorname{sgn}(\pi_2 \circ \pi_1) = \operatorname{sgn}(\pi_2) \cdot \operatorname{sgn}(\pi_1)$ und $\operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\pi)$ für die zu π inverse Permutation π^{-1} . Aufgrund der letztgenannten Eigenschaft können wir ohne Vorzeichenwechsel statt über alle $a_{\pi(1),1} \cdots a_{\pi(n),n}$ genauso über alle $a_{1,\pi(1)} \cdots a_{n,\pi(n)}$ summieren. Somit ergibt sich als Formel für die Determinante der n Vektoren $\mathbf{a}_j \in \mathbb{R}^n$:

$$\det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{1\pi(1)} \cdot \dots \cdot a_{n\pi(n)}.$$

Die Summe erstreckt sich über alle Permutationen π der Menge $M_n = \{1, 2, \dots, n\}$. Per Definitionem bilden diese die Menge S_n , die man auch die **symmetrische Gruppe** vom Rang n nennt. Wie wir aus Mathematik 1 wissen, gilt $|S_n| = n!$. Die Bezeichnung Gruppe ist deshalb gerechtfertigt, weil bezüglich der Verkettung \circ alle an eine Gruppe gestellten Bedingungen erfüllt sind: Je zwei Elementen $\pi_1, \pi_2 \in S_n$ ist ein Element $\pi_1 \circ \pi_2 \in S_n$ zugeordnet, es gilt das Assoziativgesetz $(\pi_1 \circ \pi_2) \circ \pi_3 = \pi_1 \circ (\pi_2 \circ \pi_3)$ (wie generell für die Verkettung von Abbildungen), es gibt ein neutrales Element (nämlich die identische Abbildung auf M_n), und zu jedem $\pi \in S_n$ gibt es ein Inverses $\pi^{-1} \in S_n$, nämlich die Umkehrabbildung von π .

Eine kleine Lücke in den Begriffen *gerade* und *ungerade* Permutation ist noch auszuräumen. Und zwar wäre es a priori denkbar, dass eine Permutation π einerseits zum Beispiel durch drei, auf anderem Wege auch durch vier Vertauschungen erzeugbar ist. In diesem Fall wäre $\operatorname{sgn}(\pi)$ nicht eindeutig definiert und daher die Formel sinnlos. Tatsächlich ist das aber nicht möglich. Jedem $\pi \in S_n$ können wir nämlich die Anzahl $A(\pi)$ der sogenannten Fehlstände von π zuordnen. Das seien jene Paare (i, j) mit $1 \leq i < j \leq n$ und $\pi(i) > \pi(j)$, für die also π die Reihenfolge umdreht. Man überlegt sich schnell (Übung), dass eine Vertauschung die Anzahl der Fehlstände um genau 1 verändert, also von gerade auf ungerade oder umgekehrt. Weil $A(\pi)$ nur eine Möglichkeit hat, nämlich gerade oder ungerade zu sein, ist für dieses π auch die Anzahl der Vertauschungen, mit denen man π aus der identischen Permutation erzeugen kann, entweder gerade (wenn $A(\pi)$ gerade ist) oder ungerade (sonst), egal auf welche Weise man dies tut.

Übungsaufgabe 88. (E) Prüfen Sie obige Behauptung nach: Sei τ eine sogenannte *Transposition*, d.h. eine Vertauschung von zwei Elementen $i < j \in M_n$. Für jede Permutation $\pi \in S_n$ gilt dann $|A(\tau \circ \pi) - A(\pi)| = |A(\pi \circ \tau) - A(\pi)| = 1$.

Wir können nun die **Determinante einer quadratischen $n \times n$ -Matrix A** definieren, indem wir

$$\det(A) := \det(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n).$$

mit den Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ der Matrix A setzen. Manchmal schreibt man auch $|A|$ für $\det A$.

Hervorzuheben ist der folgende **Multiplikationssatz für Determinanten**:

Satz 1.4.5.1. *Für zwei quadratische $n \times n$ -Matrizen A und B gilt*

$$\det(AB) = \det(A) \det(B) \quad (\text{Multiplikativität der Determinante}).$$

Für die identische Matrix I_n ist $\det I_n = 1$. Ist A regulär mit Inverser A^{-1} folgt insbesondere $\det(A) \neq 0$ und $\det(A^{-1}) = (\det A)^{-1}$.

Beweis. Gegeben die Matrix A , so betrachten wir die Abbildung $B \mapsto \det(AB)$. Indem wir B als n -Tupel seiner Spaltenvektoren $\mathbf{b}_1, \dots, \mathbf{b}_n$ interpretieren, können wir diese Zuordnung auch auffassen als $m : V^n \rightarrow \mathbb{R}$ mit $V = \mathbb{R}^n$. Wie man sich schnell überlegt (Übung), ist m multilinear, also eindeutig festgelegt durch seine Werte $m(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n})$ mit $1 \leq i_1, i_2, \dots, i_n \leq n$. Außerdem vererbt sich die Schiefsymmetrie von \det auf m , also ist $m(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) = \text{sgn}(\pi)m(\mathbf{e}_1, \dots, \mathbf{e}_n)$ für Permutationen $\pi : j \mapsto i_j$ und $= 0$ sonst (d.h. wenn $j \mapsto i_j$ nicht bijektiv ist). Wegen $m(\mathbf{e}_1, \dots, \mathbf{e}_n) = \det A$, zeigt dies, dass sich die Multilinearform m von der Multilinearform \det (bei Identifikation von n -Tupeln von Spaltenvektoren mit der aus ihnen gebildeten Matrix) nur um den multiplikativen Faktor $\det A$ unterscheidet, also

$$\det(AB) = m(\mathbf{b}_1, \dots, \mathbf{b}_n) = \det(A) \det(B).$$

Weiters gilt tatsächlich $\det I_n = \det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$, woraus wegen der bewiesenen Multiplikativität von \det aus $AA^{-1} = I_n$ sofort

$$\det A \det A^{-1} = \det(AA^{-1}) = \det I_n = 1,$$

also $\det A \neq 0$ und $\det(A^{-1}) = (\det A)^{-1}$ folgt. □

Folgerung 1.4.5.2. *Sei V ein n -dimensionaler Vektorraum, $f : V \rightarrow V$ linear, B_1 und B_2 zwei Basen von V , A_1 die Matrixdarstellung von f bezüglich B_1 , A_2 bezüglich B_2 . Dann gilt $\det A_1 = \det A_2$.*

Beweis. Der Basiswechsel zwischen B_1 und B_2 wird gemäß 1.3.5 durch eine quadratische Matrix S mit Inverser S^{-1} beschrieben, so dass gilt $B_2 = SB_1S^{-1}$. Aus Satz 1.4.5.1 folgt damit

$$\det(B_2) = \det(SB_1S^{-1}) = \det(S) \det(B_1) \det(S)^{-1} = \det(B_1).$$

□

Hieraus ergibt sich die Invarianz der Determinante unter Basiswechsel: Weil also die Determinanten von Matrixdarstellungen einer linearen Abbildung $f : V \rightarrow V$ eines endlichdimensionalen Vektorraums bezüglich beliebiger Basen übereinstimmen, ist es sinnvoll auch von der **Determinante von f** , symbolisch $\det(f)$ oder $\det f$, zu sprechen.

Die Multiplikativität der Determinante vererbt sich natürlich von Matrizen auf lineare Abbildungen. Das hat eine sehr überzeugende geometrische Interpretation: Ist $V = \mathbb{R}^n$, so lässt sich $\det f$ als das n -dimensionale Volumen jenes Körpers (Parallelepiped) $f(W_n)$ auffassen, auf das der von den kanonischen Basisvektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$ aufgespannte n -dimensionale Einheitswürfel W_n durch f abgebildet wird. Die Zahl $\det(f)$ ist also jener Faktor, mit dem f Volumina vergrößert. Verkettet man mehrere Abbildungen, so müssen sich die zugehörigen Faktoren aufmultiplizieren, also $\det(g \circ f) = \det(g) \cdot \det(f)$ für $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Ein beliebiger n -dimensionaler Vektorraum V ist zu \mathbb{R}^n isomorph, also übertragen sich alle rein algebraisch beschreibbaren Sachverhalte (wie etwa die Definitionen der Determinante) von \mathbb{R}^n auf V .

Zur besseren Übersicht eine Zusammenfassung:

Satz 1.4.5.3. *Für jedes positive $n \in \mathbb{N}$ kann die Determinante in dreifacher Weise als Funktion aufgefasst und wie folgt charakterisiert werden.*

1. Als Funktion

$$\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}, \quad (\mathbf{a}_1, \dots, \mathbf{a}_n) \mapsto \det(\mathbf{a}_1, \dots, \mathbf{a}_n),$$

ist \det multilinear, schiefsymmetrisch und erfüllt $\det(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$. Diese Eigenschaften legen \det eindeutig fest. Und zwar gilt für $\mathbf{a}_j = (a_{1j}, \dots, a_{nj})$, $j = 1, \dots, n$, die Formel

$$\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{1\pi(1)}, \dots, a_{n\pi(n)}.$$

Geometrische Interpretation: $|\det(\mathbf{a}_1, \dots, \mathbf{a}_n)|$ ist das n -dimensionale Volumen (= Lebesguesches Maß) des von den Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ aufgespannten Parallelepipeds (der Menge aller Linearkombinationen $\sum_{i=1}^n r_i \mathbf{a}_i$ mit $0 \leq r_i \leq 1$).

2. Als Funktion

$$\det : \mathcal{M}(n, n) \rightarrow \mathbb{R}, \quad A \mapsto \det(A)$$

ordnet jeder quadratischen $n \times n$ -Matrix A die reelle Zahl $\det(\mathbf{a}_1, \dots, \mathbf{a}_n)$ (siehe Aussage 1) zu, wobei \mathbf{a}_j für $j = 1, \dots, n$ den j -ten Spaltenvektor von A bezeichnet. Und zwar gilt für $A = (a_{ij})$ die Formel

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{1\pi(1)}, \dots, a_{n\pi(n)}.$$

Geometrische Interpretation: $|\det A|$ ist der Faktor mit dem n -dimensionale Volumina durch die von A dargestellte lineare Abbildung vergrößert werden. Diese Zahl hängt nicht davon ab, ob man sich dabei auf die kanonische oder irgendeine andere Basis des \mathbb{R}^n bezieht.

3. Ist V irgendein n -dimensionaler Vektorraum, so ordnet \det als Funktion

$$\det : L(V, V) \rightarrow \mathbb{R}, \quad f \mapsto \det f,$$

einer linearen Abbildung $f : V \rightarrow V$ die Determinante $\det(A)$ (siehe Aussage 2) jener Matrix A zu, die f bezüglich irgendeiner Basis B von V darstellt. Diese Zahl hängt nicht von der speziellen Wahl der Basis B ab.

Geometrische Interpretation für $V = \mathbb{R}^n$: $|\det(f)|$ ist jener Faktor, um den f Volumina vergrößert.

Wie schon an früherer Stelle beobachtet, ändert sich am Wert der Summe in der zweiten Aussage von Satz 1.4.5.3 nichts, wenn man die Reihenfolge der Indizes, sprich Zeilen und Spalten vertauscht:

Proposition 1.4.5.4. Für eine quadratische Matrix A gilt $\det(A) = \det(A^T)$, Determinanten einer Matrix und ihrer Transponierten stimmen also überein.

Eine Folgerung betrifft orthogonale Matrizen A (siehe 1.4.4), für deren Transponierte A^T ja $AA^T = I_n$ gilt. Wegen $\det A^T = \det A$ und der Multiplikativität der Determinante folgt daraus

$$(\det A)^2 = \det A^T \cdot \det A = \det A^T A = \det I_n = 1,$$

also $\det A = \pm 1$.

Proposition 1.4.5.5. Für jede orthogonale Matrix A gilt $\det A = 1$ oder $\det A = -1$.

Eine der wichtigsten Eigenschaften der Determinante bezieht sich darauf, wann sie den Wert 0 annimmt. Um das zu sehen zunächst ein Hilfssatz.

Hilfssatz 1.4.5.6. Ist unter den Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$ ein \mathbf{a}_i Linearkombination der übrigen (sind die Vektoren also linear abhängig), so folgt $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$.

Beweis. Sei zunächst $\mathbf{a}_n = \sum_{j=1}^{n-1} r_j \mathbf{a}_j$ Linearkombination der anderen (also $i = n$). Weil \det in der letzten Komponente linear ist, gilt $\det(\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \mathbf{0}) = 0$. Wegen der Schiefsymmetrie beobachten wir $\det(\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \mathbf{a}_j) = 0$ für $j = 1, \dots, n-1$, weil die j -te und die n -te Komponente übereinstimmen. Wegen der Linearität in der letzten Komponente ändert sich daran nichts, wenn man diese mit dem Skalar r_j multipliziert, also $\det(\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, r_j \mathbf{a}_j) = 0$. Die Linearität von \det in der letzten Komponente erlaubt es, für $j = 1, \dots, n-1$ zu $\det(\mathbf{a}_1, \dots, \mathbf{a}_{n-1}, \sum_{j=1}^{n-1} r_j \mathbf{a}_j) = 0$ zu summieren. Also gilt die Behauptung des Hilfssatzes für $i = n$. Wegen der Schiefsymmetrie von \det folgt mittels Vertauschung die entsprechende Aussage für alle anderen Komponenten. \square

Satz 1.4.5.7. Sei wieder $n \in \mathbb{N}$ und $n > 0$.

1. Für Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$ ist $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$ genau dann, wenn die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ linear abhängig sind.

1 Lineare Algebra

2. Für eine quadratische $n \times n$ -Matrix A gilt $\det(A) = 0$ genau dann, wenn die Spaltenvektoren von A linear abhängig sind. Das ist genau dann der Fall, wenn A singulär ist.

3. Ist V ein n -dimensionaler Vektorraum und $f : V \rightarrow V$ eine lineare Abbildung, so gilt $\det(f) = 0$ genau dann, wenn $\text{rg}(f) < n$, wenn also f nicht bijektiv ist.

Beweis. Aufgrund des in Satz 1.4.5.3 dargestellten Zusammenhangs genügt es, eine der drei Behauptungen zu beweisen. Wir wählen die erste.

Sind die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ linear abhängig, so ist einer von ihnen als Linearkombination der anderen darstellbar. Also folgt $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = 0$ wegen Hilfssatz 1.4.5.6.

Sind die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ hingegen linear unabhängig, so bilden sie, wenn man sie als Spaltenvektoren zusammenfasst, eine reguläre $n \times n$ -Matrix A . Laut Folgerung 1.4.5.1 gilt daher $\det(\mathbf{a}_1, \dots, \mathbf{a}_n) = \det A \neq 0$. \square

Sprechweisen für Matrizen wie Zeilen, Spalten, Zeilenzahl, Elemente, etc. überträgt man in naheliegender Weise auf Determinanten.

Abschließend noch eine Bemerkung zur geometrischen Deutung des Vektorproduktes $\mathbf{c} := \mathbf{a} \times \mathbf{b}$ zweier Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, siehe auch 1.4.2. Bekanntlich (und wie sich sehr leicht nachrechnen lässt), steht \mathbf{c} normal sowohl auf \mathbf{a} als auch auf \mathbf{b} . Will man nachweisen, dass die Länge von \mathbf{c} mit dem Zahlenwert F der Fläche des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms übereinstimmt, genügt es $|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})| = \|\mathbf{c}\|^2$ nachzuweisen. Denn $|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})|$ ist in diesem Fall das Volumen eines Prismas mit \mathbf{a} und \mathbf{b} als Grundseiten und \mathbf{c} als Höhe, nach der Regel „Volumen = Grundfläche mal Höhe“ also $|\det(\mathbf{a}, \mathbf{b}, \mathbf{c})| = F\|\mathbf{c}\|$, woraus mit obigem $F = \|\mathbf{c}\|$ folgt. Die geometrische Anschauung lehrt, dass \mathbf{c} durch diese beiden Eigenschaften bis auf das Vorzeichen eindeutig bestimmt ist. Und dieses wird durch die Forderung $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) \geq 0$ festgelegt, welche anschaulich der Rechtsschraubenregel entspricht. All das lässt sich nun leicht im Rahmen einer Übungsaufgabe nachrechnen:

Übungsaufgabe 89. (T) Verwenden Sie die Definition des Vektorproduktes aus 1.4.2, um für $\mathbf{a}, \mathbf{b}, \mathbf{c} := \mathbf{a} \times \mathbf{b} \in \mathbb{R}^3$ folgende Gleichungen nachzurechnen:

1. $\mathbf{a} \cdot \mathbf{c} = 0$
2. $\mathbf{b} \cdot \mathbf{c} = 0$
3. $\det(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \|\mathbf{c}\|^2 \geq 0$

1.5 Lineare Gleichungssysteme

Lineare Gleichungssysteme in m Variablen sind gegeben durch eine $n \times m$ -Matrix $A = (a_{ij})$, die Koeffizientenmatrix des Systems, und eine rechte Seite $\mathbf{b} \in \mathbb{R}^n$, den inhomogenen Anteil. Dabei sind jene Vektoren $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$ gesucht, für die die Matrixbeziehung $A\mathbf{x} = \mathbf{b}$ gilt. Anders formuliert: Gesucht sind jene $\mathbf{x} \in \mathbb{R}^m$ mit

$f_A(\mathbf{x}) = \mathbf{b}$ für die durch A bezüglich der kanonischen Basen dargestellten linearen Abbildung $f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$. Aufgrund dieses Zusammenhangs (siehe auch 1.5.1) wird die bisher entwickelte Theorie der Linearen Algebra verfügbar. Im Zentrum des Abschnitts steht das Eliminationsverfahren von Gauß. Es dient nicht nur zu einer sehr befriedigenden Darstellung der Lösungsmenge eines beliebigen linearen Gleichungssystems (siehe 1.5.2 für den quadratischen regulären Fall, 1.5.5 für den allgemeinen). Es ermöglicht auch die Berechnung von inversen Matrizen (1.5.3) und Determinanten (1.5.4).

1.5.1 Problemstellung

Inhalt in Kurzfassung: Die Lösung eines linearen Gleichungssystems mit einer Koeffizientenmatrix A (linke Seite des Systems) und einem inhomogenen Anteil \mathbf{b} (rechte Seite des Systems) lässt sich deuten als das Aufsuchen des Urbildes von \mathbf{b} unter der linearen Abbildung, die durch A dargestellt wird. Geometrisch definiert jede lineare Gleichung in m Variablen eine $m - 1$ -dimensionale Hyperebene, folglich ist die Lösungsmenge eines linearen Gleichungssystems die Schnittmenge dieser Hyperebenen.

Ein **lineares Gleichungssystem** S aus n skalaren Gleichungen in den Variablen x_1, \dots, x_m ist gegeben durch eine $n \times m$ -Matrix $A = (a_{ij})$ (die **System-** oder **Koeffizientenmatrix** des Systems S) und einen Vektor $\mathbf{b} \in \mathbb{R}^n$, den **inhomogenen** Anteil des Systems. Die **Lösungsmenge** L_S von S ist definiert durch

$$L_S := \{\mathbf{x} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{b}\}.$$

Jedes $\mathbf{x} \in L_S$ heißt **Lösung** von S . In dieser Formel sind \mathbf{x} und \mathbf{b} als Spaltenvektoren mit den Komponenten x_1, \dots, x_m bzw. b_1, \dots, b_n zu lesen. L_S lässt sich auch als die Menge jener m -Tupel (x_1, x_2, \dots, x_m) von (in unserem Fall reellen) Zahlen x_i verstehen, für die die n Gleichungen

$$S : \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,m}x_m = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,m}x_m = b_2 \\ \vdots \\ a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,m}x_m = b_i \\ \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,m}x_m = b_n \end{cases}$$

erfüllt sind. Ersetzt man die rechte Seite \mathbf{b} durch den Nullvektor $\mathbf{o} \in \mathbb{R}^n$, so entsteht das zugehörige **homogene System** S_h , im Gegensatz zum **inhomogenen** System mit $\mathbf{b} \neq \mathbf{o}$.

Ist $f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ die lineare Abbildung, die durch A dargestellt wird, so ist $L_S = f_A^{(-1)}(\mathbf{b})$ das Urbild von \mathbf{b} unter f_A und $L_{S_h} = \ker(f_A)$ der Kern von f_A . Insbesondere ist L_{S_h} also ein Unterraum von \mathbb{R}^m . Die Lösungsmengen von homogenem und inhomogenem System stehen in folgendem einfachen Zusammenhang, den wir zwecks zahlreicher Anwendungen allgemein für lineare Abbildungen formulieren:

Proposition 1.5.1.1. Sei $f : V_1 \rightarrow V_2$ linear. Gibt es ein $\mathbf{x}_0 \in V_1$, das auf ein vorgegebenes $\mathbf{b} \in V_2$ abgebildet wird, das also $f(\mathbf{x}_0) = \mathbf{b}$ erfüllt, so gilt

$$L := f^{-1}(\mathbf{b}) = \{\mathbf{x} \in V_1 : f(\mathbf{x}) = \mathbf{b}\} = \mathbf{x}_0 + \ker(f) = \{\mathbf{x}_0 + \mathbf{x}_1 : \mathbf{x}_1 \in \ker(f)\}.$$

Beweis. Behauptet wird die Mengengleichheit $L = \mathbf{x}_0 + \ker(f)$. Für alle $\mathbf{x} \in L$ gilt

$$\mathbf{b} = f(\mathbf{x}) = f(\mathbf{x}_0 + \mathbf{x} - \mathbf{x}_0) = f(\mathbf{x}_0) + f(\mathbf{x} - \mathbf{x}_0) = \mathbf{b} + f(\mathbf{x} - \mathbf{x}_0),$$

folglich $f(\mathbf{x} - \mathbf{x}_0) = \mathbf{o}$, d.h. $\mathbf{x} - \mathbf{x}_0 \in \ker(f)$ oder $\mathbf{x} \in \mathbf{x}_0 + \ker(f)$. Diese Überlegung beweist die Inklusion $L \subseteq \mathbf{x}_0 + \ker(f)$. Umgekehrt wird jedes Element \mathbf{x} von der Gestalt $\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1$ mit $\mathbf{x}_1 \in \ker(f)$ auf

$$f(\mathbf{x}) = f(\mathbf{x}_0 + \mathbf{x}_1) = f(\mathbf{x}_0) + f(\mathbf{x}_1) = f(\mathbf{x}_0) + \mathbf{o} = \mathbf{b}$$

abgebildet, liegt also in L . □

Die Lösungsmenge L lässt sich also als der in irgendeinen Punkt $\mathbf{x}_0 \in L$ verschobene Unterraum $\ker(f)$ von V_1 verstehen. L lässt sich somit sinnvoll beschreiben, indem man erstens irgendeinen speziellen Lösungsvektor \mathbf{x}_0 (eine sogenannte **partikuläre Lösung**) angibt und zweitens eine Basis der Lösungsmenge $L_h := \ker(f)$ des zugehörigen homogenen Problems. Ist $f = f_A$, also das lineare System $S: A\mathbf{x} = \mathbf{b}$ gegeben, so wollen wir genau das als das Lösen von S verstehen: Angabe einer (partikulären) Lösung \mathbf{x}_0 und einer Basis von $L_h = L_{S_h}$, der Lösungsmenge des zugehörigen homogenen Systems.

Auch eine geometrische Deutung liegt nahe: Eine lineare Gleichung in m Variablen hat als Lösung eine $m - 1$ -dimensionale (Hyper-)Ebene im \mathbb{R}^m . Deshalb ist das Lösen eines linearen Gleichungssystems aus n Gleichungen gleichbedeutend mit der geometrischen Aufgabe, den Durchschnitt von n Ebenen im \mathbb{R}^m zu bestimmen. Die Anschauung lehrt, dass verschiedene Situationen paralleler oder nicht paralleler Ebenen auftreten können. Zum Beispiel sind drei Ebenen im \mathbb{R}^3 denkbar, von denen je zwei nicht parallel zueinander sind, von denen aber je zwei eine Schnittgerade haben, die zur dritten parallel ist. In diesem (nicht „typischen“) Fall gibt es keinen gemeinsamen Schnittpunkt und somit keine Lösung des zugrunde liegenden Gleichungssystems – im Gegensatz zur typischen Situation, dass drei windschief zueinander stehende Ebenen im Raum genau einen Schnittpunkt haben. Bei mehr als drei Variablen, also $n > 3$, und entsprechend höherdimensionalen geometrischen Deutungen sind auch noch kompliziertere Situationen möglich. Das ruft nach einer systematischen Vorgehensweise.

Mit dem Eliminationsverfahren von Gauß steht eine sehr effektive Methode zur Verfügung. Neben dem Lösen von linearen Gleichungssystemen dient sie auch der Berechnung von Inversen und von Determinanten quadratischer Matrizen. Wir beginnen mit einem recht einfachen Beispiel, anhand dessen bereits das Wichtigste sichtbar wird.

1.5.2 Ein einfaches Beispiel

Inhalt in Kurzfassung: Anhand eines einfachen Systems aus drei linearen Gleichungen in drei Variablen werden die wichtigsten Schritte im Eliminationsverfahren von Gauß

deutlich gemacht.

Man löse folgendes Gleichungssystem S aus drei Gleichungen (Zeilen z_1, z_2, z_3) in den drei Variablen x_1, x_2, x_3 :

$$S : \begin{cases} z_1 : & x_1 + 2x_2 + x_3 = 1 \\ z_2 : & 3x_1 + 5x_2 + x_3 = 0 \\ z_3 : & 7x_1 + 12x_2 + 2x_3 = 2 \end{cases}$$

Wir werden das System S in ein **äquivalentes** System S' mit Zeilen z'_1, z'_2, z'_3 umformen. Dabei bedeutet die Äquivalenz von Gleichungssystemen S und S' ganz allgemein, dass die Lösungsmengen übereinstimmen: $L_S = L_{S'}$. Bei linearen Systemen erhält man, wie wir gleich begründen werden, stets ein äquivalentes System, wenn man zu einer der Zeilen ein Vielfaches einer anderen addiert. Im vorliegenden Fall empfiehlt es sich zum Beispiel, zu z_2 das -3 -fache von z_1 zu addieren. So entsteht aus der ursprünglichen Zeile z_2 die neue Zeile $z'_2 := z_2 - 3z_1 = -3$ mit den Koeffizienten $(3 - 3 \cdot 1) = 0$ bei x_1 , $5 - 3 \cdot 2 = -1$ bei x_2 , $1 - 3 \cdot 1 = -2$ bei x_3 und der rechten Seite $0 - 3 \cdot 1 = -3$, also:

$$z'_2 : -x_2 - 2x_3 = -3$$

Der Vorteil von z'_2 gegenüber z_2 besteht darin, dass die Variable x_1 nicht mehr auftritt, womit sich das System vereinfacht hat. Allerdings müssen wir uns noch davon überzeugen, dass dabei die Lösungsmenge nicht verändert wurde. Tatsächlich: Jedes Tripel (x_1, x_2, x_3) , das z_1 und z_2 erfüllt, muss notgedrungen auch die Linearkombination $z'_2 := z_2 - 3z_1$ erfüllen. Und umgekehrt muss jedes Tripel, das $z'_1 := z_1$ und z'_2 erfüllt, auch $z_2 = z'_2 + 3z_1$ erfüllen. Weil diese Erkenntnis natürlich nicht von unserem speziellen Beispiel abhängt, können wir als allgemeines Prinzip dahinter festhalten: Jedes lineare Gleichungssystem geht in ein äquivalentes über, wenn man zu einer Zeile ein Vielfaches einer anderen addiert. Was wir mir der zweiten Zeile gemacht haben, ist natürlich genauso mit der dritten möglich. Um darin den Koeffizienten 7 bei x_1 zu **eliminieren**, setzen wir also $z'_3 := z_3 - z_1$. Somit ist das neue, zu S äquivalente System

$$S' : \begin{cases} z'_1 : & x_1 + 2x_2 + x_3 = 1 \\ z'_2 : & -x_2 - 2x_3 = -3 \\ z'_3 : & -2x_2 + -5x_3 = -5 \end{cases}$$

entstanden. Mit Ausnahme von z'_1 kommt darin die Variable x_1 nicht mehr vor. Im nächsten Schritt bringen wir auch die Variable x_2 in der letzten Zeile zum Verschwinden, indem wir $z''_3 := z'_3 - 2z'_2$ setzen. Die beiden anderen Zeilen übernehmen wir: $z''_1 := z'_1$ und $z''_2 := z'_2$. Damit sind wir beim zu S' und somit auch zu S äquivalenten System

$$S'' : \begin{cases} z''_1 : & x_1 + 2x_2 + x_3 = 1 \\ z''_2 : & -x_2 - 2x_3 = -3 \\ z''_3 : & -x_3 = 1 \end{cases}$$

angelangt. Aus z_3'' lesen wir $x_3 = -1$ ab. Das setzen wir in z_2'' ein und erhalten $-x_2 + 2 = -3$, also $x_2 = 5$. Somit wird z_1'' zu $x_1 + 10 - 1 = 1$, also $x_1 = -8$. Damit haben wir die einelementige Lösungsmenge

$$L_S = \left\{ \begin{pmatrix} -8 \\ 5 \\ -1 \end{pmatrix} \right\}$$

ermittelt.

Die beschriebene Vorgangsweise war eine erste Anwendung des **Gauß'schen Eliminationsverfahrens**. In diesem Fall zielt es offenbar darauf ab, eine Systemmatrix (hier $A'' = (a''_{ij})$) in sogenannter **Halbdiagonalform** zu generieren, d.h. wo alle Eintragungen unterhalb der Hauptdiagonale (die mit den Diagonalelementen a''_{ii} besetzt ist) wegfallen, als Formel: $a''_{ij} = 0$ wann immer $i > j$. Sind die Diagonalelemente $\neq 0$, so lässt sich dann aus der letzten Gleichung x_n berechnen, mit x_n aus der vorletzten x_{n-1} etc. Wäre hingegen beispielsweise $a''_{nn} = 0$ und die zugehörige rechte Seite $b''_n \neq 0$, so gäbe es keine Lösung der resultierenden Gleichung $0x_n = b_n$, und $L_S = \emptyset$ wäre leer. Doch bevor wir uns mit einer systematischen Untersuchung möglicher Komplikationen befassen, nutzen wir die Methode aus, um die Inverse einer Matrix zu bestimmen.

1.5.3 Berechnung der inversen Matrix

Inhalt in Kurzfassung: Für eine reguläre quadratische Matrix A lässt sich durch geringfügige Erweiterung des Verfahrens aus 1.5.2 auch die Inverse A^{-1} ermitteln.

Im Beispiel aus 1.5.2 hätten wir statt $b_1 = 1, b_2 = 0, b_3 = 2$ für die rechte Seite $\mathbf{b} = (b_1, b_2, b_3)$ auch mit irgendwelchen anderen Zahlen beginnen können, ohne an den Zeilentransformationen, die von S zu S' und von S' zu S'' geführt haben, irgendetwas ändern zu müssen. Schlussendlich hätten mit derselben Systemmatrix nur andere Werte von der rechten Seite verarbeitet werden müssen. In jedem Fall ergibt sich eine eindeutige Lösung, also eine einelementige Lösungsmenge L_S . Auch aus der Theorie hätten wir das wie folgt erschließen können: Wir wissen aus Proposition 1.5.1.1, dass die Lösungsmenge eines linearen Gleichungssystems, sofern nicht leer, immer von der Gestalt $\mathbf{x}_0 + \ker(f)$ ist. Bei einer einelementigen Lösung für irgendein \mathbf{b} (wie wir sie im Beispiel ja vorgefunden haben) muss daher $\ker(f) = \{\mathbf{o}\}$ einelementig und somit f injektiv sein. Da dem 3×3 -System eine lineare Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ entspricht, bedeutet Injektivität automatisch Bijektivität. Also muss es zu jedem \mathbf{b} aus dem Zielbereich (im Beispiel ist das \mathbb{R}^3) genau ein \mathbf{x} aus dem Definitionsbereich (im Beispiel ebenfalls \mathbb{R}^3) geben mit $f(\mathbf{x}) = \mathbf{b}$, das also die eindeutige Lösung des Systems sein muss.

Wir wenden diese Erkenntnisse auf die kanonische Basisvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ als rechte Seiten \mathbf{b} des Systems an. Im System S'' aus 1.5.2 mussten wir abschließend noch drei Gleichungen lösen, um jede der Komponenten der Lösung (x_1, x_2, x_3) explizit zu machen. Auch diese Rechnung wollen wir möglichst nur einmal durchführen. Das gelingt, indem wir die Systemmatrix gleich nicht nur auf Halb-, sondern sogar auf Diagonalform bringen und dann noch die Diagonalelemente normieren. Um dabei Schreibearbeit zu sparen, notieren wir statt der Gleichungen nur ihre Koeffizienten. Neben der rechten Seite $\mathbf{b} =$

(b_1, b_2, b_3) des Systems S aus 1.5.2 mit $b_1 = 1, b_2 = 0$ und $b_3 = 2$ verarbeiten wir also auch gleich die drei kanonischen Einheitsvektoren $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$ und $\mathbf{e}_3 = (0, 0, 1)$.

$$\begin{array}{ccc}
 \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 1 & 0 & 0 \\ 3 & 5 & 1 & 0 & 0 & 1 & 0 \\ 7 & 12 & 2 & 2 & 0 & 0 & 1 \end{array} & \longrightarrow & \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -2 & -3 & -3 & 1 & 0 \\ 0 & -2 & -5 & -5 & -7 & 0 & 1 \end{array} & \longrightarrow & \\
 \\
 \longrightarrow & \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 & 3 & -1 & 0 \\ 0 & 0 & -1 & 1 & 1 & 2 & -1 \end{array} & \longrightarrow & \begin{array}{ccc|ccc} 1 & 2 & 0 & 2 & 0 & -2 & 1 \\ 0 & 1 & 0 & 5 & 1 & -5 & 2 \\ 0 & 0 & 1 & -1 & 1 & 2 & -1 \end{array} & \longrightarrow & \\
 \\
 & & \longrightarrow & \begin{array}{ccc|ccc} 1 & 0 & 0 & -8 & -2 & 8 & -3 \\ 0 & 1 & 0 & 5 & 1 & -5 & 2 \\ 0 & 0 & 1 & -1 & 1 & 2 & -1 \end{array} & &
 \end{array}$$

Es sollte selbsterklärend sein, wie diese Notation zu verstehen ist: Der erste Zahlenblock steht für das ursprüngliche System S in 1.5.2, ergänzt um die drei zusätzlichen rechten Seiten $\mathbf{e}_1, \mathbf{e}_2$ und \mathbf{e}_3 , der zweite entsprechend für S' und der dritte für S'' . Zum vierten Block gelangt man vermittels Elimination der dritten Koeffizienten in der ersten und zweiten Zeile mit Hilfe des dritten Koeffizienten der dritten Zeile (der ja $\neq 0$ ist; andernfalls wäre die Matrix A nicht regulär) sowie Multiplikation der dritten Zeile mit -1 zwecks Normierung. Der letzte Schritt schließlich führt zur Einheitsmatrix als Koeffizientenmatrix. Schreibt man die Variablen an, so liegt das folgende zu S äquivalente System vor:

$$\begin{array}{rcl}
 1x_1 + & 0x_2 + & 0x_3 = -8 \\
 0x_1 + & 1x_2 + & 0x_3 = -5 \\
 x_1 + & 0x_2 + & 1x_3 = -1
 \end{array}$$

Daraus liest man unmittelbar die eindeutig bestimmte Lösung

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -8 \\ 5 \\ -1 \end{pmatrix}$$

des ursprünglichen Systems $\mathbf{Ax} = \mathbf{b}$ ab. Entsprechend erhält man die (eindeutig bestimmten) Vektoren

$$\mathbf{x}_1 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} 8 \\ -5 \\ 2 \end{pmatrix}, \quad \text{und} \quad \mathbf{x}_3 = \begin{pmatrix} -3 \\ 2 \\ -1 \end{pmatrix}$$

mit $\mathbf{Ax}_i = \mathbf{e}_i$ für $i = 1, 2, 3$. Das bedeutet gleichzeitig, dass die Inverse A^{-1} von A die Einheitsvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ auf diese Vektoren $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ abbildet. Die 3×3 -Matrix A^{-1}

hat also als Spaltenvektoren $\mathbf{x}_1, \mathbf{x}_2$ und \mathbf{x}_3 . Der rechte 3×3 -Block in obiger Matrizenrechnung ist folglich die Inverse

$$A^{-1} = \begin{pmatrix} -2 & 8 & -3 \\ 1 & -5 & 2 \\ 1 & 2 & -1 \end{pmatrix}.$$

Klarerweise hängt diese Methode nicht von der Dimension 3 und den speziellen Einträgen der Matrix A in diesem Beispiel ab. Damit ist gezeigt:

Satz 1.5.3.1. Sei $A = (a_{ij})$ eine reguläre $n \times n$ -Matrix. Wendet man auf das $n \times 2n$ -Schema $A|I_n$, ausführlich:

$$\begin{array}{cccc|cccc} a_{1,1} & a_{1,2} & \dots & a_{1,n} & 1 & 0 & \dots & 0 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & 0 & 1 & \dots & 0 \\ \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} & 0 & 0 & \dots & 1 \end{array},$$

das oben exemplarisch beschriebene Eliminationsverfahren von Gauß an, so erhält man schließlich ein System der Bauart $I_n|A'$, ausführlich:

$$\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & a'_{1,1} & a'_{1,2} & \dots & a'_{1,n} \\ 0 & 1 & \dots & 0 & a'_{2,1} & a'_{2,2} & \dots & a'_{2,n} \\ \dots & \dots \\ 0 & 0 & \dots & 1 & a'_{n,1} & a'_{n,2} & \dots & a'_{n,n} \end{array},$$

mit einer Matrix $A' = (a'_{i,j})$, die mit der Inversen von A übereinstimmt: $A' = A^{-1}$.

Umgekehrt: Führt das Verfahren zu einem System dieser Gestalt, so ist A tatsächlich regulär mit der Inversen $A^{-1} = A'$.

Übungsaufgabe 90. (T) Wie lauten, falls sie existieren, die Inversen folgender Matrizen:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 2 & -1 & 7 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

Übungsaufgabe 91. (T) Inversion von Dreiecksmatrizen: Berechnen Sie die Inversen von

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & d & 1 & 0 \\ 0 & e & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & f & 1 \end{pmatrix}$$

Übungsaufgabe 92. (T) Gegeben sind die beiden Matrizen

$$A = \begin{pmatrix} 3 & -4 & 0 \\ 0 & 0 & 5 \\ 4 & 3 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}.$$

1. Berechnen Sie die inversen Matrizen A^{-1} und B^{-1} sowie die transponierten Matrizen A^T und B^T .
2. Berechnen Sie die Matrizen $(AB)^{-1}$, $(BA)^T$, $((BA)^T)^{-1}$, $((AB)^{-1})^T$, indem Sie geeignete Rechenregeln zur Vereinfachung verwenden.

Übungsaufgabe 93. (T^*) Berechnen Sie die Inversen von (a, b, c, d, e, f) sind beliebige reelle Parameter)

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & 1 & 0 & 0 \\ b & 0 & 1 & 0 \\ c & 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & d & 1 & 0 \\ 0 & e & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & f & 1 \end{pmatrix}.$$

Übungsaufgabe 94. (T) Berechnen Sie für die Matrizen aus Beispiel 93 (unter zu Hilfe nehmen der Rechenregeln für die Inversen) die Matrizen $(AB)^{-1}$, $(AB)^T$, $((BA)^T)^{-1}$, $((AB)^{-1})^T$.

1.5.4 Berechnung der Determinante

Inhalt in Kurzfassung: Die Determinante einer quadratischen Matrix in Halbdagonalform ist das Produkt ihrer Diagonalelemente. Damit kann die Determinante einer beliebigen quadratischen Matrix mit Hilfe des Eliminationsverfahrens von Gauß berechnet werden. Bei diesem Verfahren ändert sich die Determinante der Systemmatrix nämlich nur in sehr kontrollierter Weise: Zeilenvertauschungen bewirken einen Vorzeichenwechsel, und Multiplikation einer Zeile mit einem Faktor bewirkt die Multiplikation der Determinante mit demselben Faktor. Addiert man zu einer Zeile lediglich Vielfache anderer Zeilen, so verändert sich die Determinante nicht.

Auch die Determinante $\det A$ einer quadratischen $n \times n$ -Matrix $A = (a_{ij})$ lässt sich aus dem Eliminationsverfahren von Gauß ablesen. Solange die elementaren Schritte nur darin bestehen, dass zu einer Zeile z_i das c -fache cz_j einer anderen Zeile z_j addiert wird, zeigt die Schiefsymmetrie der Determinante wegen

$$\det(z_1, z_2, \dots, z_i + cz_j, \dots, z_j, \dots, z_n) = \det(z_1, z_2, \dots, z_i, \dots, z_j, \dots, z_n)$$

(hier für $i < j$, jedoch genauso für $i > j$) nämlich, dass sich die Determinante überhaupt nicht verändert. Gelingt es auf diese Art, A in eine Matrix $B = (b_{ij})$ in Halbdagonalform zu transformieren, so gilt daher

$$\det A = \det B = b_{11}b_{22} \dots b_{nn} = \prod_{i=1}^n b_{ii}.$$

Für die Begründung der letzten Gleichung hat man lediglich die Determinantenformel

$$\det B = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) b_{1\pi(1)} b_{2\pi(2)} \dots b_{n\pi(n)}$$

zu Rate zu ziehen: Tritt in einem der Summanden einmal der Fall $\pi(i) \neq i$ auf, so muss auch einmal $i > \pi(i)$ gelten, was bei Halbdagonalform $b_{i\pi(i)} = 0$ impliziert und somit das Verschwinden des entsprechenden Summanden bewirkt. Folglich bleibt lediglich der Summand für die identische Permutation mit $\pi(i) = i$ für $i = 1, \dots, n$ übrig. Diese ist gerade, hat folglich positives Signum $\operatorname{sgn}(\pi) = 1$. Also gilt tatsächlich

$$\det B = b_{11}b_{22} \dots b_{nn} = \prod_{i=1}^n b_{ii},$$

d.h.: Die Determinante einer Halbdiamontrix ist das Produkt ihrer Diagonalelemente.

Zu überlegen bleibt lediglich, ob wir jede quadratische Matrix A in Halbdiamontrixform bringen können. Die Analyse des Eliminationsverfahrens zeigt, dass bei der Elimination mittels Bildung von Linearkombinationen von Zeilen vorausgesetzt wurde, dass gewisse Elemente $\neq 0$ sind. Um die Situation zu verstehen, genügt es, den ersten Schritt zu betrachten. Wir haben die erste Zeile z_1 unverändert gelassen und ihr erstes Element a_{11} verwendet, um in den anderen Zeilen z_i durch Bildung von $z'_i := z_i + c_i z_1$ mit $c_i := -\frac{a_{i1}}{a_{11}}$ eine neue Zeile mit erstem Element $a'_{i1} = a_{i1} - c_i a_{11} = 0$ zu generieren. Das ist nur für $a_{11} \neq 0$ möglich. Ist hingegen $a_{11} = 0$, so sind zwei Fälle zu unterscheiden. Ist $a_{i1} = 0$ für alle $i = 2, \dots, n$, so ist nichts zu tun. Im weiteren Verlauf des Verfahrens, wenn in den übrigen Spalten die Elemente unter der Hauptdiagonale eliminiert werden, wird sich dieser Wert $a_{11} = 0$ des linken oberen Diagonalelementes nicht mehr ändern. Die Determinante der schlussendlich resultierenden Halbdiamontrix B wird als Produkt der Diagonalelemente folglich $= 0$ sein. Wegen $\det A = \det B$ folgt daraus auch $\det A = 0$. Dieselben Überlegungen gelten, wenn die entsprechende Situation im i_0 -ten statt im ersten Schritt auftritt, genauer: Seien die ersten $i_0 - 1$ Spalten und Zeilen bereits auf Halbdiamontrixform gebracht, so dass in der resultierenden Systemmatrix $A' = (a'_{ij})$ für alle $i < i_0$ und $j > i$ bereits $a'_{ij} = 0$ gilt. (Der einfacheren Notation halber schreiben wir A' , auch wenn nicht nur ein Schritt, sondern bereits i_0 Schritte durchgeführt worden sind.) Gilt dann auch $a'_{ii_0} = 0$ für alle $i \geq i_0$, so folgt $\det A = 0$ mit dem gleichen Argument wie für $i = 1$, und wir sind fertig. Lediglich, wenn $a_{i_0 i_0} = 0$ und gleichzeitig $a'_{i i_0} \neq 0$ für ein $i > i_0$ gilt, es also unterhalb des verschwindenden Diagonalelements ein von 0 verschiedenes gibt, müssen wir anders vorgehen. Und zwar können wir einfach die aktuelle i_0 -te Zeile z'_{i_0} mit der i -ten Zeile vertauschen. Wegen der Schiefsymmetrie verändert sich dabei die Determinante lediglich um das Vorzeichen. Wir fassen zusammen:

Satz 1.5.4.1. *Jede quadratische Matrix $A = (a_{i,j})$ lässt sich durch Zeilenoperationen der Gestalt $z'_i := z_i + cz_j$ mit $j \neq i$ (Addition von Vielfachen einer anderen Zeile) und eventuell Zeilenvertauschungen auf Halbdiamontrixform $B = (b_{i,j})$ bringen, d.h. mit $b_{i,j} = 0$ für alle $i > j$. Dann gilt*

$$\det A = (-1)^k \det B = (-1)^k \prod_{i=1}^n b_{i,i} = (-1)^k b_{1,1} b_{2,2} \dots b_{n,n},$$

wobei k die Anzahl der Vertauschungen von Zeilen ist.

Ergänzende Bemerkung: Vor allem bei händischen Rechnungen ist es (etwa um unangenehm große Zahlen oder Brüche zu vermeiden) oft praktisch, Zeilen mit einem Skalar zu multiplizieren. Das verändert natürlich den Wert der Determinante, allerdings auf leicht kontrollierbare Weise. Wegen der Multilinearität multipliziert sich dabei nämlich die Determinante um denselben Faktor. Ersetzt man in einer Matrix A beispielsweise die Zeile z_i nicht durch $z'_i := z_i + cz_j$, sondern durch $z'_i := rz_i + cz_j$, so gilt auch für die resultierende Systemmatrix A' die entsprechende Beziehung $\det A' = r \det A$. Solange $r \neq 0$ garantiert ist, kann man daraus natürlich wieder die Determinante $\det A = r^{-1} \det A'$ zurückrechnen. Will man von dieser Option Gebrauch machen, hat man also nicht nur die Anzahl der Zeilenvertauschungen zu zählen, man muss auch über die Multiplikatoren Buch halten und die Determinante der schlussendlich resultierenden Diagonalmatrix durch das Produkt all dieser Multiplikatoren dividieren, um $\det A$ zu erhalten.

Zum Abschluss wollen wir das Beispiel aus 1.5.2 und 1.5.3 rekapitulieren. Dort wurde die Matrix

$$A = (a_{i,j}) = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 5 & 1 \\ 7 & 12 & 2 \end{pmatrix}$$

durch Zeilenoperationen auf Halbdagonalform

$$B = (b_{i,j}) = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Unmittelbar lesen wir $\det B = b_{11}b_{22}b_{33} = 1 \cdot 1 \cdot 1 = 1$ ab. Um auch $\det A$ zu bestimmen, müssen wir Zeilenvertauschungen zählen und Multiplikatoren registrieren. Zeilenvertauschung war keine notwendig. Beim ersten Umformungsschritt traten auch keine Multiplikatoren $\neq 1$ auf. Beim zweiten Schritt jedoch wurde (neben $z'_1 := z_1$) $z''_2 := -z'_2$ (Multiplikator $r_1 = -1$) und $z''_3 := -z'_3 + 2z'_2$ (Multiplikator $r_2 = -1$ bei z'_3) gesetzt. Also ist $\det A = \frac{\det B}{r_1 r_2}$, was in diesem Fall wegen $r_1 = r_2 = -1$ aber wieder $\det A = 1$ bedeutet.

Für $n \times n$ -Matrizen mit $n = 3$ wie A im obigen Beispiel ist der Rechenaufwand beim Eliminationsverfahren vergleichbar zu dem bei der Regel von Sarrus. Bei größerem n gibt es aber keine Regel von Sarrus. Als eine dieser Regel entsprechende Verallgemeinerung lässt sich der sogenannte **Laplacesche Entwicklungssatz** interpretieren. Er lautet (hier als „Entwicklung nach der i -ten Zeile“):

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j})$$

und ist wie folgt zu verstehen: $A = (a_{i,j})$ ist eine $n \times n$ -Matrix, i irgendeiner der Zeilenindizes $1, \dots, n$ (die analoge Formel gilt auch mit Spaltenindizes), und $A_{i,j}$ ist die Matrix, die entsteht, wenn man aus A die i -te Zeile und die j -te Spalte streicht. Auf diese Weise lässt sich die Berechnung von $n \times n$ -Determinanten auf die Berechnung von

n verschiedenen $(n - 1) \times (n - 1)$ -Determinanten zurückführen. Rekursiv können damit grundsätzlich beliebig große Determinanten berechnet werden. Untersucht man den Rechenaufwand, erkennt man allerdings schnell, dass bei dieser Methode schlussendlich insgesamt $n!$ Produkte aus jeweils n Faktoren aufsummiert werden müssen, so wie in der Definition der Determinante. Das ist übrigens kein Zufall, da sich der Entwicklungssatz mit etwas Vorzeichenspielerei aber ohne grundsätzliche Schwierigkeiten mit Hilfe der Definition der Determinante beweisen lässt. Rechnerisch ist mit dem Entwicklungssatz also keine Vereinfachung verbunden. Analysiert man vergleichsweise das Eliminationsverfahren, so erkennt man im Gegensatz dazu, dass dabei etwa $\frac{n^3}{3}$ Multiplikationen von jeweils zwei Faktoren ausreichen. Ab $n = 4$ ist das ein wesentlich geringerer Aufwand. Zur Determinantenberechnung von 4×4 - oder noch größeren quadratischen Matrizen ist daher in der Regel das Eliminationsverfahren zu empfehlen.

1.5.5 Lösung linearer Gleichungssysteme, der allgemeine Fall

Inhalt in Kurzfassung: Durch eine weitere geringfügige Modifikation des Eliminationsverfahrens von Gauß lassen sich beliebige Systeme von n linearen Gleichungen in m Variablen lösen. Die gesuchten Koordinaten m -Tupel fasst man dabei als Vektoren $\mathbf{x} = (x_1, \dots, x_m)$ auf. Damit wird es möglich, die Lösungsmenge L des Gleichungssystems als Summe $L = \mathbf{x}_0 + L_0$ einer sogenannten partikulären Lösung \mathbf{x}_0 des gegebenen (im Allgemeinen inhomogenen) Systems und der allgemeinen Lösung L_0 des homogenen Systems darzustellen. Dabei ist L_0 ein Unterraum des \mathbb{R}^n , der durch eine Basis angegeben werden kann. Daraus ergibt sich eine Parameterdarstellung von L .

Um das Eliminationsverfahren von Gauß auf beliebige Gleichungssysteme anwenden zu können überlegen wir, welche Anpassungen nötig sind im Vergleich mit den in 1.5.2, 1.5.3 und 1.5.4 besprochenen Beispielen. Dort hatten wir es durchwegs mit quadratischen Systemen bzw. Matrizen zu tun, die wir überdies an manchen Stellen als regulär voraussetzen mussten. Diese einschränkenden Bedingungen wollen wir nun fallen lassen und Systeme von n linearen Gleichungen in m Variablen mit beliebigem $m, n \in \mathbb{N}$ ausgehen. Typisch ist der Fall $n \leq m$, wo also nicht mehr Gleichungen als Variablen vorliegen. Andernfalls ist die Lösung in der Regel leer bzw., im homogenen Fall, der Nullraum $\{\mathbf{0}\}$. Bei leerer Lösungsmenge kann man nach Vektoren suchen, die das System wenigstens in möglichst guter Annäherung lösen. Das geschieht in der sogenannten Ausgleichsrechnung, siehe auch ???. Hier wollen wir uns aber mit linearen Abhängigkeiten der Zeilen beschäftigen. Dabei kann sich, wie wir gleich sehen werden, die Zeilenzahl auch reduzieren, weshalb wir a priori auch $n > m$ zulassen.

Gegeben seien also reelle Zahlen (analog mit komplexen Zahlen oder sogar Funktionen) $a_{i,k}, b_i$ für $i = 1, \dots, m$ und $k = 1, \dots, n$. Wir stellen uns die Aufgabe, alle m -Tupel $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$ zu bestimmen, die dem **linearen Gleichungssystem**

$$\begin{array}{cccccc} a_{1,1}x_1 + & a_{1,2}x_2 + & \cdots + & a_{1,m}x_m & = & b_1 \\ a_{2,1}x_1 + & a_{2,2}x_2 + & \cdots + & a_{2,m}x_m & = & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n,1}x_1 + & a_{n,2}x_2 + & \cdots + & a_{nm}x_m & = & b_n \end{array}$$

genügen. Diese Aufgabe lässt sich auch äquivalent in Matrixschreibweise darstellen: Es soll

$$\mathbf{Ax} = \mathbf{b} \quad \text{mit} \quad A = (a_{ik}), \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

gelten. A nennt man die **Systemmatrix**, die a_{ik} die **Koeffizienten**, die x_k die **Unbekannten** und \mathbf{b} den **inhomogenen Bestandteil**, der sich aus den **Störgliedern** b_i zusammensetzt. Sind alle Störglieder $b_i = 0$, so spricht man von einem **homogenen System** – im Gegensatz zu einem **inhomogenen**, wo auch $b_i \neq 0$ zugelassen ist.

Offenbar lässt sich die Aufgabe auch so interpretieren: Sei $f = f_A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ jene lineare Abbildung, die bezüglich der kanonischen Basen durch die Systemmatrix A dargestellt wird. Dann hat die **Lösungsmenge**

$$L := \{\mathbf{x} \in \mathbb{R}^m : f(\mathbf{x}) = \mathbf{b}\} = f^{(-1)}(\mathbf{b})$$

des Gleichungssystems als Elemente genau jene $\mathbf{x} \in \mathbb{R}^m$, die durch f auf $\mathbf{b} \in \mathbb{R}^n$ abgebildet werden. Diese Sichtweise ermöglicht es, wann immer es nützlich erscheint, auf die bisher entwickelte Lineare Algebra zurückzugreifen. Zum Beispiel kennen wir bereits die Gestalt von L . Entweder ist $L = \emptyset$ leer, oder von der Form

$$L = \mathbf{x}_0 + \ker f_A,$$

sofern \mathbf{x}_0 irgendeine Lösung des Systems (d.h. irgendein Element von L) ist und $\ker f_A$ wie üblich den Kern der linearen Abbildung f_A bezeichnet, der stets ein Unterraum von \mathbb{R}^m ist und für den wir der Kürze halber auch $\ker A$ schreiben. L kann daher als bestimmt gelten, sobald wir irgendein $\mathbf{x}_0 \in L$ und eine Basis von $\ker A$ gefunden haben. Wir wollen uns nun davon überzeugen, dass das Gauß'sche Eliminationsverfahren, wie wir es bereits kennen gelernt haben, mit nur geringfügigen Anpassungen genau so eine Lösung liefern kann.

Zu diesem Zweck rufen wir uns das Verfahren aus 1.5.2 in Erinnerung sowie die Diskussion in 1.5.4 über allfällige Komplikationen im Fall, dass gewisse Elemente in der Hauptdiagonale $= 0$ sein sollten. Ist $a_{11} \neq 0$, so kann man durch Ersetzung der Zeilen z_i durch die Zeilen $z'_i := z_i - \frac{a_{i1}}{a_{11}}$ für $i = 2, \dots, m$ und Beibehaltung von z_1 als z'_1 ein äquivalentes System mit einer Systemmatrix $A' = (a'_{ij})$ erzeugen, wo in der ersten Spalte alle Eintragungen unter dem Hauptdiagonalelement $a_{11} = a'_{11}$ zu $a'_{i1} = 0$ werden. Ignoriert man nach diesem ersten Eliminationsschritt (vorerst) die erste Zeile und die erste Spalte dieses Systems, so verbleibt ein Gleichungssystem, in dem sich sowohl die Zeilenzahl n als auch die Spaltenzahl m um 1 reduziert haben. Mit diesem System kann in derselben Weise fortgefahren werden. Nach endlich vielen Eliminationsschritten wird $m = 0$ oder $n = 0$. Was dann zu tun ist, werden wir gleich analysieren. Zuvor wollen wir aber alle möglichen Fälle bei einem einzelnen Eliminationsschritt besprechen, auch für $a_{11} = 0$.

Sei also $A = (a_{ij})$ die aktuell zu bearbeitende Systemmatrix mit n Zeilen und m Spalten $m, n > 0$ und $\mathbf{b} = (b_1, \dots, b_n)$ der inhomogene Bestandteil, d.h. der Vektor der Störglieder. Insgesamt sind die folgenden Fälle zu unterscheiden:

1 Lineare Algebra

- Fall 1: Ist $a_{11} \neq 0$, so wird die erste Zeile z_1 (die mit diesem Koeffizienten a_{11} beginnt) festgehalten. Ist $n > 1$, so erfolgt Elimination wie soeben beschrieben. Im nächsten Schritt wird mit dem Rest, der nach Streichung von erster Zeile und Spalte verbleibt, weiter gearbeitet. Dabei reduzieren sich m und n jeweils um 1. Ist hingegen $n = 1$, so ist nichts zu tun, und wir sind fertig. Unten werden wir die Situation am Ende des Verfahrens analysieren.
- Fall 2: Ist $a_{11} = 0$ und $a_{i_0 1} \neq 0$ für ein $i_0 > 1$, so arbeitet man mit dem äquivalenten System weiter, das man durch Vertauschung von erster und i_0 -ter Zeile erhält: $z'_1 := z_{i_0}$, $z'_{i_0} = z_1$. Somit landet man wieder bei Fall 1.
- Fall 3: Ist $a_{i_1} = 0$ für alle i und $a_{i_0 j_0} \neq 0$ für gewisse $i_0, j_0 \geq 2$, so vertauscht man die erste mit j_0 -ten Spalte. Kann $i_0 = 1$ genommen werden, landet man direkt in Fall 1, ansonsten im Fall 2, der in der dort beschriebenen Weise durch Tausch der ersten mit der i_0 -ten Zeile auf Fall 1 zurückgeführt wird. Man beachte, dass dies aber auch eine Vertauschung der Variablen x_1 und x_{j_0} bedeutet. Diese Vertauschung muss registriert und ganz am Ende des Verfahrens berücksichtigt werden. Jedenfalls landet man schließlich wieder bei Fall 1.
- Fall 4: Ist $a_{ij} = 0$ für alle i und j , jedoch $b_{i_0} \neq 0$ für ein i_0 , dann bedeutet die i_0 -te Zeile einen Widerspruch. Denn ihr entspricht ja die Gleichung $0x_1 + 0x_2 + \dots + 0x_m = b_{i_0}$, die nicht erfüllt werden kann. Folglich ist in diesem Fall die Lösungsmenge des Gleichungssystems leer und das Verfahren erfolgreich abgeschlossen, weil wir $L = \emptyset$ bestimmt haben.
- Fall 5: Ist $a_{ij} = 0$ und überdies $b_i = 0$ für alle i, j , dann sind wir fertig. Denn alle Gleichungen haben in diesem Fall die triviale Gestalt $0x_1 + 0x_2 + \dots + 0x_m = 0$, wofür alle $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$ Lösungen sind. Diese Gleichungen liefern also keine einschränkende Information über die Lösungen und können gestrichen werden. Das Verfahren ist damit abgeschlossen.

Das Verfahren kann also auf folgende Arten enden. Tritt irgendwann Fall 4 ein, so ist die Lösungsmenge leer, und wir sind fertig. Tritt irgendwann Fall 5 ein, so bestehen die letzten Zeilen nur aus Nullen, die entsprechenden Gleichungen haben keinen Informationsgehalt und sind zu streichen. Was übrig bleibt, ist dieselbe Situation wie bei einem Ende in Fall 1, der eine Matrix in Halbdagonalform liefert. Außer im Fall 4, der wegen der leeren Lösungsmenge nicht weiter interessiert, landen wir also stets bei einem System in Halbdagonalform:

$$\begin{array}{ccccccc} a_{1,1} & a_{1,2} & \dots & a_{1,n} & \dots & a_{1,m} & b_1 \\ 0 & a_{2,2} & \dots & a_{2,n} & \dots & a_{2,m} & b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{n,n} & \dots & a_{n,m} & b_n \end{array}$$

mit Diagonalelementen $a_{i,i} \neq 0$. Insbesondere ist $n \leq m$. Wurden im Zuge des Eliminationsverfahrens keine Spaltenvertauschungen vorgenommen (in diesem Fall müssten

lediglich die Variablen entsprechend unnummeriert werden), so entspricht dieser Matrix das folgende Gleichungssystem:

$$\begin{array}{cccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & + & \dots & + & a_{1,m}x_m & = & b_1 \\ & & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & + & \dots & + & a_{2,m}x_m & = & b_2 \\ & & & & \dots & & & & \dots & & & & \dots \\ & & & & & & a_{n,n}x_n & + & \dots & + & a_{n,m}x_m & = & b_n \end{array}$$

Gibt man für $m - n$ Variablen $x_{n+1}, x_{n+2}, \dots, x_m$ beliebige Werte vor, so errechnet sich daraus der Wert für x_n aus der n -ten Gleichung eindeutig als

$$x_n = a_{n,n}^{-1}(b_n - a_{n,n+1}x_{n+1} - \dots - a_{n,m}x_m).$$

Mit diesem Wert ergibt sich analog aus der $m - 1$ -ten Gleichung eindeutig

$$x_{n-1} = a_{n-1,n-1}^{-1}(b_{n-1} - a_{n-1,n}x_n - a_{n-1,n+1}x_{n+1} - \dots - a_{n-1,m}x_m).$$

Insgesamt ergibt sich also zu jeder Vorgabe von (x_{n+1}, \dots, x_m) (als frei wählbarer Parameter) eine eindeutige vollständige Lösung $\mathbf{x} = (x_1, x_2, \dots, x_m)$. Man spricht von einer $m - n$ -parametrischen Lösungsmenge.

Will man die Abhängigkeit der x_1, \dots, x_n von den $x_{n+1}, x_{n+2}, \dots, x_m$ noch deutlicher machen, so empfiehlt es sich, den linken Teil der Koeffizientenmatrix von Halbdiegonalform sogar zur Einheitsmatrix zu transformieren, was durch Ausweitung des Eliminationsverfahrens auf die Elemente oberhalb der Hauptdiagonale ja möglich ist. (Man eliminiert mit $a_{2,2} \neq 0$ nicht nur die Elemente $a_{i,2}$ mit $i = 3, \dots, n$ zu 0, sondern auch das Element $a_{1,2}$, dann mit $a_{3,3} \neq 0$ auch die Elemente $a_{1,3}$ und $a_{2,3}$ etc. Schlussendlich normiert man noch jede Zeile so, dass $a_{i,i} = 1$ für $i = 1, \dots, m$ gilt. In Koeffizientenschreibweise landet man bei einem System dieser Gestalt:

$$\begin{array}{cccccccc} 1 & 0 & \dots & 0 & a_{1,n+1} & \dots & a_{1,m} & b_1 \\ 0 & 1 & \dots & 0 & a_{2,n+1} & \dots & a_{2,m} & b_2 \\ \dots & \dots \\ 0 & 0 & \dots & 1 & a_{n,n+1} & \dots & a_{n,m} & b_n \end{array}$$

(natürlich mit anderen Werten der $a_{i,j}$ und b_i als zuvor). Übersetzt man diese Matrix wie oben in das entsprechende Gleichungssystem und bringt man die Variablen x_{n+1} bis x_m auf die rechte Seite, so erhält man:

$$\begin{array}{l} x_1 = b_1 - a_{1,n+1}x_{n+1} - a_{1,n+2}x_{n+2} - \dots - a_{1,m}x_m \\ x_2 = b_2 - a_{2,n+1}x_{n+1} - a_{2,n+2}x_{n+2} - \dots - a_{2,m}x_m \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ x_n = b_n - a_{n,n+1}x_{n+1} - a_{n,n+2}x_{n+2} - \dots - a_{n,m}x_m \end{array}$$

Weil die x_{n+1}, \dots, x_m frei gewählt werden können, dürfen wir sie als Parameter betrachten, die wir mit t_1, \dots, t_{m-n} bezeichnen. Somit lässt sich das System durch weitere Zeilen

1 Lineare Algebra

ergänzen und wird zu:

$$\begin{array}{rcccccccc}
 x_1 & = & b_1 & - & a_{1,n+1}t_1 & - & a_{1,n+2}t_2 & - & \dots & - & a_{1,m}t_{m-n} \\
 x_2 & = & b_2 & - & a_{2,n+1}t_1 & - & a_{2,n+2}t_2 & - & \dots & - & a_{2,m}t_{m-n} \\
 \dots & & \dots \\
 x_n & = & b_n & - & a_{n,n+1}t_1 & - & a_{n,n+2}t_2 & - & \dots & - & a_{n,m}t_{m-n} \\
 x_{n+1} & = & & & t_1 & & & & & & \\
 x_{n+2} & = & & & & & t_2 & & & & \\
 \dots & & \dots \\
 x_m & = & & & & & & & & & t_{m-n}
 \end{array}$$

Wir fassen darin die Spalten zu Vektoren zusammen:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \\ x_{n+1} \\ \dots \\ x_i \\ \dots \\ x_m \end{pmatrix}, \quad \mathbf{x}_0 = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \\ 0 \\ \dots \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad \mathbf{a}_i = \begin{pmatrix} -a_{1,i} \\ -a_{2,i} \\ \dots \\ -a_{n,i} \\ 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix},$$

wobei die Eintragung 1 von \mathbf{a}_i in der $n+i$ -ten Komponente steht, $i = 1, \dots, m-n$.
Damit gilt

$$\mathbf{x} = \mathbf{x}_0 - t_1\mathbf{a}_1 - t_2\mathbf{a}_2 - \dots - t_{m-n}\mathbf{a}_{m-n},$$

wobei die Parameter t_1, \dots, t_{m-n} unabhängig voneinander ganz \mathbb{R} durchlaufen. Indem wir t_i durch $-t_i$ ersetzen, können wir natürlich genauso

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \dots + t_{m-n}\mathbf{a}_{m-n},$$

schreiben.

Die \mathbf{a}_i sind linear unabhängig, bilden daher eine Basis eines Unterraums von \mathbb{R}^n . Dieser Unterraum ist die Lösungsmenge des zugehörigen homogenen Systems, weshalb wir ihn mit L_0 bezeichnen. Damit haben wir die gesuchte Darstellung

$$L = \mathbf{x}_0 + L_0 = \mathbf{x}_0 + \text{LH}(\mathbf{a}_1, \dots, \mathbf{a}_{m-n})$$

der Lösungsmenge L des linearen Gleichungssystems gefunden, nämlich als Summe einer sogenannten partikulären Lösung \mathbf{x}_0 des inhomogenen Systems und des Lösungsraumes L_0 des homogenen Systems, wobei L_0 durch die Basis $\{\mathbf{a}_1, \dots, \mathbf{a}_{m-n}\}$ gegeben ist.

Aus dieser Darstellung können wir auch den Rang der Matrix A ablesen. Sei A' die $n \times m$ -Matrix in Hauptdiagonalgestalt, auf die A durch das Eliminationsverfahren gebracht worden ist. A' stellt ein zu A äquivalentes Gleichungssystem dar, das heißt die

Lösungsmengen stimmen überein. Das gleiche gilt für die Lösungsmengen der zugehörigen homogenen Systeme, die also beide $= L_0$ sind. L_0 ist der Kern von A und hat die Basis $\mathbf{a}_1, \dots, \mathbf{a}_{m-n}$, also gilt nach der Rangformel 1.2.3.2

$$m = \operatorname{rg} A + \operatorname{def} A = \operatorname{rg} A + \dim \ker A = \operatorname{rg} A + \dim L_0 = \operatorname{rg} A + m - n,$$

woraus $\operatorname{rg} A = n$ folgt. Zusammenfassend haben wir also gefolgert:

Folgerung 1.5.5.1. *Ergibt das Eliminationsverfahren von Gauß angewendet auf eine Matrix A eine Halbdiaagonalmatrix A' mit n Zeilen, so gilt $\operatorname{rg} A = n$.*

Beispiele dazu finden sich in manchen der nachfolgenden Übungsaufgaben.

Übungsaufgabe 95. (T) Für welche Werte $\alpha, \beta \in \mathbb{R}$ besitzt das Gleichungssystem

$$\begin{aligned} 2x_1 + x_2 &= 0 \\ x_1 + 2x_2 + x_3 &= 0 \\ x_2 + 2x_3 + x_4 &= 0 \\ x_3 + \alpha x_4 &= \beta \end{aligned}$$

(a) keine Lösung (b) eine eindeutige Lösung (c) unendlich viele Lösungen? Geben Sie jeweils auch den Rang der Koeffizientenmatrix bzw. der erweiterten Koeffizientenmatrix an!

Übungsaufgabe 96. (T) Bestimmen Sie die Lösungen von Aufgabe 95 in den Fällen (b) und (c).

Übungsaufgabe 97. (T*) Ein Magisches Quadrat ist ein $n \times n$ Zahlenschema, in dem die natürlichen Zahlen 1 bis n^2 so angeordnet sind, dass die Summe über alle Zeilen, Spalten und Diagonalen gleich ist.

(a) Zeigen Sie durch Aufstellen eines geeigneten linearen Gleichungssystems, dass es kein magisches Quadrat der Größe $n = 2$ gibt.

(b) Wie groß muss die Summe über alle Zeilen, Spalten und Diagonalen in einem magischen Quadrat mit $n = 3$ sein?

(c) Bestimmen Sie alle Möglichkeiten,

$$\begin{array}{ccc} x_1 & x_2 & x_3 \\ x_4 & 5 & x_6 \\ x_7 & 1 & x_8 \end{array}$$

zu einem magischen Quadrat zu ergänzen!

Übungsaufgabe 98. (T) Geben Sie für den Vektorraum, der von den Spalten der Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & -1 & 2 & 0 \end{pmatrix}$$

erzeugt wird, eine Basis an. Bestimmen Sie sodann die Lösungsmenge des linearen Gleichungssystems

$$A\mathbf{x} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}, \quad a, b, c, d \in \mathbb{R}.$$

Übungsaufgabe 99. (T) Bestimmen Sie für die folgenden Gleichungssysteme jeweils den Rang der Koeffizientenmatrix und den Rang der erweiterten Matrix. Untersuchen Sie die Gleichungssysteme auf Lösbarkeit und geben Sie die Dimension der Lösungsräume an!

$$\begin{array}{l} 3x - 6y - 12z = 9 \quad 3x - 6y - 12z = 9 \quad 3x - 6y - 12z = 9 \\ -4x + 8y + 16z = -12 \quad -4x + y + 16z = -19 \quad -4x + 8y + 16z = 12 \end{array}.$$

Übungsaufgabe 100. (T) Lösen Sie die folgenden linearen Gleichungssysteme:

$$\begin{array}{l} 3x + 2y + z = 11 \quad x + 3y + z = 0 \quad x + 3y + z = 0 \\ 2x - 4y - 3z = -11, \quad x + 2y - z = 1, \quad x + 2y - z = 1. \\ x + 3y - 2z = 9 \quad 3x + 7y - z = 2, \quad 3x + 7y - z = 1 \end{array}$$

1.5.6 Ergänzende Bemerkungen

Inhalt in Kurzfassung: Es dient dem Verständnis der Lösung linearer Gleichungssysteme, wenn man sich klar macht, dass die Operationen des Eliminationsverfahrens von Gauß auch als Multiplikation mit regulären quadratischen Matrizen bzw., äquivalent dazu, als Anwendung von Vektorraumisomorphismen interpretiert werden können. Diese Möglichkeit soll im Folgenden beschrieben werden. Außerdem zeigt sich, dass der Rang einer Matrix mit mehreren Zahlen übereinstimmt: dem Zeilenrang, dem Spaltenrang und dem Determinantenrang. All diese Zahlen bleiben beim Eliminationsverfahren unverändert.

Oft ist es praktisch, die Äquivalenzumformungen eines linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ als Multiplikation mit einer regulären quadratischen Matrix B der entsprechenden Dimension mit Inverser B^{-1} aufzufassen. Denn in diesem Fall folgt nicht nur (vermittels Matrixmultiplikation mit B von links) aus $A\mathbf{x} = \mathbf{b}$ die Beziehung $BA\mathbf{x} = B\mathbf{b}$, sondern auch umgekehrt aus dieser (vermittels Multiplikation mit B^{-1} von links) wieder die ursprüngliche Beziehung

$$A\mathbf{x} = I_n A\mathbf{x} = (B^{-1}B)A\mathbf{x} = B^{-1}(BA\mathbf{x}) = B^{-1}B\mathbf{b} = I_n \mathbf{b}.$$

Übungsaufgabe 101. (E) Überlegen Sie sich, welche Matrix B man wählen muss, damit BA oder AB (welche der beiden?) jene Matrix ist, die entsteht, wenn man mit A folgende Operationen ausführt:

1. Vertauschung der i -ten mit der j -ten Zeile in A .
2. Vertauschung der i -ten mit der j -ten Spalte in A .

3. Multiplikation der i -ten Zeile von A mit einem Faktor c .
4. Multiplikation der i -ten Spalte von A mit einem Faktor c .
5. Ersetzung der i -ten Zeile von A durch die Summe aus i -ter und j -ter Zeile.
6. Ersetzung der i -ten Spalte von A durch die Summe aus i -ter und j -ter Spalte.

Wir haben den Rang $\operatorname{rg} f$ einer linearen Abbildung $f : V_1 \rightarrow V_2$ definiert als die Dimension des Bildes $f(V_1)$, den Rang $\operatorname{rg} A$ einer Matrix A als den Rang der durch A dargestellten linearen Abbildung. Dieser Rang $\operatorname{rg} A$ hat sich als unabhängig erwiesen von den Basen von V_1 und V_2 , die der Matrixdarstellung von f zugrunde liegen. Man kann sich daher oBdA auf die Standardvektorräume und die kanonischen Basen beziehen. Folglich stimmt der Rang $\operatorname{rg} A$ einer $n \times m$ -Matrix überein mit der Dimension des Unterraums, der von den Spaltenvektoren von A in \mathbb{R}^n erzeugt wird. Man spricht daher auch vom Spaltenrang von A , analog definiert man den Zeilenrang. Oft ist es nützlich zu wissen, dass Spalten- und Zeilenrang übereinstimmen. Um das zu beweisen bietet sich als Hilfsmittel auch noch der Determinantenrang an, die maximale Dimension einer regulären quadratischen Teilmatrix von A , die durch eventuelle Streichung von Zeilen und Spalten von A entstehen kann. Die laut Proposition 1.5.6.1 übereinstimmenden Zahlen sind also schlicht der **Rang** der Matrix A . Genauer gilt:

Proposition 1.5.6.1. *Für jede Matrix $A = (a_{i,j})$ (mit n Zeilen und m Spalten) stimmen folgende Zahlen überein:*

- der **Spaltenrang** $\operatorname{rg}_S(A)$, das ist definitionsgemäß die maximale Kardinalität einer linear unabhängigen Teilmenge der Spaltenvektoren von A (aufgefasst als Elemente des Standardvektorraumes \mathbb{R}^n).
- der **Zeilenrang** $\operatorname{rg}_Z(A)$, das ist definitionsgemäß die maximale Kardinalität einer linear unabhängigen Teilmenge der Zeilenvektoren von A (aufgefasst als Elemente des Standardvektorraumes \mathbb{R}^n).
- der **Determinantenrang** $\operatorname{rg}_D(A)$, das ist definitionsgemäß die maximale Zahl k , für die folgendes gilt: Es gibt k -elementige Mengen $I = \{1 \leq i_1 < i_2 < \dots < i_k \leq n\}$ von Zeilenindizes und $J = \{1 \leq j_1 < j_2 < \dots < j_k \leq m\}$ von Spaltenindizes derart, dass $\det A' \neq 0$ gilt für jene $k \times k$ -Matrix $A' = A'(I, J)$, die entsteht, wenn man in A nur jene Elemente a_{ij} mit $i \in I$ und $j \in J$ belässt und alle anderen Eintragungen streicht.

Beweis. Wir zeigen zunächst mittels Induktion nach k folgende **Hilfsbehauptung**: Seien die Vektoren $\mathbf{a}_1 = (a_{1,1}, \dots, a_{m,1}), \dots, \mathbf{a}_k = (a_{1,k}, \dots, a_{m,k}) \in \mathbb{R}^m$ linear unabhängig (also $k \leq m$). Dann gibt es eine k -elementige Menge $I = \{i_1, \dots, i_k\}$ von Indizes mit $1 \leq i_1 < \dots < i_k \leq m$ derart, dass die Vektoren $\mathbf{a}'_1 = (a_{i_1,1}, \dots, a_{i_k,1}), \dots, \mathbf{a}'_k = (a_{i_1,k}, \dots, a_{i_k,k}) \in \mathbb{R}^k$, die entstehen, wenn man in den \mathbf{a}_j nur jene i -ten Komponenten mit $i \in I$ belässt, linear unabhängig sind.

Beweis der Hilfsbehauptung: Für $k = 0$ ist nichts zu zeigen, weil die leere Menge linear unabhängig ist. Angenommen wir haben die Hilfsbehauptung für ein bestimmtes k bewiesen und haben $k + 1$ linear unabhängige Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_{k+1} \in \mathbb{R}^m$ vorgegeben. Laut Induktionsannahme gibt es eine k -elementige Indexmenge $I = \{i_1, \dots, i_k\}$ wie in der Hilfsbehauptung derart, dass die Vektoren $\mathbf{a}'_1, \dots, \mathbf{a}'_k \in \mathbb{R}^k$ linear unabhängig sind, also eine Basis von \mathbb{R}^k bilden. Der entsprechende Vektor $\mathbf{a}'_{k+1} = (a_{i_1 k+1}, \dots, a_{i_k k+1}) \in \mathbb{R}^k$ lässt sich in eindeutiger Weise als Linearkombination $\mathbf{a}'_{k+1} = \sum_{l=1}^k r_l \mathbf{a}'_l$ schreiben. Wäre auch $\mathbf{a}_{k+1} = \sum_{l=1}^k r_l \mathbf{a}_l \in \mathbb{R}^m$, so widerspräche das der linearen Unabhängigkeit der $\mathbf{a}_1, \dots, \mathbf{a}_{k+1}$. Also muss es eine Komponente $i_{k+1} \notin I$ mit $a_{i_{k+1} k+1} \neq \sum_{l=1}^k r_l a_{i_{k+1} l} \in \mathbb{R}^m$ geben. Die um diesen Index i_{k+1} erweiterte Indexmenge $I^+ := \{i_1, \dots, i_{k+1}\}$ erfüllt die Hilfsbehauptung für $k + 1$.

Mit der Hilfsbehauptung lässt sich nun relativ schnell die Gleichheit der drei Zahlen beweisen: Gehen wir von gewissen k linear unabhängigen Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_k$ der Matrix A aus. Laut Hilfsbehauptung gibt es dann eine Auswahl von k geeigneten Zeilen, so dass die entsprechenden Vektoren $\mathbf{a}'_1, \dots, \mathbf{a}'_k \in \mathbb{R}^k$ (Bezeichnung wie in der Hilfsbehauptung) linear unabhängig sind. Daraus folgt $\det(\mathbf{a}'_1, \dots, \mathbf{a}'_k) \neq 0$. Für $k = \text{rg}_S(A)$ zeigt diese Überlegung $\text{rg}_S(A) \leq \text{rg}_D(A)$.

Im nächsten Schritt wollen wir $\text{rg}_Z(A) \geq \text{rg}_D(A) =: k$ zeigen. Dazu wählen wir I und J und somit A' mit $\det A' \neq 0$ wie in der Definition von $\text{rg}_D(A)$ aus. Wegen 1.4.5.4 gilt

$$\det(A')^T = \det(A') \neq 0.$$

Folglich sind die Spalten von $(A')^T$, die ja gleichzeitig die Zeilen von A' sind, linear unabhängig, was $\text{rg}_Z(A) \geq k = \text{rg}_D(A)$ beweist.

Ganz symmetrisch zur bewiesenen Ungleichungskette $\text{rg}_S(A) \leq \text{rg}_D(A) \leq \text{rg}_Z(A)$ zeigt man auch $\text{rg}_Z(A) \leq \text{rg}_D(A) \leq \text{rg}_S(A)$, womit der Beweis der Proposition vollständig ist. \square

Übungsaufgabe 102. (P) Man begründe: Ein inhomogenes lineares Gleichungssystem aus n Gleichungen in m Variablen hat genau dann eine Lösung, wenn der Rang der $n \times m$ -Matrix des homogenen Systems mit dem Rang der um den inhomogenen Anteil erweiterten $n \times (m + 1)$ -Matrix übereinstimmt.

Systemmatrizen mit m Spalten und n Zeilen haben im Fall $m > n$ typischerweise mehr als einpunktige Lösungsmengen. Im Fall $n > m$ jedoch sind sie in der Regel überbestimmt und haben gar keine Lösung. Man kann jedoch stets nach einem Lösungsvektor $\mathbf{x} \in \mathbb{R}^m$ suchen, der das Gleichungssystem in möglichst guter Annäherung erfüllt. Dieser Zugang führt zur Ausgleichsrechnung und Regressionsgerade, siehe 2.4.6. Da es sich dabei um Extremwertaufgaben handelt, welche die Differentialrechnung in mehreren Variablen voraussetzen, werden wir diesen Themenkreis erst im nächsten Kapitel behandeln.

Ohne Beweis sei abschließend auch die sogenannte **Cramersche Regel** erwähnt:

Satz 1.5.6.2. Ist A eine reguläre $n \times n$ -Matrix und $A\mathbf{x} = \mathbf{b}$, $\mathbf{x} = (x_1, \dots, x_n)$, das zugehörige lineare Gleichungssystem mit inhomogenem Anteil $\mathbf{b} = (b_1, \dots, b_n)$, dann ist

die Lösung gegeben durch

$$x_k = \frac{1}{|A|} \begin{vmatrix} a_{11} & \cdots & a_{1k-1} & b_1 & a_{1k+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2k-1} & b_2 & a_{2k+1} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nk-1} & b_n & a_{nk+1} & \cdots & a_{nn} \end{vmatrix}$$

für $k = 1, \dots, n$.

1.6 Eigenwerttheorie

Dieser Abschnitt dreht sich rund um die Begriffe Eigenwert, Eigenvektor und Eigenraum. In 1.6.1 werden diese Begriffe definiert und die Rolle der charakteristischen Gleichung besprochen, welche auch der Ausgangspunkt für die einschlägigen Berechnungsmethoden ist, siehe 1.6.2. In engem Zusammenhang damit steht der Begriff der Ähnlichkeit von Matrizen. Dieser ist für uns vor allem im diagonalisierbaren Fall von Interesse. Denn dann ist bei geeignetem Basiswechsel die Transformation auf Diagonalgestalt möglich (1.6.3).

1.6.1 Definitionen und charakteristische Gleichung

Inhalt in Kurzfassung: Die Begriffe Eigenwert, Eigenvektor und Eigenraum (jeweils zu einem bestimmten Eigenwert) einer linearen Abbildung $f : V \rightarrow V$ oder einer quadratischen Matrix werden definiert.

Es sei A eine reelle oder komplexe (n, n) -Matrix. Dann sind alle reellen oder komplexen Zahlen λ zu ermitteln, zu denen es Vektoren $\mathbf{x} \neq \mathbf{o}$ in \mathbb{R}^n oder \mathbb{C}^n gibt mit

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Man definiert:

Definition 1.6.1.1. Sei V ein Vektorraum, $f : V \rightarrow V$ eine lineare Abbildung, $\mathbf{x} \in V$ mit $\mathbf{x} \neq \mathbf{o}$ und $\lambda \in \mathbb{R}$ (sehr oft auch: $\lambda \in \mathbb{C}$) mit $f(\mathbf{x}) = \lambda\mathbf{x}$. Dann heißt λ ein **Eigenwert** von f , und \mathbf{x} heißt **Eigenvektor** von f zum Eigenwert λ . Für einen beliebigen Skalar λ heißt die Menge $E_{f,\lambda}$ aller $\mathbf{x} \in V$ mit $f(\mathbf{x}) = \lambda\mathbf{x}$ auch der **Eigenraum** von f zum Eigenwert λ .

Ist A eine quadratische $n \times n$ -Matrix und $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (oder auch $f_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$) die bezüglich der kanonischen Basis durch A dargestellte lineare Abbildung, so heißen die Eigenwerte, Eigenvektoren und Eigenräume von f_A auch diejenigen von A .

Erste wichtige Beobachtungen dazu sind die folgenden.

Proposition 1.6.1.2. Sei $f : V \rightarrow V$ linear.

1. Der Eigenraum von f zu jedem Eigenwert λ ist ein Unterraum von V .

1 Lineare Algebra

2. Eigenvektoren $\mathbf{x}_1, \mathbf{x}_2, \dots$ zu verschiedenen Eigenwerten $\lambda_1, \lambda_2, \dots$ (endlich oder unendlich viele) von f sind linear unabhängig.

Beweis. 1. Aus $f(\mathbf{x}) = \lambda\mathbf{x}$ und $f(\mathbf{y}) = \lambda\mathbf{y}$ folgt $f(\mathbf{x} + \mathbf{y}) = \lambda(\mathbf{x} + \mathbf{y})$, außerdem $f(r\mathbf{x}) = \lambda(r\mathbf{x})$. Folglich ist $E_{f,\lambda}$ ein Unterraum.

2. Die zweite Aussage beweisen wir mit Induktion nach der zunächst als endlich angenommenen Anzahl n der gegebenen Eigenvektoren (Eigenwerte).

$n = 0$: Die leere Menge von Vektoren ist definitionsgemäß linear unabhängig ist.

$n = 1$: Ein einzelner Vektor bildet, sofern es sich nicht um den Nullvektor handelt, stets eine linear unabhängige Einermenge. Und \mathbf{x}_1 ist als Eigenvektor definitionsgemäß nicht der Nullvektor.

Schritt von n auf $n + 1$: Sei die Aussage für ein bestimmtes $n \in \mathbb{N}$ wahr. Wir haben sie unter dieser Annahme für $n + 1$ zu beweisen. Dazu gehen wir von einer Darstellung des Nullvektors \mathbf{o} als Linearkombination

$$(1) \quad \mathbf{o} = \sum_{i=1}^{n+1} c_i \mathbf{x}_i$$

mit irgendwelchen Koeffizienten $c_i \in \mathbb{R}$ aus. Anwendung von f liefert

$$(2) \quad \mathbf{o} = f(\mathbf{o}) = f\left(\sum_{i=1}^{n+1} c_i \mathbf{x}_i\right) = \sum_{i=1}^{n+1} c_i f(\mathbf{x}_i) = \sum_{i=1}^{n+1} c_i \lambda_i \mathbf{x}_i.$$

Wir haben zwei Fälle zu unterscheiden.

Fall 1, $\lambda_{n+1} = 0$. Dann fällt der letzte Summand in (2) weg, und wir haben

$$\mathbf{o} = \sum_{i=1}^n c_i \lambda_i \mathbf{x}_i.$$

Laut Induktionsannahme sind die Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ linear unabhängig, folglich gilt $c_i \lambda_i = 0$ für $i = 1, \dots, n$. Weil die λ_i für $i = 1, \dots, n + 1$ paarweise verschieden sind, gilt $\lambda_i \neq \lambda_{n+1} = 0$ für $i = 1, \dots, n$, woraus $c_i = 0$ für $i = 1, \dots, n$ folgt, also

$$(3) \quad \mathbf{o} = \sum_{i=1}^{n+1} c_i \mathbf{x}_i = c_{n+1} \mathbf{x}_{n+1}.$$

Als Eigenvektor ist $\mathbf{x}_{n+1} \neq \mathbf{o}$, womit auch $c_{n+1} = 0$ bewiesen ist. Also sind die Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_{n+1}$ linear unabhängig, wie behauptet.

Fall 2, $\lambda_{n+1} \neq 0$: Wir multiplizieren (1) mit λ_{n+1} und subtrahieren das Ergebnis von (2), womit der letzte Summand wegfällt:

$$\mathbf{o} = \sum_{i=1}^{n+1} c_i (\lambda_i - \lambda_{n+1}) \mathbf{x}_i = \sum_{i=1}^n c_i (\lambda_i - \lambda_{n+1}) \mathbf{x}_i.$$

Wieder verwenden wir die Induktionsannahme, dass $\mathbf{x}_1, \dots, \mathbf{x}_n$ linear unabhängig sind. Das liefert $c_i(\lambda_i - \lambda_{n+1}) = 0$ für $i = 1, \dots, n$. Wegen der Verschiedenheit der Eigenwerte ist $\lambda_i - \lambda_{n+1} \neq 0$, also $c_i = 0$ für $i = 1, \dots, n$. Somit wird (1) wieder zu (3) mit derselben Konsequenz, nämlich $c_{n+1} = 0$, was wieder die lineare Unabhängigkeit von $\mathbf{x}_1, \dots, \mathbf{x}_{n+1}$ zeigt.

Damit ist die Behauptung für endlich viele \mathbf{x}_i und λ_i bewiesen. Sie überträgt sich aber auch auf unendlich viele, weil lineare Unabhängigkeit in diesem Fall ja definitionsgemäß die lineare Unabhängigkeit je endlich vieler bedeutet. \square

Damit ist aber noch wenig über die Berechnung von Eigenwerten und -vektoren gesagt, welche Gegenstand des nun folgenden Unterabschnitts ist.

1.6.2 Berechnungsmethoden

Inhalt in Kurzfassung: Die Eigenwerte erhält man als Nullstellen des sogenannten charakteristischen Polynoms. Zu jedem Eigenwert ergibt sich der zugehörige Eigenraum als Lösungsmenge eines homogenen linearen Gleichungssystems. Die Vielfachheit als Nullstelle des charakteristischen Polynoms heißt die algebraische Vielfachheit eines Eigenwertes, die Dimension des Eigenraums heißt die geometrische Vielfachheit des Eigenwertes.

Für eine quadratische Matrix A ist \mathbf{x} genau dann ein Eigenvektor zum Eigenwert λ , wenn \mathbf{x} im Kern von $A - \lambda I_n$ liegt. Also hat A den Eigenwert λ genau dann, wenn es $A - \lambda I_n$ singulär ist, d.h. wenn $\det(A - \lambda I_n) = 0$. Daraus ergibt sich:

Proposition 1.6.2.1. *Die Eigenwerte einer reellen oder komplexen (n, n) -Matrix A sind genau die Zahlen λ in \mathbb{R} oder \mathbb{C} mit*

$$\text{rang}(A - \lambda I) < n$$

oder

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Betrachtet man in Proposition 1.6.2.1 λ als Variable und multipliziert man die Determinante gemäß der Summenformel aus 1.4.5.3 aus, so erhält man ein Polynom in λ vom Grad n , dessen Nullstellen also gesucht sind.

Dieses Polynom $\det(A - \lambda I)$ nennt man das **charakteristische Polynom** von A und die Gleichung

$$\det(A - \lambda I) = 0$$

die **charakteristische Gleichung** von A . Also:

Satz 1.6.2.2. *Die Eigenwerte einer $n \times n$ -Matrix A sind genau die Nullstellen des charakteristischen Polynoms von A . Insbesondere gibt es für ungerades n stets einen reellen Eigenwert.*

Beweis. Die erste Aussage folgt unmittelbar aus der vorangegangenen Diskussion. Die zweite folgt deshalb, weil reelle Polynome ungeraden Grades immer wenigstens eine reelle Nullstelle haben. \square

Übungsaufgabe 103. (E) Begründen Sie die im Beweis von Satz 1.6.2.2 verwendete Aussage, dass jedes reelle Polynom $p(x)$ ungeraden Grades mindestens eine reelle Nullstelle hat. Hinweis: Argumentieren Sie, dass $p(x)$ für betragsmäßig große negative und positive Werte entgegengesetztes Vorzeichen hat und verwenden Sie die Stetigkeit von f sowie den Zwischenwertsatz.

Übungsaufgabe 104. (E) Interpretieren Sie Satz 1.6.2.2 im Fall $n = 3$ geometrisch, insbesondere für orthogonales A .

Definition 1.6.2.3. Sei λ ein Eigenwert von A , also eine Nullstelle des charakteristischen Polynoms $p(x) = \det(A - xI)$. Die Vielfachheit von λ als Nullstelle von p heißt auch seine **algebraische Vielfachheit**, die Dimension des Eigenraums $E_{A,\lambda}$ heißt **geometrische Vielfachheit** von λ als Eigenwert von A . Analog sind algebraische und geometrische Vielfachheit von Eigenwerten linearer Abbildungen $f : V \rightarrow V$ eines endlichdimensionalen Vektorraums V auf sich definiert.

Beispiel. Man bestimme die Eigenwerte und die zugehörigen Eigenvektoren der Matrix

$$\begin{pmatrix} 6 & 4 & 4 \\ -1 & 0 & -1 \\ -2 & 0 & 0 \end{pmatrix}.$$

Charakteristische Gleichung:

$$\begin{vmatrix} 6 - \lambda & 4 & 4 \\ -1 & -\lambda & -1 \\ -2 & 0 & -\lambda \end{vmatrix} = -\lambda^3 + 6\lambda^2 - 12\lambda + 8 = -(\lambda - 2)^3 = 0.$$

Man erhält also den einzigen Eigenwert $\lambda = 2$ mit der algebraischen Vielfachheit 3. Die zugehörigen Eigenvektoren \mathbf{x} sind die Lösungen von folgendem Gleichungssystem

$$\begin{aligned} 4x_1 + 4x_2 + 4x_3 &= 0 \\ -x_1 - 2x_2 - x_3 &= 0 \\ -2x_1 - 2x_3 &= 0. \end{aligned}$$

Gaußelimination liefert

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ -x_2 &= 0 \\ x_2 &= 0, \end{aligned}$$

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_2 &= 0 \\ 0 &= 0, \end{aligned}$$

also $x_3 = -t$, $x_2 = 0$, $x_1 = t$ ($t \in \mathbb{R}$, beliebig).

Lösung:

$$\mathbf{x} = t \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad t \in \mathbb{R} \text{ bel.}$$

Es gibt zu λ also nur einen linear unabhängigen Eigenvektor. Somit hat λ die geometrische Vielfachheit 1.

Algebraische und geometrische Vielfachheit eines Eigenwerts müssen also nicht übereinstimmen. Die geometrische kann auch kleiner sein, nicht aber größer als die algebraische.

Als Beispiele ohne reelle Eigenwerte dienen Drehungen in der Ebene um einen Winkel α mit $\sin(\alpha) \neq 0$, der also kein ganzzahliges Vielfaches von π ist. Das charakteristische Polynom der zugehörigen Drehmatrix

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat die Gestalt

$$\det(A - \lambda I) = \begin{vmatrix} \cos(\alpha) - \lambda & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) - \lambda \end{vmatrix} = (\cos(\alpha) - \lambda)^2 + \sin^2(\alpha) = \lambda^2 - 2\lambda \cos(\alpha) + 1.$$

Die Nullstellen sind nach der quadratischen Lösungsformel gegeben durch

$$\lambda_{1,2} = \cos(\alpha) \pm \sqrt{\cos^2(\alpha) - 1} = \cos(\alpha) \pm \sqrt{-\sin^2(\alpha)} = \cos(\alpha) \pm i \sin(\alpha),$$

für $\sin(\alpha) \neq 0$ (d.h. $\alpha \neq k\pi$ mit $k \in \mathbb{Z}$), also nicht reell.

Eine interessante geometrische Einsicht ergibt sich, wenn man quadratische $n \times n$ -Matrizen A ungerader Dimension n betrachtet, wo also auch das charakteristische Polynom ungeraden Grad n hat. Ein solches Polynom hat jedenfalls mindestens eine reelle Nullstelle. A hat also einen reellen Eigenwert λ mit zugehörigen Eigenvektoren, die durch A lediglich um den Faktor λ gestreckt werden, für $\lambda = 1$ sogar fest bleiben. Weitere Eigenvektoren müssen aber nicht auftreten, wie das Beispiel von Drehungen um eine Drehachse als eindimensionalem Eigenraum belegen.

Übungsaufgabe 105. (T) Bestimmen Sie die Eigenwerte und Eigenvektoren folgender Matrizen. Für die Matrix C müssen nur die Eigenwerte (und nicht die Eigenvektoren) berechnet werden.

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ -2 & -1 & 1 \\ 0 & -2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 4 & -2 & -2 & 3 \end{pmatrix}$$

Übungsaufgabe 106. (P) Geben Sie für die folgenden, geometrisch definierten, linearen Abbildungen im \mathbb{R}^2 jeweils alle Eigenwerte und -vektoren an:

1 Lineare Algebra

1. Spiegelung am Ursprung.
2. Spiegelung an der Gerade mit der Gleichung $ax + by = 0$.
3. Orthogonale Projektion auf die Gerade mit der Gleichung $ax + by = 0$.
4. Streckung um den Faktor λ .
5. Drehung um den Winkel ϕ .

Übungsaufgabe 107. (E) Die Drehung im \mathbb{R}^3 um die Achse \vec{n} (mit $\vec{n} \cdot \vec{n} = 1$) und den Winkel φ wird durch

$$R_{\vec{n},\varphi}(\vec{x}) = \vec{n}(\vec{n} \cdot \vec{x}) + (\vec{n} \times \vec{x}) \times \vec{n} \cos \varphi + (\vec{n} \times \vec{x}) \sin \varphi$$

beschrieben. Berechnen Sie Spur, Eigenwerte und Eigenvektoren von $R_{\vec{n},\varphi}$.

Anleitung: Untersuchen Sie zuerst den Axialvektor \vec{n} und überlegen Sie sich, dass die Ebene N mit Normalvektor \vec{n} fix bleibt, d.h. wenn $\vec{x} \cdot \vec{n} = 0$, dann auch $R_{\vec{n},\varphi}(\vec{x}) \cdot \vec{n} = 0$. Rechnen Sie nach, dass für festes x_0 die beiden Vektoren $\vec{b}_1 = (\vec{n} \times \vec{x}_0) \times \vec{n}$ und $\vec{b}_2 = (\vec{n} \times \vec{x}_0)$ eine Orthogonalbasis von N bilden. Wie sieht nun $R_{\vec{n},\varphi}$ bezüglich der Basis $\{\vec{b}_1, \vec{b}_2, \vec{n}\}$ aus?

Übungsaufgabe 108. (T) Berechnen Sie für folgende orthogonale Matrizen Determinante, Drehwinkel und -achse:

$$T_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad T_2 = \frac{1}{15} \begin{pmatrix} 5 & 10 & 10 \\ -2 & 11 & -10 \\ -14 & 2 & 5 \end{pmatrix}$$

Übungsaufgabe 109. (T) Überprüfen Sie, für welche Parameterwerte α die folgende Matrix eine Drehung im \mathbb{R}^3 beschreibt und bestimmen Sie Drehachse und -winkel:

$$\frac{1}{9} \begin{pmatrix} 8 & 1 & \alpha \\ -\alpha & -4 & 7 \\ -1 & -8 & -4 \end{pmatrix}.$$

1.6.3 Ähnlichkeit und Diagonalform

Inhalt in Kurzfassung: Zwei $n \times n$ -Matrizen A und B heißen ähnlich, wenn sie bezüglich geeigneter Basen von \mathbb{R}^n dieselbe lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ darstellen. Das ist genau dann der Fall, wenn es eine der Basistransformation entsprechende reguläre Matrix S mit $B = SAS^{-1}$ gibt. Von besonderem Interesse sind dabei diagonalisierbare Matrizen A , die nämlich ähnlich sind einer Diagonalmatrix. Das ist definitionsgemäß eine quadratische $n \times n$ -Matrix $A = (a_{ij})$ mit $a_{ij} = 0$ für alle $i \neq j$, die von 0 verschiedene Eintragungen also nur in der Diagonale hat. Für eine Diagonalmatrix sind die $\lambda_i := a_{ii}$ genau die Eigenwerte von A mit den kanonischen Basisvektoren \mathbf{e}_i als zugehörigen Eigenvektoren. Eine Matrix A ist also diagonalisierbar, wenn es eine reguläre Matrix S

mit $D = SAS^{-1}$ bzw. $A = S^{-1}DS$ gibt (Diagonalisierung von A). Manche Operationen lassen sich mit diagonalisierbaren Matrizen leichter ausführen, insbesondere die Berechnung von Potenzen: $A^n = (SDS^{-1})^n = SD^nS^{-1}$, wobei D^n die Diagonalmatrix mit den Diagonalelementen λ_i^n ist. Nicht alle quadratischen Matrizen sind diagonalisierbar, sie sind aber ähnlich zu einer Matrix in sogenannter Jordanscher Normalform. Jedenfalls diagonalisierbar sind symmetrische Matrizen. Sie besitzen sogar eine orthogonale Basis aus Eigenvektoren zu (durchwegs reellen) Eigenwerten.

Wir wissen bereits, wie sich ein Basiswechsel auf die Darstellung linearer Abbildungen f durch Matrizen auswirkt. Geht es um Transformationen $f : V \rightarrow V$ eines n -dimensionalen Vektorraums V , und sind A und B die Darstellungen von f bezüglich zweier (geordneter) Basen, so gilt $B = SAS^{-1}$ bzw. $A = S^{-1}BS$ mit einer regulären $n \times n$ -Matrix S , welche dem Basiswechsel entspricht, siehe 1.3.5.

Diese Beziehung zwischen den Matrizen A und B motiviert den Begriff der Ähnlichkeit von Matrizen.

Definition 1.6.3.1. *Zwei quadratische $n \times n$ -Matrizen heißen **ähnlich**, symbolisch $A \sim B$, wenn es eine reguläre $n \times n$ -Matrix S gibt mit $B = SAS^{-1}$.*

Stets gilt $A \sim IAI^{-1}$, also $A \sim A$ (\sim ist reflexiv). Aus $A \sim B$ folgt stets $B \sim A$ (denn $B = SAS^{-1}$ kann zu $A = S^{-1}BS$ umgeschrieben werden; \sim ist symmetrisch). Außerdem folgt aus $A \sim B$ und $B \sim C$ auch $A \sim C$ (denn $B = SAS^{-1}$ und $C = TBT^{-1}$ liefert $C = TSAS^{-1}T^{-1} = (TS)A(TS)^{-1}$; \sim ist transitiv). Deshalb ist \sim eine sogenannte **Äquivalenzrelation**. Das hat zur Folge, dass man die Menge aller $n \times n$ -Matrizen in Klassen einteilen kann, von denen jede aus zueinander ähnlichen Matrizen besteht, während zwei Matrizen aus verschiedenen Klassen nie äquivalent zueinander sind.

Übungsaufgabe 110. *(E) Man beweise die hier implizit ausgesprochene Behauptung: Jede Äquivalenzrelation (eine reflexive, symmetrische, transitive Relation) auf einer beliebigen Menge gibt in der beschriebenen Weise Anlass zu einer Zerlegung dieser Menge in Klassen derart, dass zwei Elemente genau dann in Relation zueinander stehen, wenn sie zur selben Klasse gehören.*

Übungsaufgabe 111. *(T) Geben Sie die **Matrixdarstellung** jener linearen Abbildung $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ an, welche einer Drehung um 90° gegen den Uhrzeigersinn um die Achse in Richtung $(1, 0, 1)$ entspricht!*

Übungsaufgabe 112. *(T) Die Abbildung f werde bezüglich der Basis $\mathbf{e}_1, \mathbf{e}_2$ durch folgende Matrix A beschrieben:*

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}.$$

1. Transformieren Sie die Matrix A in jene Matrix B , welche die Abbildung f bezüglich jener Basis beschreibt, die aus den beiden Vektoren $\begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix}$ und $\begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix}$ besteht. (Dabei sei $(\phi \in [0, 2\pi])$ irgendein fester Winkel.)

1 Lineare Algebra

2. Wählen Sie ϕ so, dass f bezüglich einer neuen Basis durch eine Diagonalmatrix beschrieben wird. Wie lauten die Eigenvektoren von f bezüglich dieser Basis?

Diagonalmatrizen sind besonders einfach und übersichtlich zu handhaben. Es ist daher naheliegend zu versuchen, zu einer Matrix eine ähnliche Diagonalmatrix zu finden. Das ist nicht immer möglich. Wie man sich unmittelbar überlegt, gilt aber:

Proposition 1.6.3.2. *Es sei A eine reelle oder komplexe (n, n) -Matrix. A ist genau dann ähnlich zu einer Diagonalmatrix, wenn sie n linear unabhängige Eigenvektoren besitzt. Gilt in diesem Fall etwa*

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \dots, A\mathbf{x}_n = \lambda_n\mathbf{x}_n$$

mit den linear unabhängigen Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$, so gilt mit der Matrix S , deren Spaltenvektoren $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ sind,

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Beispiel. Man diagonalisiere die Matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}.$$

Eigenwerte:

$$\begin{aligned} & \begin{vmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 4 \\ -1 & -1 & -2-\lambda \end{vmatrix} = \\ & = (2-\lambda)(3-\lambda)(-2-\lambda) - 4 - 2 + (2-\lambda)4 + (2+\lambda)2 + (3-\lambda) \\ & = (1-\lambda)(1+\lambda)(\lambda-3) = 0 \end{aligned}$$

$$\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 3$$

Eigenvektoren:

$$\lambda_1 = 1: \begin{array}{rcl} x_1 + x_2 + x_3 & = & 0 \\ 2x_1 + 2x_2 + 4x_3 & = & 0 \\ -x_1 - x_2 - 3x_3 & = & 0 \end{array}, \text{ also } \mathbf{x}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad (\text{und alle Vielfachen})$$

$$\lambda_2 = -1: \begin{array}{rcl} 3x_1 + x_2 + x_3 & = & 0 \\ 2x_1 + 4x_2 + 4x_3 & = & 0 \\ -x_1 - x_2 - x_3 & = & 0 \end{array} \quad \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$\lambda_3 = 3 : \begin{array}{rcl} -x_1 + x_2 + x_3 & = & 0 \\ 2x_1 & + & 4x_3 = 0 \\ -x_1 - x_2 - 5x_3 & = & 0 \end{array} \quad \mathbf{x}_3 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

$\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ sind linear unabhängig. Damit bildet man die Matrix

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

Es folgt

$$S^{-1}AS = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Ist etwa A die Matrix einer Lineartransformation bezüglich der Basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ von \mathbb{R}^3 , und führt man die Vektoren $\mathbf{x}_1, \mathbf{x}_2$ und \mathbf{x}_3 von oben als Elemente einer neuen Basis ein, so nimmt die Lineartransformation die Gestalt

$$\mathbf{y}' = B\mathbf{x}' \quad \text{mit} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

an, wird also besonders einfach dargestellt.

Beispiel (Matrizenpotenzen). Es sei A wie eben. Man berechne A^n für großes n . Es ist

$$A^n = AA \cdots A = SS^{-1}ASS^{-1}AS \cdots S^{-1}ASS^{-1} = SB^nS^{-1} = S \begin{pmatrix} 1 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 3^n \end{pmatrix} S^{-1}.$$

Übungsaufgabe 113. (T) Berechnen Sie die Potenzen A^4 und A^{-12} von $A := \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ sowie deren Eigenvektoren und Eigenwerte!

Übungsaufgabe 114. (T) Berechnen Sie für die Matrix $A := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ induktiv alle Potenzen A^2, A^3, \dots, A^n . Wie lautet der Wert der unendlichen Reihe

$$\exp(A) := \frac{1}{0!}I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \frac{1}{4!}A^4 + \dots$$

Übungsaufgabe 115. (T) Berechnen Sie für die Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$$

$\exp(A) := \frac{1}{0!}I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \frac{1}{4!}A^4 + \dots$ und $\exp(B)$, indem Sie auf Diagonalform transformieren.

Bemerkung. Ist eine (n, n) -Matrix A nicht ähnlich zu einer Diagonalmatrix, so versucht man, andere Standardmatrizen von vergleichsweise einfacher Bauart zu definieren, so dass jede beliebige quadratische Matrix zu (genau) einer solchen ähnlich ist. Arbeitet man über den komplexen Zahlen (dabei spielt der Fundamentalsatz eine entscheidende Rolle) so ist stets die Transformation auf die sogenannte **Jordan'sche Normalform** möglich.

Definition 1.6.3.3. Eine quadratische Matrix $A = (a_{i,j})$ heißt **symmetrisch**, wenn sie mit ihrer Transponierten übereinstimmt, wenn also $A^T = A$ bzw., äquivalent, $a_{i,j} = a_{j,i}$ für alle Paare (i, j) von Zeilen- und Spaltenindizes gilt.

Nicht offensichtlich ist der folgende Satz, den wir seiner Bedeutung für später wegen angeben, wenn auch ohne Beweis:

Satz 1.6.3.4. Zu jeder symmetrischen (n, n) -Matrix A gibt es eine Orthonormalbasis aus Eigenvektoren zu reellen Eigenwerten. A ist also ähnlich zu einer reellen Diagonalmatrix D über eine orthogonale Basistransformation S :

$$D = S^{-1}AS$$

Dieser Satz wird im Zusammenhang mit Extremwertaufgaben in mehreren Variablen noch eine bemerkenswerte Rolle spielen, Schlagwort Hauptachsentransformation.

2 Differentialrechnung in mehreren Variablen

Die Übertragung der Differentialrechnung auf Funktionen in mehreren Variablen und mit Werten in einem höherdimensionalen Raum orientiert sich an der aus Mathematik 1 bereits bekannten Idee der lokalen Approximation differenzierbarer reeller Funktionen durch lineare. Dieses Paradigma der Linearisierung zieht sich durch große Teile der gesamten Mathematik und beherrscht insbesondere das vorliegende Kapitel. Man könnte es frei nach Galilei¹ zur Sentenz komprimieren: Was linear ist, mit Hilfe der Linearen Algebra berechnen; was nicht linear ist, mit Hilfe der Differentialrechnung linear machen!

Um die Konzepte nicht nur formal zu verstehen, sondern auch anschaulich zu erfassen, beginnen wir mit einem Abschnitt über die Geometrie im \mathbb{R}^n (siehe 2.1). Die eigentliche Differentialrechnung beginnt mit Abschnitt 2.2, wo die grundlegenden Definitionen sowie die wichtigsten Tatsachen und Beispiele präsentiert werden. In späteren Anwendungen werden wir die Ergebnisse aus 2.3 über nichtlineare Gleichungssysteme in mehreren Variablen und die damit verbundenen Hauptsätze über Umkehr- und über implizite Funktionen verwenden bzw. darauf verweisen. 2.4, der letzte Abschnitt des Kapitels, beschäftigt sich mit dem Problemkreis Extremwertaufgaben.

Ergänzend noch eine Bemerkung zu einer Uneinheitlichkeit in der Notation: Im Kapitel über Lineare Algebra sind wir der weit verbreiteten Konvention gefolgt, die Zahl der Zeilen einer Matrix mit n , die der Spalten mit m zu bezeichnen. Das entspricht Funktionen $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$. In der Differentialrechnung in mehreren Variablen wird deren Anzahl meist mit n bezeichnet. Das wiederum entspricht Funktionen $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}^m$, d.h. mit vertauschten Rollen von m und n . Große Teile dieses Kapitels (noch nicht zu Beginn) folgen dieser Tradition. Es empfiehlt sich daher, stets darauf zu achten, was gerade m und was n ist, und gedanklich weniger an diesen beiden Buchstaben zu kleben als an den Konzepten Definitionsbereich bzw. Zielbereich.

2.1 Nichtlineare Geometrie im \mathbb{R}^n

In diesem Abschnitt werden wichtige Beispiele von höherdimensionalen Funktionen $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^m$ behandelt und mit ihnen verknüpfte geometrische Vorstellungen und Intuitionen entwickelt. Nach einer Motivation (2.1.1) und der Übertragung begrifflicher Grundlagen wie Grenzwert von Folgen und Stetigkeit von Funktionen vom

¹Galileo Galilei (1564-1642) forderte für eine moderne, empirisch-experimentelle Naturwissenschaft, die von den Früchten der Mathematik profitieren kann: „Was messbar ist, messen; was nicht messbar ist, messbar machen!“ Galilei gilt damit als ein Ahnherr der wissenschaftlichen Revolution der Neuzeit, die etwa mit ihm und seiner Generation einsetzte.

eindimensionalen Fall (siehe Mathematik 1) auf den höherdimensionalen (2.1.2) werden verschiedene Fälle für m und n durchgespielt: reellwertige Funktionen in m Variablen in 2.1.3, Kurven und Flächen und ihre Parametrisierung in 2.1.4 sowie für den Fall $m = n$ Koordinatentransformationen in 2.1.5 und Vektorfelder in 2.1.6.

2.1.1 Motivation

Inhalt in Kurzfassung: Die Programmatik des vorliegenden Kapitels lässt sich als eine Synthese eindimensionaler Differentialrechnung und höherdimensionaler Linearer Algebra deuten, aus der die höherdimensionale Differentialrechnung entsteht.

Im Kapitel über Lineare Algebra haben wir uns bereits mit Abbildungen zwischen höherdimensionalen Bereichen, insbesondere zwischen den Vektorräumen $\mathbb{R}^m, \mathbb{R}^n, \dots$ mit unterschiedlichen Dimensionen m, n, \dots beschäftigt. In der *Linearen* Algebra beschränkt man sich dabei im Wesentlichen aber auf lineare Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, die sich in natürlicher Weise durch Matrizen darstellen und damit effizient untersuchen lassen.

Aus Mathematik 1 wissen wir, dass viele interessante Abbildungen (Funktionen) nicht linear sind. Insbesondere im Kapitel über Differentialrechnung haben wir aber gesehen, dass es immerhin möglich ist, viele nichtlineare Abbildungen durch lineare so gut zu approximieren, dass sich bei weiterer Entwicklung der Theorie dabei ungeahnte Möglichkeiten eröffnen. Man erinnere sich beispielsweise daran, wie wir über den Mittelwertsatz der Differentialrechnung unter Verwendung höherer Ableitungen zu Taylorreihen und Potenzreihendarstellungen der wichtigsten Funktionen gelangt sind.

Im vorliegenden Kapitel geht es um eine Verbindung der Möglichkeiten aus der Differentialrechnung in einer Variablen mit den Methoden der Linearen Algebra. Damit lassen sich auch nichtlineare höherdimensionale Phänomene mathematisch erfassen.

Bei der Entwicklung der Differentialrechnung in mehreren Variablen erweist es sich immer wieder als hilfreich, die in der Linearen Algebra gewonnenen geometrischen Intuitionen weiter zu entwickeln und auf nichtlineare Phänomene auszuweiten.

Mathematik erweist sich immer dann als besonders fruchtbar, wenn begriffliche und methodische Strenge mit lebendiger Anschauung in Verbindung tritt. In einem ersten Schritt werden wir deshalb wichtige Begriffe wie Folgenreizwerte, Stetigkeit etc. vom eindimensionalen auf den mehrdimensionalen Fall übertragen und geometrisch deuten. Sodann werden wir typische Beispiele mit variierenden Dimensionen m und n durchspielen.

2.1.2 Höherdimensionale Folgen und stetige Funktionen

Inhalt in Kurzfassung: Weil die n -dimensionalen Räume \mathbb{R}^n mittels des euklidischen Abstandsbegriffs zu metrischen Räumen gemacht werden können, lassen sich grundlegende Begriffe wie Grenzwert und Stetigkeit samt wichtigen Eigenschaften mühelos vom ein- auf den höherdimensionalen Fall übertragen.

Haben wir in Mathematik 1 vor allem reelle Funktionen, also Funktionen $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}$ untersucht, interessieren wir uns nun für Funktionen, deren Definitions- und Zielbereich höherdimensional sein dürfen. Unsere Objekte sind also Funktionen

$$f : D \rightarrow \mathbb{R}^m \quad \text{mit} \quad D \subseteq \mathbb{R}^n.$$

Nichts hindert uns daran, im Bildbereich jede Komponente einzeln zu betrachten. Wir schreiben $f_i : D \rightarrow \mathbb{R}$, $i = 1, \dots, m$, für jene reellwertigen Funktionen mit

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_m(x_1, \dots, x_n) \end{pmatrix}, \quad \mathbf{x} = (x_1, \dots, x_n) \in D.$$

Die f_i heißen auch die **Komponentenfunktionen** von f .

Viele interessante Eigenschaften von f lassen sich auf solche der f_i zurückführen. Das vereinfacht die Untersuchungen, weil man sich dann auf den psychologisch und notationell einfacheren Fall $m = 1$ konzentrieren kann. Die Dimension n des Definitionsbereichs von f lässt sich nicht so häufig in sinnvoller Weise auf $n = 1$ reduzieren. Ganz wesentliche Aspekte von Funktionen in mehreren Variablen gehen dabei verloren. Oft reicht aber ein Studium der Situation für $n = 2$ aus, um die wesentlichen Unterschiede zu verstehen. Wem höherdimensionale Funktionen Unbehagen verschaffen, möge also zunächst an die Situation $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^2$ denken. Aber auch der allgemeine Fall hat psychologische Vorteile, weil in der Notation die Symmetrie in den n Komponenten oft deutlicher zum Ausdruck kommt. Hin und wieder werden wir der speziellen Darstellungsweise mit $n = 2$ den Vorzug geben, häufiger der allgemeinen.

Eine kurze Wiederholung aus Mathematik 1 ist am Platz: Wir beginnen mit dem Begriff der Stetigkeit. Er lässt sich auf Funktionen zwischen beliebigen metrischen Räumen übertragen. Das wurde auch schon in Mathematik 1 angedeutet. Die Räume \mathbb{R}^n sind metrische Räume, weil (neben anderen oft praktischen Metriken) die (in geometrischer Hinsicht natürliche) euklidische Metrik zur Verfügung steht. Diese Zusammenhänge sollen nun rekapituliert werden, zunächst aus Kapitel 1.

Für einen Vektor $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ ist, motiviert durch den Satz von Pythagoras, die **euklidische Länge** oder **Norm** von \mathbf{x} definiert durch

$$\|\mathbf{x}\| := \sqrt{\sum_{i=1}^n x_i^2}.$$

Diese Norm nimmt niemals negative Werte an, und $\|\mathbf{x}\| = 0$ gilt nur für $\mathbf{x} = \mathbf{o}$. Außerdem ist die Norm homogen im Sinne von $\|r\mathbf{x}\| = |r| \cdot \|\mathbf{x}\|$ und erfüllt die Dreiecksungleichung $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. Daraus folgt, dass durch

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$$

2 Differentialrechnung in mehreren Variablen

eine **Metrik** oder ein **Abstand** $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_0^+$ zwischen \mathbf{x} und \mathbf{y} (interpretiert als Punkte im \mathbb{R}^n) definiert wird. Das bedeutet definitionsgemäß, dass $d(\mathbf{x}, \mathbf{y}) = 0$ nur für $\mathbf{x} = \mathbf{y}$ gilt, außerdem $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ und $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.

Damit lässt sich die Definition des Grenzwertes von Folgen vom Eindimensionalen direkt übertragen:

Definition 2.1.2.1. Eine Folge von Punkten $\mathbf{x}_n \in \mathbb{R}^m$, $n \in \mathbb{N}$, **konvergiert** gegen $\mathbf{x} \in \mathbb{R}^m$, wenn es zu jedem $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$ die Ungleichung $d(\mathbf{x}_n, \mathbf{x}) = \|\mathbf{x}_n - \mathbf{x}\| < \varepsilon$ gilt. In diesem Fall heißt \mathbf{x} der **Grenzwert** der Folge, symbolisch:

$$\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}$$

Explizit bedeutet das (Wiederholung der in Mathematik 1 gepflegten Schreibweise mit logischen Quantoren):

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : d(\mathbf{x}_n, \mathbf{x}) < \varepsilon$$

Wann immer nützlich und auf offensichtliche Weise möglich, werden wir neben dem Grenzwert selbst auch mit ihm verwandte Begriffe aus dem Eindimensionalen sinngemäß übernehmen. Der Begriff der Stetigkeit sei wegen seiner besonderen Wichtigkeit explizit hervorgehoben:

Definition 2.1.2.2. Eine Funktion $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, (oder allgemeiner zwischen zwei metrischen Räumen) heißt **stetig** im Punkt $\mathbf{x}_0 \in D$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass für alle $\mathbf{x} \in D$ gilt:

$$\text{Aus } d(\mathbf{x}_0, \mathbf{x}) < \delta \text{ folgt } d(f(\mathbf{x}_0), f(\mathbf{x})) < \varepsilon.$$

Ist f in allen Punkten $\mathbf{x}_0 \in D$ stetig, so heißt f (schlechthin) stetig. Auch hier als Formel mit logischen Quantoren:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall \mathbf{x} \in D : d(\mathbf{x}_0, \mathbf{x}) < \delta \rightarrow d(f(\mathbf{x}_0), f(\mathbf{x})) < \varepsilon$$

Wie im Eindimensionalen lässt sich Stetigkeit durch **Folgenstetigkeit** charakterisieren:

Proposition 2.1.2.3. Eine Funktion $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, (oder allgemeiner zwischen zwei metrischen Räumen) ist stetig im Punkt $\mathbf{x} \in D$ genau dann, wenn für jede Folge von Punkten $\mathbf{x}_n \in D$ mit $\lim_{n \rightarrow \infty} \mathbf{x}_n = \mathbf{x}$ auch $\lim_{n \rightarrow \infty} f(\mathbf{x}_n) = f(\mathbf{x})$ gilt.

Der Beweis von Proposition 2.1.2.3 erfolgt ganz analog zum Eindimensionalen und darf zur Übung empfohlen werden.

Übungsaufgabe 116. (E) Beweisen Sie Proposition 2.1.2.3.

Ebenso lassen sich weitere wichtige Sätze aus der eindimensionalen Analysis auf höhere Dimensionen übertragen, wobei die Beweise nur geringfügige Modifikationen erfordern. Zum Beispiel ist die Verkettung stetiger Funktionen selbst wieder stetig. Weitere Beispiele sind der **Satz von Bolzano-Weierstraß**: Jede beschränkte unendliche Folge in \mathbb{R}^n , $n \in \mathbb{N}$, hat einen Häufungspunkt in \mathbb{R}^n . Außerdem: Zu jedem Häufungspunkt einer Folge lässt sich eine gegen diesen konvergente Teilfolge auswählen. Des weiteren gilt der **Satz vom Maximum**: Reellwertige stetige Funktionen nehmen auf einem kompakten (d.h. in \mathbb{R}^n : abgeschlossenen und beschränkten) Definitionsbereich ein Minimum und ein Maximum an.

Übungsaufgabe 117. (E) Wiederholen Sie die im obigen Absatz verwendeten Begriffe und versuchen Sie, die Behauptungen zu beweisen.

Konvergenz und Stetigkeit lassen sich auf die einzelnen Komponenten reduzieren:

Proposition 2.1.2.4. Sei eine Folge von Punkten $\mathbf{x}_n = (x_{n,1}, \dots, x_{n,m}) \in \mathbb{R}^m$, $n \in \mathbb{N}$, gegeben, außerdem $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$. Dann gilt

$$\lim_{n \rightarrow \infty} \mathbf{x}_n = \lim_{n \rightarrow \infty} \begin{pmatrix} x_{n,1} \\ \vdots \\ x_{n,m} \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} =: \mathbf{x}$$

genau dann, wenn für alle $i = 1, \dots, m$ die Konvergenz

$$\lim_{n \rightarrow \infty} x_{n,i} = x_i$$

vorliegt.

Eine Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ ist im Punkt $\mathbf{x} \in D$ stetig genau dann, wenn jede der Komponentenfunktionen $f_i : D \rightarrow \mathbb{R}$ von f , $i = 1, \dots, m$, im Punkt \mathbf{x} stetig ist.

Übungsaufgabe 118. (E) Beweisen Sie Proposition 2.1.2.4.

Laut Proposition 2.1.2.4 genügt es im Hinblick auf Stetigkeit, reellwertige Funktionen $f : D \rightarrow \mathbb{R}$ zu beherrschen, die allerdings von mehreren Variablen x_1, \dots, x_n abhängen können. In den meisten interessanten Fällen greifen hier ganz ähnliche Argumente wie im Eindimensionalen. Anstatt die Theorie in aller Ausführlichkeit nochmals zu entwickeln, sei das an einem einfachen Beispiel illustriert:

Beispiel 2.1.2.5. Die Funktion $f(x_1, x_2) := \sin(x_1 x_2)$ ist stetig auf ganz \mathbb{R}^2 . Begründung: Der Grenzwertsatz für Produkte aus Mathematik 1 besagt in Verbindung mit Proposition 2.1.2.3 (Folgenstetigkeit), dass die Funktion $m : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x_1, x_2) \mapsto x_1 x_2$ stetig auf ganz \mathbb{R}^2 ist. Weiters ist die Funktion $\sin : \mathbb{R} \rightarrow \mathbb{R}$, $y \mapsto \sin y$ stetig. Weil $f = \sin \circ m$ die Verkettung der stetigen Funktionen \sin und m ist, folgt auch die Stetigkeit von f selbst.

Argumentationen dieser Art werden wir in Zukunft mit kurzen Wendungen abhandeln wie: „Die Funktion f setzt sich in stetiger Weise aus stetigen Funktionen zusammen, ist also selbst stetig.“

2.1.3 Reellwertige Funktionen – Funktionsgebirge

Inhalt in Kurzfassung: Reellwertige Funktionen $f : D \rightarrow \mathbb{R}$ auf einem n -dimensionalen Bereich $D \subseteq \mathbb{R}^n$ lassen sich im Fall $n = 2$ sehr gut als Funktionsgebirge veranschaulichen. Diese Veranschaulichung liefert eine der wichtigsten Intuitionen für die Differentialrechnung in mehreren Variablen, der wir uns in späteren Abschnitten zuwenden werden. Von $n = 2$ auf $n > 2$ gelingt die Übertragung der gewonnenen Intuition weitgehend durch Analogieschlüsse.

Wie zum Beispiel in Proposition 2.1.2.4 zum Ausdruck gekommen ist, lassen sich viele Fragen im Zusammenhang mit höherdimensionalen Funktionen $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, zurückführen auf den Fall $m = 1$. Sodann kann man die meisten interessanten Phänomene, die für $n = 1$ noch nicht auftreten, schon für $n = 2$ sehr gut veranschaulichen. Deshalb ist es wertvoll, sich den Fall $n = 2$ und $m = 1$ besonders intensiv vor Augen zu führen und diesen durch Analogieschlüsse auf die Situation mit $n \geq 3$ zu übertragen.

Dabei stellt man sich den Definitionsbereich $D \subseteq \mathbb{R}^2$ als Teilmenge einer Grundebene (\mathbb{R}^2 als x - y -Ebene) vor. Die Funktionswerte $f(x, y)$ von f an Stellen $(x, y) \in D$ werden dann auf einer senkrechten, d.h. auf die x - y -Achse normal stehenden z -Achse aufgetragen. Der Zusammenhang $z = f(x, y)$ kommt anschaulich dadurch zum Ausdruck, dass der Punkt (x, y, z) als zu f gehörig markiert wird. (Das stimmt sogar streng formal, weil ja $f : D \rightarrow \mathbb{R}$ die Menge aller Punkte $((x, y), f(x, y)) = (x, y, f(x, y)) = (x, y, z)$ mit $z = f(x, y)$ ist.)

Man kann sich f also als Fläche = Oberfläche einer Landschaft vorstellen, der ein Punkt (x, y, z) angehört, wenn $z = f(x, y)$ die Seehöhe über den Ortskoordinaten x und y (die näherungsweise auch als geographische Länge und Breite aufgefasst werden können) ist. Bei negativen Funktionswerten sinkt die Landschaft unter das Nullniveau. Stellen wir uns das Nullniveau als Meeresspiegel vor, so beschreiben negative Werte von f die Tiefe des Meeresgrundes.

Stetigkeit von f bedeutet anschaulich, dass es keine senkrechten Wände gibt. Um eine solche Vorstellung wachzurufen, sprechen wir auch kurz von **Funktionsgebirgen**. Das passt sehr gut auf die Situation $m = 1$ und $n = 2$, wird wie gesagt aber auch für größere Werte von n hilfreiche Dienste leisten.

Übungsaufgabe 119. (T) Gegeben ist die Funktion

$$f(x, y) = \frac{2x^4y^4}{x^4 + y^4} \text{ für } (x, y) \neq (0, 0).$$

Bestätigen Sie, dass f in $(0, 0)$ stetig fortsetzbar ist. Geben Sie die Fortsetzung an!

Übungsaufgabe 120. (T) Gegeben ist die Funktion

$$f(x, y) = \frac{2x^2y^2}{x^4 + y^4} \text{ für } (x, y) \neq (0, 0).$$

Bestätigen Sie, dass f in $(0, 0)$ nicht stetig fortsetzbar ist, indem Sie zeigen, dass

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

nicht existiert.

Zur Unterscheidung von Vektorfeldern (deren Werte nicht Skalare, sondern Vektoren sind, siehe auch 2.1.6) nennt man reell-, d.h. skalarwertige Abbildungen $f : D \rightarrow \mathbb{R}$ ($D \subseteq \mathbb{R}^n$) auch **Skalarfelder**.

2.1.4 Kurven und Flächen

Inhalt in Kurzfassung: Kurven und Flächen im herkömmlichen Sinn lassen sich anschaulich als gekrümmte Geraden oder Ebenen beschreiben. Für eine hinreichend präzise mathematische Definition sind Parametrisierungen unerlässlich. Das sind Abbildungen von eindimensionalen Intervallen, zwei- oder auch k -dimensionalen Bereichen in den \mathbb{R}^n , typischerweise mit $k < n$. Von Parametrisierungen verlangt man mindestens Stetigkeit, oft aber auch (stetige oder auch höhere) Differenzierbarkeit.

Im Kapitel über Lineare Algebra haben wir geometrische Objekte mit linearer Struktur beschrieben, die zum Beispiel als Lösungsmengen linearer Gleichungssysteme auftreten. Sie haben die Gestalt $\mathbf{v} + U$, wobei U ein Unterraum der Dimension $\dim U = k$ eines Vektorraums V der Dimension $\dim V = n \geq k$ ist, meist $V = \mathbb{R}^n$. Ist $B = \{\mathbf{b}_1, \dots, \mathbf{b}_k\}$ eine Basis von U , so beschreibt

$$\mathbf{v} + U = \{\mathbf{v} + \mathbf{u} : \mathbf{u} \in U\} = \left\{ \mathbf{v} + \sum_{i=1}^k r_i \mathbf{b}_i : r_i \in \mathbb{R} \right\}$$

die k -dimensionale Ebene $\mathbf{v} + U$ als Teilmenge des n -dimensionalen Raumes \mathbb{R}^n . Die Punkte $\mathbf{v} + \sum_{i=1}^k r_i \mathbf{b}_i$ von $\mathbf{v} + U$ hängen von den k reellen Parametern r_1, \dots, r_k ab. Man spricht deshalb auch von einer **Parameterdarstellung**, in diesem Fall von $\mathbf{v} + U$. Für $n = 3$ haben wir es mit Teilmengen des Anschauungsraumes zu tun, Ebenen im üblichen Sinn für $k = 2$ und Geraden für $k = 1$. Der Extremfall $k = 3$ beschreibt den gesamten Raum, $k = 0$ den Punkt, genauer die einelementige Menge, die nur den Punkt \mathbf{v} enthält. Für $n = 2$ ist $k = 1$ der einzige nichttriviale Fall: eine Gerade in der Ebene \mathbb{R}^2 .

Weil wir uns ab nun nicht auf lineare Gebilde beschränken wollen, betrachten wir Lösungen nicht nur von linearen Gleichungssystemen. Ein einfaches und wohlbekanntes Beispiel ist die Gleichung $x^2 + y^2 = 1$. Bekanntlich handelt es sich bei der Lösungsmenge

$$K := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\} \subseteq \mathbb{R}^2$$

um den Einheitskreis K in der Ebene, das ist die Menge aller Punkte mit Abstand 1 vom Koordinatenursprung. Eine Parameterdarstellung des Kreises ist gegeben durch

$$K = \left\{ \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} : t \in [0, 2\pi) \right\}.$$

2 Differentialrechnung in mehreren Variablen

Bei dieser Darstellung sind alle Punkte $\begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$ verschieden. Jedenfalls wird K als Bildmenge unter einer Funktion f dargestellt, in diesem Fall von

$$f : D \rightarrow \mathbb{R}^2, \quad D = [0, 2\pi] \subseteq \mathbb{R} = \mathbb{R}^1, \quad f(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Dehnt man den Definitionsbereich D aus, so werden Punkte mehrmals durchlaufen, die Menge K ändert sich aber nicht. Oft ist es von Interesse, dass D kompakt (abgeschlossen und beschränkt) ist, was in diesem Fall mit $D = [0, 2\pi]$ der Fall wäre. Dann stimmten nur Anfangspunkt und Endpunkt überein: $f(0) = (1, 0) = f(2\pi)$. Generell beschreibt man Kurven im \mathbb{R}^n vorzugsweise mit Hilfe von stetigen Funktionen $f : D \rightarrow \mathbb{R}^n$ mit zusammenhängendem Definitionsbereich $D \subseteq \mathbb{R}$, häufig einem Intervall $D = [a, b]$ ($a < b \in \mathbb{R}$).

Ähnlich lassen sich gekrümmte Flächen im \mathbb{R}^3 als stetige Bilder zweidimensionaler Teilbereiche $D \subseteq \mathbb{R}^2$ der Ebene darstellen. Wenn wir uns zum Beispiel an den Kugelkoordinaten orientieren, so können wir die Oberfläche der Einheitskugel, die sogenannte zweidimensionale Sphäre S^2 , die durch

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

definiert ist, auch als Bild $S^2 = f(D)$ des Rechtecks $D := [0, 2\pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \subseteq \mathbb{R}^2$ unter der Funktion

$$f : (\alpha, \beta) \mapsto \begin{pmatrix} \cos(\alpha) \cos(\beta) \\ \sin(\alpha) \cos(\beta) \\ \sin(\beta) \end{pmatrix}$$

in Parameterform darstellen. Die Punkte $\begin{pmatrix} \cos(\alpha) \cos(\beta) \\ \sin(\alpha) \cos(\beta) \\ \sin(\beta) \end{pmatrix} \in S^2$ hängen nun von zwei

Parametern ab, nämlich von α und β .

Eine allgemeine Definition, die all diesen Situationen gerecht wird, ist die folgende.

Definition 2.1.4.1. Eine **Kurve** oder auch ein **Weg** im n -dimensionalen Raum \mathbb{R}^n ist eine auf einem Intervall $[a, b]$ definierte stetige Funktion $f : [a, b] \rightarrow \mathbb{R}^n$. Dabei heißen $f(a)$ und $f(b)$ **Anfangs-** bzw. **Endpunkt** der Kurve, und man nennt f auch einen Weg von $f(a)$ nach $f(b)$.

Eine Menge $M \subseteq \mathbb{R}^n$ heißt **wegzusammenhängend**, wenn es zu je zwei Punkten $\mathbf{a}, \mathbf{b} \in M$ einen Weg von \mathbf{a} nach \mathbf{b} gibt.

Unter einer k -dimensionalen **Fläche** im \mathbb{R}^n versteht man eine stetige Funktion $f : D \rightarrow \mathbb{R}^n$ mit einem Definitionsbereich $D \subseteq \mathbb{R}^k$. (Bei herkömmlichen zweidimensionalen Flächen im \mathbb{R}^3 ist $n = 3$ und $k = 2$.)

Zu Definition 2.1.4.1 sind mehrere Bemerkungen am Platz:

Erstens sei ausdrücklich betont, dass eine Kurve bzw. Fläche nicht allein durch die Punktmenge $f(D)$ gegeben ist, sondern dass immer auch eine Parametrisierung mitzudenken ist – obwohl man oft, ungenau, auch die Punktmenge als Kurve bzw. Fläche

bezeichnet. Diese Ungenauigkeit wird in solchen Fällen oft in Kauf genommen, wo man tatsächlich nur die Menge $f(D)$ im Sinn hat und sich kaum für die spezielle Gestalt der Funktion f , die in diesem Zusammenhang die **Parametrisierung** der Fläche genannt wird, interessiert. Man wäre genauso zufrieden mit irgendeiner anderen Parametrisierung $f_1 : D_1 \rightarrow \mathbb{R}^n$, sofern nur $f_1(D_1) = f(D)$ gilt. Entscheidend ist, dass es für $f(D)$ überhaupt eine Parametrisierung mit gewissen Eigenschaften (stetig, differenzierbar, stetig differenzierbar, siehe 2.2) gibt.

Außerdem beachte man, dass Definition 2.1.4.1 sehr allgemein gehalten ist. So sind beispielsweise konstante Funktionen zugelassen, wodurch auch einzelne Punkte als Bilder von Bereichen beliebiger Dimension k und somit als k -dimensionale Flächen auftreten. Das entspricht sicher nicht der Intention, weil die Dimension der Bildmenge 0, also kleiner ist als die des Definitionsbereichs. Doch auch das umgekehrte Phänomen tritt auf: Es gibt sogenannte **raumfüllende Kurven** wie sie erstmals von Peano konstruiert wurden. Das sind stetige Funktionen $f : [0, 1] \rightarrow \mathbb{R}^n$, die also auf dem 1-dimensionalen Einheitsintervall definiert sind, deren Bild aber zum Beispiel das volle Einheitsquadrat oder auch der volle k -dimensionale Einheitswürfel ist. So ein f kann allerdings, (wie sich mit Hilfe der Bogenlänge zeigen kann) zum Beispiel nicht stetig differenzierbar sein. Will man sich auf Situationen beschränken, die der Intuition besser entsprechen, ist es sinnvoll, zusätzliche Voraussetzungen an f und/oder D zu machen. Ist f stetig differenzierbar, dann kann die Dimension von $f(D)$ (hier in einem anschaulichen, nicht exakt definierten Sinne zu verstehen) nicht größer sein als die von D . Ist f' (siehe spätere Abschnitte) nur an wenigen Punkten singulär und auch D regulär (zum Beispiel im Sinn von $\bar{D}^\circ = D$, d.h. D ist der Abschluss seines Inneren; außerdem natürlich $D \neq \emptyset$), so kann die Dimension von $f(D)$ umgekehrt auch nicht kleiner werden als die von D .

Übungsaufgabe 121. (P) Parametrisieren Sie folgende ebene Bereiche:

1. Viereck im \mathbb{R}^2 mit den Eckpunkten $(0, 0)$, $(1, 1)$, $(3, 2)$ und $(4, -3)$. Hinweis: Zerlegen in Dreiecke.
2. Jene Punkte im \mathbb{R}^2 , welche innerhalb des Kreises mit Mittelpunkt $(0, \frac{1}{2})$ und Radius 1 liegen und eine y -Koordinate ≥ 0 besitzen.

Übungsaufgabe 122. (P) Parametrisieren Sie das Tetraeder mit den Eckpunkten $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$.

Übungsaufgabe 123. (P) Parametrisieren und skizzieren Sie folgende Flächen im \mathbb{R}^3

1. Kreis in der y, z Ebene mit Mittelpunkt $(0, 1, 2)$ und Radius $r = 3$,
2. Dreieck mit den Eckpunkten $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$,
3. Kreis in der Ebene $E : x + y + z = 1$ mit Mittelpunkt $(1, 0, 0)$ und Radius $r = 1$,
4. Kegelmantel des (Doppel)-Kegels um die z -Achse mit Spitze im Ursprung und Öffnungswinkel 45°

5. Kegelmantel des (Doppel)-Kegels um die Gerade $g : (1, 0, 0) + t(1/1/1)$ mit Spitze im Punkt $(1, 0, 0)$ und Öffnungswinkel 45°

Übungsaufgabe 124. (E) Stellen Sie weitere interessante, nichtlineare geometrische Objekte Ihrer Wahl (zum Beispiel Ellipsen, Hyperbeln, Ellipsoide, Paraboloiden etc.) sowohl durch Gleichungen bzw. Gleichungssysteme als auch in Parameterform als Flächen dar.

2.1.5 Koordinatentransformationen

Inhalt in Kurzfassung: Die aus Mathematik 1 bereits bekannten Polarkoordinaten in der Ebene sowie die Kugelkoordinaten im Raum werden als Beispiele für eine Funktion vom n -Dimensionalen ins n -Dimensionale für $n = 2$ bzw. $n = 3$ rekapituliert.

Für Funktionen $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ wurde der einfachste Fall, nämlich $m = n = 1$, in Mathematik 1 behandelt, der Fall $m = 1$ und $n > 1$ in 2.1.3, $n = 1, 2$ und $m > n$ in 2.1.4. Als weiteren Spezialfall von besonderem Interesse betrachten wir nun die Situation $m = n > 1$, insbesondere für $m = n = 2, 3$. Zwei sehr unterschiedliche geometrische Interpretationen sind von besonderer Bedeutung. In 2.1.6 werden wir uns mit der zweiten, nämlich mit Vektorfeldern beschäftigen. Zunächst aber behandeln wir Koordinatentransformationen, und zwar anhand der bereits aus Mathematik 1 bekannten Beispiele Polar- (für $n = 2$) und Kugelkoordinaten (für $n = 3$):

$n = 2$, Polarkoordinaten: Ein Punkt $\mathbf{x} = (x, y) \in \mathbb{R}^2$ der Ebene mit den **kartesischen Koordinaten** x und y lässt sich in der Form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos(\alpha) \\ r \sin(\alpha) \end{pmatrix}$$

darstellen mit den sogenannten **Polarkoordinaten** r und α . Wegen

$$x^2 + y^2 = r^2 \cos^2(\alpha) + r^2 \sin^2(\alpha) = r^2(\cos^2(\alpha) + \sin^2(\alpha)) = r^2$$

ist dabei r^2 und somit $|r|$ durch \mathbf{x} eindeutig bestimmt. Auf r selbst überträgt sich die Eindeutigkeit, wenn man $r \geq 0$ verlangt. Auch α wird unter einer geeigneten Einschränkung, am einfachsten $\alpha \in [0, 2\pi)$, eindeutig. Geometrische Interpretation: Der Radius oder Betrag r gibt den Abstand des Punktes \mathbf{x} vom Koordinatenursprung \mathbf{o} an, der Winkel α die Richtung, in die man von \mathbf{o} geradlinig nach \mathbf{x} kommt. Für unsere weiteren Zwecke liegt es nahe, den Zusammenhang zwischen kartesischen und Polarkoordinaten als Funktion

$$f : \mathbf{a} = \begin{pmatrix} r \\ \alpha \end{pmatrix} \mapsto \begin{pmatrix} x(r, \alpha) \\ y(r, \alpha) \end{pmatrix} := \begin{pmatrix} r \cos(\alpha) \\ r \sin(\alpha) \end{pmatrix}$$

aufzufassen. Definiert ist f auf ganz \mathbb{R}^2 . Besonders interessant ist der kleinere, d.h. nicht ganz \mathbb{R}^2 umfassende Definitionsbereich $D = [0, \infty) \times [0, 2\pi] \subseteq \mathbb{R}^2$. Dort ist f nämlich immer noch surjektiv, wenn auch nicht injektiv. Immerhin ist Injektivität nur auf einer

kleinen Menge verletzt: Für $(r_1, \alpha_1) \neq (r_2, \alpha_2)$ ist $f(r_1, \alpha_1) = f(r_2, \alpha_2)$ nur dann möglich, wenn $r_1 = r_2 = 0$ oder $\{\alpha_1, \alpha_2\} \subseteq \{0, 2\pi\}$. Umgekehrt ist die Situation auf der in D dichten Teilmenge $D_0 := (0, \infty) \times [0, 2\pi)$. Da ist f tatsächlich injektiv, dafür nur *fast* surjektiv, weil im Bildbereich der Koordinatenursprung fehlt.

$n = 3$, Kugelkoordinaten: Im dreidimensionalen Fall tritt zu r und α noch ein dritter Parameter β . Die Funktion

$$f : \mathbf{a} = \begin{pmatrix} r \\ \alpha \\ \beta \end{pmatrix} \mapsto \begin{pmatrix} x(r, \alpha, \beta) \\ y(r, \alpha, \beta) \\ z(r, \alpha, \beta) \end{pmatrix} := \begin{pmatrix} r \cos(\alpha) \cos(\beta) \\ r \sin(\alpha) \cos(\beta) \\ r \sin(\beta) \end{pmatrix}$$

ist auf ganz \mathbb{R}^3 definiert, auf der abgeschlossenen Menge $D := [0, \infty) \times [0, 2\pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ surjektiv nach \mathbb{R}^3 und fast injektiv, auf der in D dichten Menge $D_0 := (0, \infty) \times [0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2})$ im strengen Sinn injektiv, dafür nur fast surjektiv (wieder fehlt der Nullvektor im Bildbereich). Geometrische Interpretation: Wieder ist r der Abstand vom Koordinatenursprung, α entspricht, auf den Globus übertragen, der geographischen Länge, β der geographischen Breite.

In beiden Fällen, Polar- wie auch Kugelkoordinaten, liegt eine surjektive Funktion $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ vor, die auf D_0 injektiv ist. Das wird vor allem in der Integralrechnung (Kapitel 4) wichtig sein, weil der Bereich $D \setminus D_0$, wo Injektivität verletzt ist, jeweils eine Nullmenge ist (n -dimensionales Maß 0 hat) und deshalb bei der Integration nicht stört. Auf dem Weg zur Integralrechnung werden wir vorher noch an geeigneter Stelle den Zusammenhang mit der Funktionaldeterminante (der Determinante der Ableitung von f) aufgreifen.

2.1.6 Vektorfelder

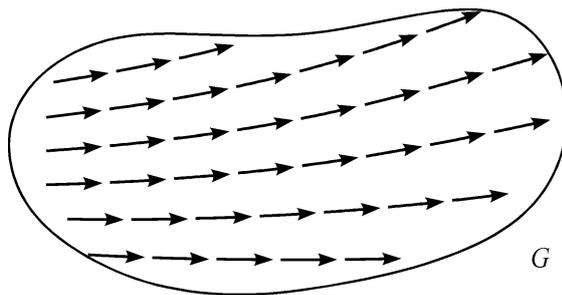
Inhalt in Kurzfassung: Ein Vektorfeld ordnet einem n -dimensionalen Vektor, den man meist als Punkt im Raum auffasst, einen anderen n -dimensionalen Vektor zu, der zum Beispiel eine Kraft oder eine Strömung, die an diesem Punkt wirkt, repräsentieren kann.

So wie Koordinatentransformationen sind formal auch Vektorfelder Abbildungen vom n -Dimensionalen ins n -Dimensionale. Allerdings ist ihre Interpretation eine völlig andere als bei Koordinatentransformationen, was auch ganz andere daran anschließende Begriffsbildungen und Untersuchungen nach sich ziehen wird. Meist notieren wir Vektorfelder als $\mathbf{v} : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ (vorzugsweise für $n = 2, 3$).

Im Vordergrund steht die Vorstellung, dass ein Vektorfeld jedem Punkt $\mathbf{x} \in D \subseteq \mathbb{R}^n$ wieder einen Vektor $\mathbf{v}(\mathbf{x})$ im \mathbb{R}^n zuordnet. Oft wird vorausgesetzt, dass D eine offene und **zusammenhängende** Teilmenge des \mathbb{R}^n ist. Eine solche Menge D nennt man auch ein **Gebiet** und bezeichnet sie deshalb gerne mit dem Buchstaben G .²

²Wir erinnern an die Definition aus Mathematik 1: Eine Teilmenge O eines metrischen Raumes X heißt offen, wenn es zu jedem $x \in O$ einen positiven ε gibt derart, dass die Kugel um x vom Radius ε ganz in O enthalten ist.

Wann ein $O \subseteq X$ zusammenhängend ist, lässt sich exakt über das Gegenteil definieren: O heißt



Eine häufige Interpretation von $\mathbf{v}(\mathbf{x})$ ist die als im Punkt \mathbf{x} ansetzender Kraftvektor innerhalb eines Kraftfeldes \mathbf{v} , das zum Beispiel ein Gravitations-, ein elektrisches oder magnetisches Feld sein kann. In solchen Situationen ist im Definitionsbereich von \mathbf{v} die Quelle des Feldes (Massepunkt, elektrische Ladung) ausgenommen. Andere wichtige Anwendungen haben Vektorfelder in der Strömungslehre. Dort repräsentiert $\mathbf{v}(\mathbf{x})$ zum Beispiel die Bewegung einer Flüssigkeit oder eines Gases im Punkt \mathbf{x} . Auch Wärmeströmungen lassen sich mit Vektorfeldern modellieren. Man spricht von **stationären** Feldern, wenn keine zeitliche Veränderung in die Modellierung eingebaut ist. Andernfalls hinge $\mathbf{v} : (\mathbf{x}, t) \mapsto \mathbf{v}(\mathbf{x}, t)$ auch von der Zeit t ab. Hier konzentrieren wir uns auf stationäre Vektorfelder.

Für dreidimensionale Vektorfelder $\mathbf{v} : D \rightarrow \mathbb{R}^3$ mit $D \subseteq \mathbb{R}^3$ verwenden wir meist die Schreibweise

$$\mathbf{v} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} u(x, y, z) \\ v(x, y, z) \\ w(x, y, z) \end{pmatrix}$$

mit Komponentenfunktionen $u, v, w : D \rightarrow \mathbb{R}$.

2.2 Die Ableitung im Höherdimensionalen

In diesem Abschnitt geht es um den wichtigsten Begriff dieses Kapitels, nämlich um die Ableitung $f'_{\mathbf{x}_0}$ einer Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ an einer Stelle $\mathbf{x}_0 \in D$. Eine erste, stark heuristische Annäherung erfolgt in 2.2.1, wo vor allem die Richtungs- und partiellen Ableitungen einer reellwertigen Funktion definiert werden. Durch Abstraktion und Übertragung von Bekanntem aus Mathematik 1 gelangen wir in 2.2.2 zur allgemeinen Definition von $f'_{\mathbf{x}_0}$ als linearer Approximation von f in der Nähe von \mathbf{x}_0 . Die Darstellung dieser linearen Abbildung $f'_{\mathbf{x}_0}$ als Matrix mit den partiellen Ableitungen von f als Eintragungen ist Hauptgegenstand von 2.2.3. Wichtige, bereits bekannte Sachverhalte und Rechenregeln aus der eindimensionalen Differentialrechnung haben sehr natürliche Verallgemeinerungen ins Mehrdimensionale, welche den Inhalt von 2.2.4 bilden. Weil eine lebendige Anschauung für die – je nach Dimension unterschiedlichen – geometrischen Interpretationen der Ableitung für das weitere Verständnis entscheidend ist, werden in

nicht zusammenhängend, wenn es die Vereinigung $O = A \cup B$ zweier nichtleerer disjunkter Mengen A und B darstellen lässt, von denen keine einen Randpunkt der anderen enthält. Wenn A und B positiven Abstand haben, ist das beispielsweise sicher der Fall.

2.2.5 verschiedene Fälle von m und n unter diesem Gesichtspunkt durchgespielt. Einer dieser Fälle, nämlich Gradientenfelder (auch Potentialfelder; hier ist $m = 1$, wobei $n = 2, 3$ die interessantesten Fälle sind), wird in 2.2.6 wesentlich vertieft.

2.2.1 Richtungs- und partielle Ableitungen

Inhalt in Kurzfassung: Anhand von Funktionsgebirgen werden die Begriffe der Richtungsableitung und der partiellen Ableitungen nach den verschiedenen Variablen motiviert und definiert. Durch Iteration des Prozesses des partiellen Differenzierens erhält man die höheren partiellen Ableitungen.

Wir wollen uns der Frage nach der Differentiation im Höherdimensionalen zunächst am Beispiel reellwertiger Funktionen

$$f : D \rightarrow \mathbb{R}, \quad f : \mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x}) = f(x_1, \dots, x_n), \quad \text{mit } D \subseteq \mathbb{R}^n$$

annähern. Oft denken wir zwecks leichter Veranschaulichung mittels Funktionsgebirgen (siehe 2.1.3) besonders an den Fall $n = 2$ und schreiben für die Variablen dann gerne auch x und y statt x_1 und x_2 , also

$$f : D \rightarrow \mathbb{R}, \quad f : \mathbf{x} = (x, y) \mapsto f(\mathbf{x}) = f(x, y), \quad \text{mit } D \subseteq \mathbb{R}^2.$$

Im Eindimensionalen gibt $f'(x_0)$ den Anstieg einer reellen Funktion f in einem Punkt x_0 entlang der x -Achse an. Bei (beispielsweise) 2-dimensionalem Definitionsbereich gibt es aber nicht nur die eine Richtung entlang der x -Achse, sondern auch die entlang der y -Achse sowie unendlich viele Richtungen dazwischen. Jede dieser Richtungen lässt sich durch irgendeinen Vektor $\neq \mathbf{o}$ repräsentieren. Zweckmäßigerweise wählen wir einen Einheitsvektor, d.h. einen (nicht notwendig achsenparallelen) Vektor \mathbf{e} der Länge $\|\mathbf{e}\| = 1$. Völlig analoge Überlegungen wie in Mathematik 1 für die gewöhnliche Ableitung einer differenzierbaren reellen Funktion führen uns zum verallgemeinerten Differentialquotienten

$$\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}_0) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{e}) - f(\mathbf{x}_0)}{t}$$

als Anstieg von f im Punkt $\mathbf{x}_0 \in D$ in Richtung \mathbf{e} . (Wollten wir die Richtung \mathbf{e} von der entgegengesetzten Richtung $-\mathbf{e}$ unterscheiden, könnten wir im Grenzwert $t \rightarrow 0$ der obigen Definition zusätzlich $t > 0$ voraussetzen.) Wegen $\|\mathbf{e}\| = 1$ lässt sich der Quotient deuten als Änderung des Funktionswertes dividiert durch die Änderung $t = \|t\mathbf{e}\| = \|\mathbf{x} - \mathbf{x}_0\|$ von $\mathbf{x} := \mathbf{x}_0 + t\mathbf{e}$ in Richtung \mathbf{e} . Existiert dieser Grenzwert, so heißt $\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}_0)$ die **Richtungsableitung** von f entlang oder in Richtung von \mathbf{e} . Man beachte, dass sich die Richtungsableitung $\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}_0)$ auch als ganz gewöhnliche Ableitung im Sinne der eindimensionalen Differentialrechnung deuten lässt, nämlich als

$$\frac{\partial f}{\partial \mathbf{e}}(\mathbf{x}_0) = g'(0) \quad \text{mit der reellen Funktion } g(t) := f(\mathbf{x}_0 + t\mathbf{e}).$$

2 Differentialrechnung in mehreren Variablen

Ein besondere Rolle spielen die Richtungsableitungen

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) := \frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0)$$

entlang der kanonischen Einheitsvektoren $\mathbf{e}_i \in \mathbb{R}^n$, $i = 1, \dots, n$, die sogenannten ***i*-ten partiellen Ableitungen** von f . Gebräuchlich sind auch die Schreibweisen f_{x_i} , $\frac{df}{dx_i}$ und $\frac{\partial}{\partial x_i} f$ für $\frac{\partial f}{\partial x_i}$. Den aus den partiellen Ableitungen aufgebauten Vektor nennt man auch den **Gradienten** $\text{grad}_{\mathbf{x}_0} f$ von f in \mathbf{x}_0 , genauer:

$$\text{grad}_{\mathbf{x}_0} f := \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \end{pmatrix} \in \mathbb{R}^n.$$

Je nach Kontext kann $\text{grad}_{\mathbf{x}_0} f$ als Zeilen- oder auch Spaltenvektor wie hier aufgefasst werden. Er lässt sich, wie wir an späterer Stelle (siehe 2.2.3.3) begründen werden, als Richtung des größten Anstiegs von f an der Stelle \mathbf{x}_0 interpretieren. Existiert der Gradient in jedem Punkt des Definitionsbereichs, so liegt eine Abbildung $\text{grad} f : D \rightarrow \mathbb{R}^n, \mathbf{x} \mapsto \text{grad}_{\mathbf{x}} f$ vor, die wir, wie wir später sehen werden, auch mit der Ableitung von f identifizieren können. Es wird sich nämlich zeigen, dass die partiellen Ableitungen die wesentliche Information über das, was wir gleich als Ableitung von f definieren werden, enthalten.

Warum das so ist, lässt sich sehr leicht anhand von Funktionsgebirgen illustrieren. So, wie bei einer Funktion in einer Variablen die Tangente an den Funktionsgraphen die Ableitung repräsentiert, tut das bei zwei Variablen die Tangentialebene $\tau = \tau_{f, \mathbf{x}_0}$ ans Funktionsgebirge. Eine Ebene ist aber bereits durch zwei Richtungsableitungen eindeutig bestimmt, und solche ergeben sich aus den partiellen Ableitungen. Für $n = 2$ und $\mathbf{x}_0 = (x_0, y_0)$ ergibt sich mit ihrer Hilfe nämlich folgende Parameterdarstellung:

$$\tau = \tau_{f, \mathbf{x}_0} = \left\{ \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ f(\mathbf{x}_0) \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(\mathbf{x}_0) \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(\mathbf{x}_0) \end{pmatrix}, \quad s, t \in \mathbb{R} \right\}$$

Durchlaufen die Parameter s und t alle reellen Zahlen, so durchläuft \mathbf{x} alle Punkte der Tangentialebene an den Graphen von f an der Stelle \mathbf{x}_0 , d.h. im Punkt $(x_0, y_0, f(x_0, y_0))$. Ganz Analoges gilt statt für zwei auch für $n > 2$ Variablen, wenn man sämtliche n partiellen Ableitungen von f an der Stelle $\mathbf{x}_0 \in \mathbb{R}^n$ verwendet.

Bevor wir zum allgemeinen Begriff der Ableitung kommen, definieren wir für den späteren Gebrauch noch die **höheren partiellen Ableitungen**. Für eine reellwertige Funktion $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ ist die i -te partielle Ableitung (sofern sie existiert) wieder eine reellwertige Funktion

$$\frac{\partial f}{\partial x_i} : D \rightarrow \mathbb{R}.$$

Hat diese Funktion selbst wieder partielle Ableitungen, kann der Prozess der partiellen Differentiation iteriert werden. Durch n -maliges partielles Differenzieren nach einer oder verschiedenen der Variablen erhält man die **höheren partiellen Ableitungen der Ordnung n** . Für $n = 2$, $x_1 = x$ und $x_2 = y$ beispielsweise schreibt man

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2}{\partial x^2} f, \quad f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2}{\partial y \partial x} f, \dots$$

Die allgemeine Definition ergibt sich rekursiv nach n :

$$\frac{\partial^{n+1} f}{\partial x_{i_{n+1}} \partial x_{i_n} \dots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_{n+1}}} \left(\frac{\partial^n f}{\partial x_{i_n} \dots \partial x_{i_1}} \right)$$

2.2.2 Definition der Ableitung als lineare Approximation

Inhalt in Kurzfassung: Die Definition der Ableitung einer (eindimensionalen) reellen Funktion als Differentialquotient lässt sich nicht unmittelbar auf höhere Dimensionen übertragen. Aus Mathematik 1 kennen wir aber auch die Deutung als (hinreichend gute) lineare Approximation, die sich sehr wohl auf den allgemeinen Fall übertragen lässt. Damit ergeben sich in natürlicher Weise Begriffe wie Differenzierbarkeit etc.

Wir greifen die Parameterdarstellung aus 2.2.1 nochmals auf und deuten sie als (zunächst inhomogene lineare, d.h. eigentlich **affine**) Funktion

$$(s, t) \mapsto \begin{pmatrix} x_0 \\ y_0 \\ f(\mathbf{x}_0) \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(\mathbf{x}_0) \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(\mathbf{x}_0) \end{pmatrix}.$$

Weil in den ersten beiden Komponenten f gar nicht vorkommt, ist die dritte Komponente die interessante. Wir schreiben sie an als Funktion

$$z(s, t) := f(\mathbf{x}_0) + l(s, t) \quad \text{mit} \quad l(s, t) := s \frac{\partial f}{\partial x}(\mathbf{x}_0) + t \frac{\partial f}{\partial y}(\mathbf{x}_0).$$

Darin ist die Funktion l eine (homogene) lineare Funktion, die wir, sofern sie f hinreichend gut approximiert, als **Ableitung** von f an der Stelle \mathbf{x}_0 ansehen werden. Was in diesem Zusammenhang *hinreichend gut* bedeutet, ergibt sich in Analogie zur Charakterisierung der Ableitung im Eindimensionalen (siehe Mathematik 1). Dazu eine kurze Veranschaulichung des Problems.

Beispiel: Man stelle sich etwa ein Gebäude mit kreuzförmigen Grundriss und mit einem Satteldach vor, dessen Langhäuser parallel zu den Koordinatenachsen verlaufen mit dem Ursprung $\mathbf{o} = (0, 0)$ eines zweidimensionalen Koordinatensystems in der Grundebene direkt unter dem Schnittpunkt der beiden Teile des Dachfirstes. Die Funktion $f(x, y)$ beschreibe die Höhe des Daches über dem Punkt (x, y) . Weil der Dachfirst sowohl in x - als auch in y -Richtung waagrecht verläuft, sind beide partiellen Ableitungen von f

2 Differentialrechnung in mehreren Variablen

in \mathbf{o} gleich 0. Die obige Parameterdarstellung für die Tangentialebene ergibt jene Ebene, in der der (rechtwinkelige) Dachfirst liegt. Wir sind aber nicht geneigt, diese Ebene als Tangentialebene des Daches am First (genauer: im Punkt $(0, 0, f(0, 0))$) zu akzeptieren, und zwar aus ganz analogem Grund, aus dem wir im Eindimensionalen die x -Achse nicht als Tangente an den Funktionsgraphen von $f(x) := |x|$ im Punkt $x_0 = 0$ akzeptieren. Wir verlangen nämlich sowohl von einer Tangente als auch von einer Tangentialebene, dass sie eine bessere Approximation der Funktion f ist als im Beispiel mit $|x|$ bzw. im Beispiel mit dem Dachfirst. Wir lernen, dass an gewissen Stellen auch dann partielle Ableitungen existieren können, wenn die Funktion insgesamt an dieser Stelle nicht differenzierbar ist.

Die Antwort auf die Frage, was nun die angemessene Approximationsgüte ist, die man von der Ableitung erwarten darf, ergibt sich durch unmittelbare Verallgemeinerung der Situation im Eindimensionalen. Dazu erinnern wir uns zunächst an die Definition der Ableitung $f'(x_0)$ einer reellen Funktion f an der Stelle x_0 als Differentialquotient

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Diese Definition wäre in höheren Dimensionen nicht sinnvoll, weil eine Division nur dann einen Sinn hat, wenn im Nenner Zahlen stehen, nicht aber Vektoren. Es gibt aber einen Ausweg: In Mathematik 1 haben wir gesehen, dass die Existenz dieses Grenzwertes äquivalent dazu ist, dass es eine lineare Funktion l gibt, nämlich $l : x \mapsto f'(x_0)x$, so dass das Fehlerglied $r(x)$ in der Darstellung

$$f(x) = f(x_0) + l(x - x_0) + r(x)$$

von kleinerer Größenordnung ist als $x - x_0$. Diese Beziehung lässt sich problemlos auf höhere Dimensionen übertragen. Sogar auf Funktionen, deren Werte nicht in \mathbb{R} sondern in einem höherdimensionalen Standardvektorraum \mathbb{R}^m liegen. Die Ableitung ist dann allerdings nicht mehr eine Zahl wie im Eindimensionalen der Differentialquotient, sondern eine lineare Abbildung. Die allgemeine Definition lautet also:

Definition 2.2.2.1. Sei $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ und \mathbf{x}_0 ein innerer Punkt von D . Eine lineare Abbildung $l : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt **Ableitung** oder **Differential von f an der Stelle \mathbf{x}_0** , symbolisch $l = f'_{\mathbf{x}_0}$, wenn gilt:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + l(\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x}) = f(\mathbf{x}_0) + f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x})$$

mit einem Restterm (Fehlerterm) $r(\mathbf{x})$, der für $\mathbf{x} \rightarrow \mathbf{x}_0$ sehr schnell gegen $\mathbf{o} \in \mathbb{R}^m$ konvergiert, genauer:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|r(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

In diesem Fall heißt f im Punkt \mathbf{x}_0 **differenzierbar**. Ist f in allen Punkten $\mathbf{x} \in D$ differenzierbar (in diesem Fall muss D offen sein), so heißt f (schlechthin) differenzierbar auf D . In diesem Fall heißt die Funktion $f' : D \rightarrow L(\mathbb{R}^n, \mathbb{R}^m)$, $f' : \mathbf{x} \mapsto f'_{\mathbf{x}}$ die **Ableitung von f** .

In der letzten Beziehung kann im Zähler des Bruches statt $\|r(\mathbf{x})\|$ auch einfacher $r(\mathbf{x})$ geschrieben werden, wenn gleichzeitig auf der rechten Seite die Zahl 0 durch den Nullvektor \mathbf{o} ersetzt wird.

Die Formulierung „Eine lineare Abbildung ...“ in Definition 2.2.2.1 scheint die Möglichkeit offen zu lassen, dass es mehrere solche Abbildungen l_1 und l_2 geben könnte. In Wahrheit ist das aber nicht der Fall, wie aus der folgenden Eindeutigkeitsaussage hervorgeht, durch die auch die Schreibweise $f'_{\mathbf{x}_0}$ ihre Rechtfertigung erfährt.

Proposition 2.2.2.2. *Die Ableitung $f'_{\mathbf{x}_0}$ einer Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ in einem inneren Punkt \mathbf{x}_0 von D ist eindeutig bestimmt.*

Beweis. Nimmt man

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + l_1(\mathbf{x} - \mathbf{x}_0) + r_1(\mathbf{x}) \quad \text{und} \\ f(\mathbf{x}) &= f(\mathbf{x}_0) + l_2(\mathbf{x} - \mathbf{x}_0) + r_2(\mathbf{x}) \end{aligned}$$

mit Restgliedern r_1 und r_2 entsprechend Definition 2.2.2.1 an, so liefert Differenzbildung $(l_1 - l_2)(\mathbf{x} - \mathbf{x}_0) = r_2(\mathbf{x}) - r_1(\mathbf{x})$. Wir halten ein beliebiges $\mathbf{x}_1 \neq \mathbf{x}_0$ aus \mathbb{R}^n fest und setzen $\mathbf{x}_\varepsilon := \mathbf{x}_0 + \varepsilon(\mathbf{x}_1 - \mathbf{x}_0)$. Dann gilt

$$\begin{aligned} \varepsilon(l_1 - l_2)(\mathbf{x}_1 - \mathbf{x}_0) &= (l_1 - l_2)(\mathbf{x}_\varepsilon - \mathbf{x}_0) = r_2(\mathbf{x}_\varepsilon) - r_1(\mathbf{x}_\varepsilon) = \frac{r_2(\mathbf{x}_\varepsilon) - r_1(\mathbf{x}_\varepsilon)}{\|\mathbf{x}_\varepsilon - \mathbf{x}\|} \|\mathbf{x}_\varepsilon - \mathbf{x}\| = \\ &= \frac{r_2(\mathbf{x}_\varepsilon) - r_1(\mathbf{x}_\varepsilon)}{\|\mathbf{x}_\varepsilon - \mathbf{x}\|} \varepsilon \|\mathbf{x}_1 - \mathbf{x}_0\|, \end{aligned}$$

nach kürzen von ε in der Gleichung also

$$(l_1 - l_2)(\mathbf{x}_1 - \mathbf{x}_0) = \frac{r_2(\mathbf{x}_\varepsilon) - r_1(\mathbf{x}_\varepsilon)}{\|\mathbf{x}_\varepsilon - \mathbf{x}\|} \|\mathbf{x}_1 - \mathbf{x}_0\|.$$

Im Grenzübergang $\varepsilon \rightarrow 0$ ist rechts der zweite Faktor konstant, während der erste gegen \mathbf{o} konvergiert, also ist $(l_1 - l_2)(\mathbf{x}_1 - \mathbf{x}_0) = \mathbf{o}$. Weil $\mathbf{x}_1 \neq \mathbf{x}_0$ ein beliebiger Vektor aus \mathbb{R}^n war, folgt $l_1 = l_2$. □

2.2.3 Funktionalmatrix und -determinante

Inhalt in Kurzfassung: Weil die Ableitung $f'_{\mathbf{x}_0}$ einer Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ mit Komponentenfunktionen f_1, \dots, f_m an der Stelle \mathbf{x}_0 eine lineare Funktion $f'_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist, hat sie (bezüglich der kanonischen Basen) eine Darstellung als $m \times n$ -Matrix $A = (a_{ij})$. Die Eintragungen a_{ij} von A erweisen sich gerade als die partiellen Ableitungen: $a_{ij} = \frac{\partial f_i}{\partial x_j}$. Im Falle $m = 1$ (wenn also f reellwertig ist) folgt daraus $f'_{\mathbf{x}_0} = \text{grad}_{\mathbf{x}_0} f$. Mit Hilfe der Cauchy-Schwarzschen Ungleichung erkennt man damit, dass, wie schon an früherer Stelle angekündigt, der Gradient in die Richtung des größten Anstiegs von f weist.

Die Bestimmung der Ableitung einer Funktion $f : \mathbb{R}^n \supset D \rightarrow \mathbb{R}^m$ in einem Punkt $\mathbf{x}_0 \in D$ als lineare Abbildung läuft auf die Frage nach der Matrixdarstellung der linearen

2 Differentialrechnung in mehreren Variablen

Abbildung $f'_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ (bezüglich der kanonischen Basen von \mathbb{R}^n und \mathbb{R}^m) hinaus. Zunächst überlegt man sich sehr leicht, dass sich die Ableitung f' von f in natürlicher Weise aus den Ableitungen der Komponentenfunktionen f_1, \dots, f_m von f zusammensetzt:

Proposition 2.2.3.1. *Eine Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ und*

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}, \quad f_i : D \rightarrow \mathbb{R} \quad \text{für } i = 1, \dots, m,$$

ist im (inneren) Punkt \mathbf{x}_0 von D genau dann differenzierbar, wenn alle f_i in \mathbf{x}_0 differenzierbar sind. In diesem Fall gilt

$$f'_{\mathbf{x}_0} = \begin{pmatrix} (f'_1)_{\mathbf{x}_0} \\ \vdots \\ (f'_m)_{\mathbf{x}_0} \end{pmatrix}.$$

Die Zeilenvektoren in der Darstellung von $f'_{\mathbf{x}_0}$ als $m \times n$ -Matrix sind folglich die Darstellungen der Ableitungen $(f'_i)_{\mathbf{x}_0}$ der Komponentenfunktionen f_i an der Stelle \mathbf{x}_0 .

Der Beweis ist im Wesentlichen eine Anwendung von Proposition 2.1.2.4 (Übung).

Übungsaufgabe 125. (E) *Beweisen Sie Proposition 2.2.3.1.*

Wir konzentrieren daher unsere Überlegungen auf den Fall $m = 1$, d.h. einer reellwertigen Funktion

$$f : D \rightarrow \mathbb{R}, \quad f : \mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x}) = f(x_1, \dots, x_n), \quad \text{mit } D \subseteq \mathbb{R}^n.$$

Für diesen Fall haben wir die wichtigste Vorarbeit schon in 2.2.1 geleistet. Kombiniert man nämlich die dortigen Überlegungen mit der Beziehung zwischen Differentialquotient und linearer Approximation im Eindimensionalen, so wird klar, dass die lineare Approximation von f an der Stelle \mathbf{x}_0 entlang der kanonischen Einheitsvektoren \mathbf{e}_i gerade mit dem Faktor $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$ zu versehen ist. Für die lineare Abbildung $f'_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}$ gilt also $f'_{\mathbf{x}_0} : \mathbf{e}_i \mapsto \frac{\partial f}{\partial x_i}(\mathbf{x}_0)$. Da wir somit die Werte von $f'_{\mathbf{x}_0}$ auf der kanonischen Basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ kennen, ergibt sich für $f'_{\mathbf{x}_0}$ (bezüglich der kanonischen Basis) daraus die Darstellung

$$f'_{\mathbf{x}_0} = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_i}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right) = \text{grad}_{\mathbf{x}_0} f$$

als $1 \times n$ -Matrix. Sie stimmt mit dem Gradienten von f an der Stelle \mathbf{x}_0 , aufgefasst als Zeilenvektor, überein. Aus der Differenzierbarkeit von f in \mathbf{x}_0 folgt also insbesondere die Existenz der partiellen Ableitungen von f in \mathbf{x}_0 . Dass die Umkehrung nicht allgemein gelten kann, haben wir uns anhand des Beispiels mit dem Dachfirst in 2.2.2 klar gemacht. Sie gilt aber, wenn man von den partiellen Ableitungen nicht nur Existenz, sondern sogar Stetigkeit in einer Umgebung von \mathbf{x}_0 voraussetzt. Dank Proposition 2.2.3.1 gilt sogar:

Satz 2.2.3.2. Gegeben seien eine Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ und

$$f(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix}, \quad f_i : D \rightarrow \mathbb{R} \quad \text{für } i = 1, \dots, m,$$

und ein innerer Punkt \mathbf{x}_0 von D . Dann gilt:

1. Ist f differenzierbar in \mathbf{x}_0 , so existieren alle partiellen Ableitungen

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0), \quad i = 1, \dots, m, \quad j = 1, \dots, n,$$

von f im Punkt \mathbf{x}_0 . Ist $A = (a_{ij})$ jene $m \times n$ -Matrix, die $f'_{\mathbf{x}_0}$ bezüglich der kanonischen Basis darstellt, so gilt

$$a_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0).$$

2. Umgekehrt gilt: Gibt es eine Umgebung von \mathbf{x}_0 , in der die partiellen Ableitungen von f nicht nur existieren, sondern sogar stetig sind, dann ist f in \mathbf{x}_0 differenzierbar. (Insbesondere ist die erste Aussage des Satzes anwendbar, und die Matrix A für $f'_{\mathbf{x}_0}$ hat die dort angegebene Gestalt.)
3. Die Zeilenvektoren in der Darstellung von $f'_{\mathbf{x}_0}$ als $m \times n$ -Matrix sind die Darstellungen der Ableitungen $(f'_i)_{\mathbf{x}_0} = \text{grad}_{\mathbf{x}_0}(f_i)$ der Komponentenfunktionen f_i .

Beweis. (Skizze) Wegen Proposition 2.2.3.1 dürfen wir uns auf den Fall $m = 1$ konzentrieren. Dann ist die erste Behauptung des Satzes gerade die Erkenntnis aus den vorangegangenen Überlegungen.

Die zweite Behauptung lässt sich mit Hilfe folgender Idee beweisen (Skizze): Weil \mathbf{x}_0 ein innerer Punkt des Definitionsbereichs D von f ist und weil die partiellen Ableitungen stetig sind, gibt es einen in D enthaltenen achsenparallelen Quader Q um \mathbf{x}_0 , innerhalb dessen sich die partiellen Ableitungen von f sehr wenig ändern, also fast mit ihren Werten in \mathbf{x}_0 übereinstimmen. Sei nun $\mathbf{x} \in Q$ beliebig. Wir suchen jene Punkte $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n = \mathbf{x} \in Q$ auf, so dass sich für alle $j = 1, \dots, n$ die Punkte \mathbf{x}_{j-1} und \mathbf{x}_j höchstens in der j -ten Komponente unterscheiden. Nach dem Mittelwertsatz der eindimensionalen Differentialrechnung ist

$$f(\mathbf{x}_j) \approx f(\mathbf{x}_{j-1}) + \|\mathbf{x}_j - \mathbf{x}_{j-1}\| \frac{\partial f}{\partial x_j}(\mathbf{x}_0).$$

Diese Approximationen lassen sich zu

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \text{grad}_{\mathbf{x}_0} f(\mathbf{x} - \mathbf{x}_0)$$

zusammenfügen, wobei die dem Zeilenvektor $\text{grad}_{\mathbf{x}_0} f$ entsprechende lineare Abbildung auf den Vektor $\mathbf{x} - \mathbf{x}_0$ anzuwenden ist. Wenn man sich auch noch die Mühe macht, die

2 Differentialrechnung in mehreren Variablen

Approximationsgüte aus den einzelnen Komponenten auf die n -fache Summe hochzurechnen, so bestätigt dies genau jene Beziehung aus Definition 2.2.2.1, die den Gradienten $\text{grad}_{\mathbf{x}_0} f$ als Matrixdarstellung von $f'_{\mathbf{x}_0}$ ausweist.

Diese Überlegung auf jede einzelne Zeile angewendet beweist auch die dritte Behauptung des Satzes. \square

Die Matrix $A = (a_{ij})$ in Satz 2.2.3.2 mit den Eintragungen $a_{ij} = \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$ heißt auch die **Funktionalmatrix** von f an der Stelle \mathbf{x}_0 . Ist $m = n$, so handelt es sich um eine quadratische Matrix, deren Determinante **Funktionaldeterminante** von f in \mathbf{x}_0 heißt. Für eine lineare Abbildung gibt die Determinante den multiplikativen Faktor an, mit dem Volumina verzerrt werden. Folglich lässt sich für $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ die Funktionaldeterminante $\det f'_{\mathbf{x}_0}$ (wenigstens für stetig differenzierbares f) als Faktor der Volumsvergrößerung durch f in der Nähe von \mathbf{x}_0 deuten.

Satz 2.2.3.2 erklärt auch die in 2.2.1 bereits angekündigte Deutung des Gradienten $\text{grad} f(\mathbf{x}_0)$ als steilsten Anstieg einer reellwertigen Funktion f in n Variablen im Punkt \mathbf{x}_0 . Um das zu sehen sei $\mathbf{e} = (e_1, \dots, e_n) \in \mathbb{R}^n$ ein beliebiger Vektor der Länge $\|\mathbf{e}\| = 1$. Ist f in \mathbf{x}_0 differenzierbar, so gilt

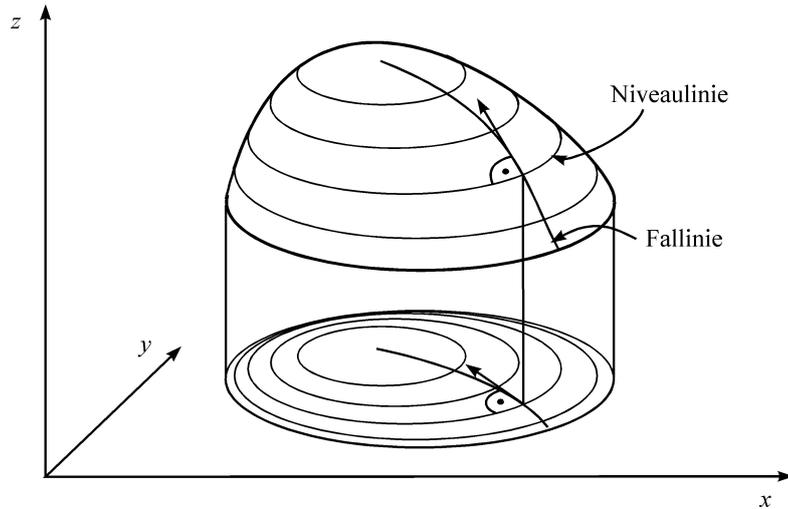
$$\frac{f(\mathbf{x}_0 + t\mathbf{e}) - f(\mathbf{x}_0)}{t} \approx \frac{1}{t} f'_{\mathbf{x}_0}(t\mathbf{e}) = \frac{1}{t} \text{grad}_{\mathbf{x}_0}(f)(t\mathbf{e}) = \text{grad}_{\mathbf{x}_0}(f)\mathbf{e},$$

Wobei die Annäherung \approx im Grenzwert $t \rightarrow 0$ zur Gleichheit wird. Der mittlere Anstieg von f ausgehend von \mathbf{x}_0 in Richtung \mathbf{e} ist also das Skalarprodukt des Gradienten mit dem Einheitsvektor \mathbf{e} . Das Skalarprodukt $\text{grad}_f(\mathbf{x}_0) \cdot \mathbf{e}$ lässt sich mit Hilfe der Cauchy-Schwarzschen Ungleichung

$$|\text{grad}_{\mathbf{x}_0}(f) \cdot \mathbf{e}| \leq \|\text{grad}_{\mathbf{x}_0}(f)\| \cdot \|\mathbf{e}\| = \|\text{grad}_{\mathbf{x}_0}(f)\|,$$

abschätzen, wobei Gleichheit genau dann gilt, wenn \mathbf{e} in dieselbe Richtung weist wie $\text{grad}_f(\mathbf{x}_0)$. Folglich gilt:

Proposition 2.2.3.3. *Ist die reellwertige Funktion $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$, im inneren Punkt \mathbf{x}_0 des Definitionsbereichs differenzierbar, so weist $\text{grad}_{\mathbf{x}_0}(f)$ in die Richtung der größten Richtungsableitung. Ihr Betrag stimmt mit der Norm $\|\text{grad}_{\mathbf{x}_0}(f)\|$ des Gradienten überein.*



Übungsaufgabe 126. (P) Berechnen Sie den Gradienten der Funktion f im Punkt P und geben Sie eine geometrische Interpretation:

1. $f(x, y) = 2x - 3y$, $P = (-1, 2)$
2. $f(x, y) = \sin \sqrt{x^2 + y^2}$, $P = (0, \pi)$
3. $f(x, y) = e^x(\cos y + \sin y)$, $P = (0, 0)$

Übungsaufgabe 127. (T) Gegeben ist die Funktion

$$f(x, y) = \begin{cases} \frac{2x^3y^2}{x^4+y^4} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0). \end{cases}$$

Zeigen Sie, indem Sie in die entsprechende Definition einsetzen, dass alle Richtungsableitungen von f im Punkt $(0, 0)$ existieren. Warum ist f trotzdem nicht im Sinne von Definition 2.2.2.1 an $(0, 0)$ differenzierbar? Hinweis: Berechnen Sie grad_f an der Stelle $(0, 0)$.

Übungsaufgabe 128. (T*) Gegeben ist die Funktion

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} + c, \quad \text{mit } c \in \mathbb{R}, \mathbf{x}, \mathbf{b} \in \mathbb{R}^n \text{ und } \mathbf{A} \in \mathbb{R}^{n \times n}$$

eine symmetrische Matrix. Berechnen Sie alle Richtungsableitungen von f im Punkt \mathbf{x} . Wann gilt $\text{grad } f(\mathbf{x}) = \mathbf{o}$?

Übungsaufgabe 129. (T) Berechnen Sie die Richtungsableitung der Funktion $f(x, y) = \ln(\sqrt{x^2 + y^2})$ in einem Punkt $(x_0, y_0) \neq (0, 0)$ in Richtung des Nullpunktes.

Übungsaufgabe 130. (T*) Ein Skifahrer fährt in der Falllinie der Ebene $6x + 8y - z = 22$. Geben Sie die Gleichung der Bahn und deren Neigung gegen die x, y -Ebene an (Die Schwerkraft wirke in Richtung der z -Achse). Wie hängt die Position $\mathbf{x}(t) = (x(t), y(t), z(t))$ von der Zeit ab, wenn der Start bei $t = 0$ im Punkt $\mathbf{x}_0 = (3, 2, 12)$ erfolgt? (Von der Reibung ist der Einfachheit halber abzusehen.)

Übungsaufgabe 131. (T) In welche Richtung muss ein der Skifahrer (vgl. Aufgabe 130) vom Punkt $(1, 1, 0)$ aus starten, wenn er auf einer durch die Gleichung $z = 3 - x^2 - 2y^2$ bestimmten Piste in Richtung des stärksten Gefälles abfahren will?

2.2.4 Entfaltung der Differentialrechnung

Inhalt in Kurzfassung: In Analogie zur Differentialrechnung in einer Variablen (siehe Mathematik 1) können auch im Fall mehrerer Variablen wichtige Konzepte entwickelt und Sachverhalte bewiesen werden. Beispielsweise ist jede differenzierbare Funktion erst recht stetig. Außerdem gilt weiterhin die Kettenregel: Die Ableitung einer Verkettung ist die Verkettung der Ableitungen. Einen neuartigen Aspekt zeigt der Schwarzsche Vertauschbarkeitssatz, wonach es im Fall stetiger Differenzierbarkeit bei höheren partiellen Ableitungen nach verschiedenen Variablen nicht auf die Reihenfolge der Differentiation ankommt. Der Begriff der höheren Ableitung in mehreren Variablen ist insofern komplizierter als bei nur einer Variablen, weil man Räume linearer Abbildungen von mit jeder Differentiation wachsender Dimension betrachten muss. Trotzdem können auch höhere Ableitungen in befriedigender Weise durch die entsprechenden höheren partiellen Ableitungen beschrieben werden. Zum Beispiel wird die zweite Ableitung einer reellwertigen Funktion in n Variablen durch die $n \times n$ -Matrix sämtlicher partieller Ableitungen zweiter Ordnung, die sogenannte Hesse-Matrix, beschrieben.

Wie im Eindimensionalen folgt auch im allgemeinen Fall aus Differenzierbarkeit stets Stetigkeit:

Proposition 2.2.4.1. *Ist die Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ im Punkt $\mathbf{x}_0 \in D$ differenzierbar, dann auch stetig.*

Übungsaufgabe 132. (E) Beweisen Sie Proposition 2.2.4.1. Hinweis: Der Beweis im Eindimensionalen lässt sich fast unmittelbar übertragen.

Die Berechnung der Ableitung einer höherdimensionalen Funktion f besteht in der Ermittlung der Funktionalmatrix. Deren Eintragungen sind die partiellen Ableitungen, die sich wiederum als Ableitung eindimensionaler reeller Funktionen berechnen lassen. Folglich besteht keine dringende Notwendigkeit, für die Differentialrechnung in mehreren Variablen neue Differentiationsregeln herzuleiten. Die aus Mathematik 1 bekannten Regeln reichen meistens aus. Mit ihrer Hilfe lassen sich aber auch zahlreiche griffige Regeln für die höherdimensionale Differentiation herleiten. Hier begnügen wir uns damit, die folgende, sehr einprägsame **allgemeine Kettenregel** zu besprechen. Denn an ihr lässt sich die Wirkungsmacht der linearen Approximation, die der Differentialrechnung als sehr universelles Paradigma zugrundeliegt, besonders deutlich beobachten.

Satz 2.2.4.2. *Die Funktion $f : D_f \rightarrow \mathbb{R}^m$ mit $D_f \subseteq \mathbb{R}^n$ sei differenzierbar in $\mathbf{x}_0 \in D_f$, die Funktion $g : D_g \rightarrow \mathbb{R}^k$ mit $f(D_f) \subseteq D_g \subseteq \mathbb{R}^m$ sei differenzierbar in $f(\mathbf{x}_0) \in D_g$. Dann ist die Verkettung $g \circ f : D_f \rightarrow \mathbb{R}^k$ differenzierbar in \mathbf{x}_0 , und es gilt*

$$(g \circ f)'_{\mathbf{x}_0} = g'_{f(\mathbf{x}_0)} \circ f'_{\mathbf{x}_0}.$$

(Merkregel: Die Ableitung der Verkettung ist die Verkettung der Ableitungen.)

Beweis. (Beweisidee) Der Beweis der Kettenregel gelingt mit denselben Ideen wie für die eindimensionale Kettenregel: Nach Definition der Ableitung angewendet auf $f'_{\mathbf{x}_0}$ und $g'_{f(\mathbf{x}_0)}$ gilt

$$\begin{aligned}(g \circ f)(\mathbf{x}) &= g(f(\mathbf{x})) = g(f(\mathbf{x}_0) + f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0) + r_f(\mathbf{x})) = \\ &= g(f(\mathbf{x}_0)) + g'_{f(\mathbf{x}_0)}(f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0) + r_f(\mathbf{x})) + r_g(f(\mathbf{x})) = \\ &= g(f(\mathbf{x}_0)) + g'_{f(\mathbf{x}_0)}(f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)) + g'_{f(\mathbf{x}_0)}(r_f(\mathbf{x})) + r_g(f(\mathbf{x}))\end{aligned}$$

Mit Fehlerfunktionen r_f und r_g , die für $\mathbf{x} \rightarrow \mathbf{x}_0$ schnell genug klein werden, um zu garantieren, dass auch der Restterm

$$r_{g \circ f}(\mathbf{x}) := g'_{f(\mathbf{x}_0)}(r_f(\mathbf{x})) + r_g(f(\mathbf{x}))$$

hinreichend schnell klein wird, damit in der resultierenden Darstellung

$$(g \circ f)(\mathbf{x}) = (g \circ f)(\mathbf{x}_0) + (g'_{f(\mathbf{x}_0)} \circ f'_{\mathbf{x}_0})(\mathbf{x} - \mathbf{x}_0) + r_{g \circ f}(\mathbf{x})$$

die Definition für die Ableitung erfüllt ist. Somit folgt $(g \circ f)'_{\mathbf{x}_0} = g'_{f(\mathbf{x}_0)} \circ f'_{\mathbf{x}_0}$. Etwas genauer: Man hat

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|r_{g \circ f}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

nachzuweisen. Wegen $r_{g \circ f}(\mathbf{x}) := g'_{f(\mathbf{x}_0)}(r_f(\mathbf{x})) + r_g(f(\mathbf{x}))$ genügt es, die entsprechende Konvergenz für jeden der beiden Summanden $g'_{f(\mathbf{x}_0)}(r_f(\mathbf{x}))$ und $r_g(f(\mathbf{x}))$ nachzuweisen. Für den ersten Summanden verwendet man die Differenzierbarkeit von f in \mathbf{x}_0 , also

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|r_f(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

Diese Konvergenz gegen \mathbf{o} bleibt auch nach Anwendung der linearen Funktion $g'_{f(\mathbf{x}_0)}$ auf den Zähler bestehen.

Für den zweiten Summanden verwendet man, dass aus der Differenzierbarkeit erst recht die Stetigkeit von f , also $f(\mathbf{x}) \rightarrow f(\mathbf{x}_0)$ für $\mathbf{x} \rightarrow \mathbf{x}_0$ folgt und somit, wegen der Differenzierbarkeit von g , auch der zweite Restterm $r_g(f(\mathbf{x}))$ die geforderte Konvergenzeigenschaft hat. \square

Übungsaufgabe 133. (E) Vervollständigen Sie den Beweis von Satz 2.2.4.2, indem Sie die angedeuteten Abschätzungen ausführlich durchführen.

Wir illustrieren die Kettenregel am Beispiel $k = n = 1$, wo die Komposition $g \circ f$ also eine eindimensionale reelle Funktion ist, die jedoch über m Zwischenergebnisse (die den m Komponentenfunktionen f_i von f entsprechen) ermittelt wird. In dieser Situation hängt also $g = g(y_1, \dots, y_m)$ von m Variablen ab, deren Werte durch die Komponentenfunktionen $f_i : D \rightarrow \mathbb{R}$, $i = 1, \dots, m$, von $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}$ vermittelt werden. Die Komposition $h = g \circ f : \mathbb{R} \rightarrow \mathbb{R}$ hat daher die Bauart

$$h = g \circ f : x \mapsto g(f_1(x), \dots, f_m(x)).$$

2 Differentialrechnung in mehreren Variablen

Sei nun $x_0 \in D$. Laut Kettenregel ist $h'(x_0) = g'_{f(x_0)} \circ f'_{x_0}$. Wegen $k = n = 1$ werden beide involvierten Ableitungen durch m -dimensionale Vektoren dargestellt,

$$g'_{f(x_0)} = \text{grad}_{\mathbf{x}_0}(g) = \left(\frac{\partial g}{\partial y_1}(f(x_0)), \dots, \frac{\partial g}{\partial y_m}(f(x_0)) \right)$$

als Zeilenvektor,

$$f'_{x_0} = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x_0) \\ \vdots \\ \frac{\partial f_m}{\partial x}(x_0) \end{pmatrix}$$

als Spaltenvektor. Nach den Regeln der Matrizenmultiplikation ist das gesuchte Produkt $h'(x_0) = g'_{f(x_0)} \circ f'_{x_0}$ folglich als Skalarprodukt dieser beiden Vektoren zu berechnen und wird zu

$$h'(x_0) = g'_{f(x_0)} \circ f'_{x_0} = \sum_{i=1}^m \frac{\partial g}{\partial y_i}(f(x_0)) \cdot \frac{\partial f_i}{\partial x}(x_0).$$

Gern schreibt man kürzer und sehr einprägsam:

$$(g(f_1(x), \dots, f_m(x)))' = \frac{dg}{dx} = \sum_{i=1}^m \frac{dg}{df_i} \cdot \frac{df_i}{dx}$$

Übungsaufgabe 134. (E) Es sei $\mathbf{x} \rightarrow \tilde{\mathbf{x}}$ eine lineare Variablen-Transformation im \mathbb{R}^2 oder \mathbb{R}^3 , also $\tilde{\mathbf{x}} = S\mathbf{x}$ mit einer regulären 2×2 bzw. 3×3 Matrix S . Rechnen Sie nach: Ist $u(\mathbf{x})$ eine differenzierbare Funktion und $\tilde{u}(\tilde{\mathbf{x}}) := u(\mathbf{x})$ die transformierte Funktion, dann gilt

$$S^T \tilde{u}'(\tilde{\mathbf{x}}) = u'(\mathbf{x}).$$

(Man beachte, dass auf der linken Seite die Ableitungen $\frac{\partial}{\partial \tilde{x}}$ nach den gewellten Koordinaten betrachtet werden, auf der rechten Seite hingegen die Ableitungen $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ nach den ungewellten Koordinaten.) Man schreibt daher auch kurz $\nabla' = (S^{-1})^T \nabla$, vgl. Abschnitt 4.4.1.

Von besonderem Interesse für spätere Anwendungen ist der **Schwarzsche Vertauschbarkeitssatz**:

Satz 2.2.4.3. Die reellwertige Funktion $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^n$ möge von $n \geq 2$ Variablen x_1, \dots, x_n abhängen. Sei \mathbf{x}_0 ein innerer Punkt von D mit einer Umgebung U , in der sämtliche partiellen Ableitungen bis zur Ordnung m existieren und stetig sind. Dann hängen die partiellen Ableitungen bis zur Ordnung m nicht von der Reihenfolge der Differentiation ab. (Speziell für $n = m = 2$, $x_1 = x$ und $x_2 = y$ bedeutet das $f_{xy} = f_{yx}$ auf U .)

Beweis. (Andeutung) Der Beweis erfordert einige technische Arbeit, und wir begnügen uns hier mit einer Andeutung. Zunächst überzeugt man sich davon, dass es genügt, den explizit hervorgehobenen Fall $m = n = 2$, $x_1 = x$ und $x_2 = y$ zu behandeln. Um diesen zu beweisen, wählt man um \mathbf{x}_0 ein kleines achsenparalleles Rechteck $V \subseteq U$ so,

dass sich innerhalb von V die involvierten Ableitungen nur wenig ändern, was wegen der Stetigkeitsvoraussetzung möglich ist. Jedes $\mathbf{x} \in V$ bestimmt ein achsenparalleles Rechteck $R_{\mathbf{x}} \subseteq V$ mit den Eckpunkten \mathbf{x}_0 und \mathbf{x} . Es gibt zwei Wege von \mathbf{x}_0 entlang des Randes von $R_{\mathbf{x}}$. Beide geben Anlass zu einer Approximation von $f(\mathbf{x})$ mit Hilfe erster und zweiter Ableitungen von f , etwas genauer: eine Anwendung des Mittelwertsatzes auf die ersten Ableitungen bringt die zweiten ins Spiel. In einem Fall tritt f_{xy} auf, im anderen f_{yx} . Setzt man diese beiden Approximationen geschickt in Beziehung zueinander, kann man im Grenzübergang, wenn sich V auf \mathbf{x}_0 zusammenzieht, auf $f_{xy}(\mathbf{x}_0) = f_{yx}(\mathbf{x}_0)$ schließen. \square

Nachdem wir schon mit höheren partiellen Ableitungen gearbeitet haben, wollen wir uns auch noch überlegen, wie generell höhere Ableitungen einer Funktion $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$ zu verstehen sind. Es liegt nahe, nach einer Iteration des Ableitungsprozesses zu fragen. Dazu ist lediglich ein kleines Hindernis zu überwinden. Das besteht darin, dass die Ableitung f' eine auf D definierte Funktion ist, deren Werte aber nicht in \mathbb{R}^m liegen, sondern lineare Abbildungen $f'_{\mathbf{x}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ sind. Wir können uns aber Erkenntnisse aus der Linearen Algebra zunutze machen, nämlich dass die Menge $L(\mathbb{R}^n, \mathbb{R}^m)$ aller linearen Abbildungen von \mathbb{R}^n nach \mathbb{R}^m einen $m \cdot n$ -dimensionalen Vektorraum bilden, dessen Elemente bezüglich der kanonischen Basen Darstellungen als $m \times n$ -Matrizen haben. Also liegt eine natürliche Interpretation der Ableitung f' von f als Funktion

$$f' : D \rightarrow \mathbb{R}^{m \cdot n}, \quad \mathbf{x} \mapsto f'_{\mathbf{x}}$$

vor. Auf so eine Funktion sind die Begriffe Stetigkeit, Differenzierbarkeit etc. anwendbar. Wir beobachten: Der Definitionsbereich bleibt gegenüber der ursprünglichen Funktion f unverändert, der Zielbereich hingegen wächst bei jedem Differentiationsschritt in der Dimension um einen Faktor n (die Anzahl der Variablen, von denen f abhängt). Somit kann man rekursiv definieren:

Definition 2.2.4.4. Für eine Funktion $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, sind die höheren Ableitungen $f^{(k)}$, $k \in \mathbb{N}$, (sofern existent) rekursiv definiert durch

$$f^{(0)} := f : D \rightarrow \mathbb{R}^m, \quad f^{(k+1)} := (f^{(k)})' : D \rightarrow \mathbb{R}^{n^{k+1} \cdot m}.$$

Die Funktion f heißt **k -mal oder k -fach differenzierbar**, sofern $f^{(k)}$ existiert. Existiert $f^{(k)}$ für alle $k \in \mathbb{N}$, so heißt f **unendlich oft differenzierbar**.

Für die erste Ableitung so einer Funktion f ist laut Proposition 2.2.3.1 die Differenzierbarkeit von f äquivalent mit der jeder der m Komponentenfunktionen von f . Außerdem ist nach Satz 2.2.3.2 die Existenz aller partiellen Ableitungen notwendig für Differenzierbarkeit, ihre Stetigkeit sogar hinreichend. Überdies ist nach Proposition 2.1.2.4 auch die Stetigkeit einer Funktion mit höherdimensionaler Zielmenge äquivalent zur Stetigkeit in jeder einzelnen Komponente, folglich ist Stetigkeit der partiellen Ableitungen sogar äquivalent mit stetiger Differenzierbarkeit. Diese Zusammenhänge lassen sich mittels Induktion nach k mühelos von der ersten auf die k -te Ableitung von f übertragen. Analoges gilt für den Satz von Schwarz, 2.2.4.3. Wir fassen zusammen:

2 Differentialrechnung in mehreren Variablen

Satz 2.2.4.5. Für eine Funktion $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, und ihre mehrfache Differenzierbarkeit an einer Stelle $\mathbf{x}_0 \in D$ betreffend gelten folgende Aussagen:

1. Aus der k -fachen Differenzierbarkeit von f in \mathbf{x}_0 folgt die Existenz sämtlicher partieller Ableitungen von f in \mathbf{x}_0 bis zur Ordnung k .
2. Existieren sämtliche partiellen Ableitungen von f bis zur Ordnung k in einer Umgebung von \mathbf{x}_0 und sind sie dort stetig, so ist f auch k -mal differenzierbar.
3. Die Funktion ist genau dann k -mal stetig differenzierbar in einer Umgebung U von \mathbf{x}_0 , wenn auf U alle partiellen Ableitungen von f bis zur Ordnung k existieren und stetig sind.
4. Ist f in einer Umgebung U von \mathbf{x}_0 k -mal stetig differenzierbar, so hängen dort alle partiellen Ableitungen bis zur Ordnung k von f davon ab, nach welcher Variablen wie oft differenziert worden ist, nicht aber von der Reihenfolge der Differentiation.

Setzen sich die Komponentenfunktionen von f so aus einfachen, beliebig oft differenzierbaren Bausteinen zusammen, dass die Differentiationsregeln aus der eindimensionalen Differentialrechnung beliebig oft anwendbar sind, ist f also auch als höherdimensionale Funktion unendlich oft differenzierbar.

Im Fall $m = 1$ (also f reellwertig) ist $f^{(k)} : D \rightarrow \mathbb{R}^{n^k}$. Sehr gut kennen wir bereits die Situation $k = 1$: Die erste Ableitung $f^{(1)} = f' : D \rightarrow \mathbb{R}^n$ wird durch den Gradienten $\mathbf{x} \mapsto \text{grad}_{\mathbf{x}}(f)$ dargestellt. Für $k = 2$ ist $f^{(k)} = f^{(2)}$ gegeben durch sämtliche $\frac{\partial^2 f}{\partial x_i \partial x_j}$ ($i, j = 1, \dots, n$), die sich in natürlicher Weise in einer $n \times n$ -Matrix darstellen lassen, der sogenannten **Hesse-Matrix** von f . Ist f zweimal stetig differenzierbar, so ist die Hesse-Matrix nach Satz 2.2.4.3 symmetrisch. Die Hesse-Matrix wird uns im Zusammenhang mit der Extremwertbestimmung von Funktionen in mehreren Variablen wieder begegnen.

Die Behandlung größerer Werte von $k \geq 3$ erfordert den Einsatz von Objekten, die alle partiellen Ableitungen bis zur Ordnung k enthalten. Damit ist eine Verbindung zu den Tensoren höherer Stufe aus 1.4.2 hergestellt, die wir hier allerdings nicht weiter verfolgen.

Übungsaufgabe 135. (E) Es sei $f(x, y)$ definiert durch $f(x, y) = \frac{x^3 y - x y^3}{x^2 + y^2}$ für $(x, y) \neq (0, 0)$ und $f(0, 0) = 0$. Überprüfen Sie die partiellen Ableitungen f_{xy} und f_{yx} auf Stetigkeit am Ursprung und zeigen Sie, dass $f_{xy}(0, 0) \neq f_{yx}(0, 0)$.

Übungsaufgabe 136. (T) Gegeben ist die Funktion $f(x, y) = y\sqrt{2x^2 + y^2}$. Entscheiden Sie ob f stetig differenzierbar ist und geben Sie alle ersten partiellen Ableitungen an! Skizze!

Übungsaufgabe 137. (T) Gegeben sind $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = \frac{x}{y}$ und $g : \mathbb{R} \rightarrow \mathbb{R}^2$ mit $g(x) = (x^2 + 1, x^2 - 1)$. Berechnen Sie die Ableitungen von f und g sowie auf zwei Arten die von $g \circ f$ und $f \circ g$. Hinweis: Kettenregel.

Übungsaufgabe 138. (T) Entscheiden Sie, ob die Funktion

$$u(x, t) = \cosh(x + 2t)$$

eine Lösung der Schwingungsgleichung $u_{tt} = 4u_{xx}$ ist.

Übungsaufgabe 139. (T) Entscheiden Sie, ob die Funktion

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{t}}$$

eine Lösung der Wärmeleitungsgleichung $u_t = u_{xx}$ ist.

Übungsaufgabe 140. (T) Berechnen Sie nach der Kettenregel

1. $\frac{df}{dt}, \frac{d^2f}{dt^2}$ für $f(x, y) = 1 - x^2 - y^2$, $x(t) = \sin t$, $y(t) = \cos t$
2. $\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}$ für $f(x, y) = y \ln x$, $x(u, v) = \sqrt{u^2 - v^2}$, $y(u, v) = \sqrt{u^2 + v^2}$.

2.2.5 Geometrische Interpretation der Ableitung

Inhalt in Kurzfassung: Um die Intuition für die Bedeutung der Ableitung einer Funktion $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, zu entwickeln, werden verschiedene Dimensionen m und n durchgespielt: $m = n = 1$ (Anstieg einer reellen Funktion, Wiederholung aus Mathematik 1), $m = 1$ und $n = 2$ (Funktionsgebirge), $m = 1$ und $n = 3$ (Funktion als Potential mit Gradientenfeld), $n = 1$ und $m = 2, 3$ (Kurven mit Tangenten), $n = 2$ und $m = 3$ (Fläche mit Tangentialebene im Raum), $m = n = 2$ (ebene Koordinatentransformation, insbesondere Polarkoordinaten; alternativ: ebene Vektorfelder), $m = n = 3$ (räumliche Koordinatentransformation, insbesondere Kugelkoordinaten; alternativ: Vektorfelder im Raum).

Wir wollen nun die geometrische Interpretation der Ableitung höherdimensionaler Funktionen $f : D \rightarrow \mathbb{R}^m$, $D \subseteq \mathbb{R}^n$, besprechen. Dabei orientieren wir uns an der Gliederung aus Abschnitt 2.1 nach verschiedenen Werten von m und n , beginnend mit 2.1.3.

Reelle Funktionen ($m = n = 1$): Dieser Fall war Gegenstand der eindimensionalen Differentialrechnung in Mathematik 1. Die Ableitung $f'(x)$ von f an einer Stelle $x \in D \subseteq \mathbb{R}$ gibt den Anstieg des Funktionsgraphen, genauer: der Tangente an den Funktionsgraphen von f im Punkt x an.

Funktionsgebirge ($m = 1, n = 2$): Den Fall eines Funktionsgebirges haben wir in 2.2.1 und 2.2.2 zum Ausgangspunkt genommen, um den Begriff der Ableitung einer höherdimensionalen Funktion zu entwickeln. Aus den dort angestellten Überlegungen ergibt sich die Interpretation von f' als Tangentialebene, genauer: Für $\mathbf{x}_0 = (x_0, y_0) \in D$ ist der Graph der Funktion $\mathbf{x} \mapsto f(\mathbf{x}_0) + f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)$ die Tangentialebene an (den Graph

2 Differentialrechnung in mehreren Variablen

von) f im Punkt $(x_0, y_0, f(x_0, y_0))$.

Potentiale ($m = 1, n = 3$): Man kann versuchen, die Situation vom Fall $m = 1, n = 2$ durch Analogieschlüsse zu übertragen. Für beide Fälle gibt es jedoch noch eine andere wichtige Interpretation. Man kann eine reellwertige Funktion $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^3$ nämlich häufig als Potential interpretieren, das zum Beispiel das Energieniveau $f(\mathbf{x}_0)$ an einem Punkt $\mathbf{x}_0 = (x_0, y_0, z_0)$ im Raum angibt. Dann ist die Matrixdarstellung von $f'_{\mathbf{x}_0}$ der Gradientenvektor

$$\text{grad}_{\mathbf{x}_0} f = (f_x(\mathbf{x}_0), f_y(\mathbf{x}_0), f_z(\mathbf{x}_0)),$$

der den größten Anstieg von f angibt. Die Ableitung $f' : D \rightarrow \mathbb{R}^3$ ist also als Vektorfeld interpretierbar, das man in dieser Situation, insbesondere wenn f zweimal stetig differenzierbar ist, ein **Potentialfeld** nennt. Alternative Bezeichnungen sind **Gradientenfeld** oder **konservatives Vektorfeld**. In 2.2.6 und auch an späteren Stellen werden wir uns noch ausführlicher damit beschäftigen.

Kurven ($m = 2, 3, n = 1$): In diesem Fall interpretieren wir

$$f : t \mapsto \begin{pmatrix} f_1(t) \\ f_2(t) \end{pmatrix} \quad \text{beziehungsweise} \quad f : t \mapsto \begin{pmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \end{pmatrix}$$

als (Parametrisierung einer) Kurve in der Ebene ($m = 2$) beziehungsweise im Raum ($m = 3$). Wie schon in Mathematik 1 besprochen lässt sich der Vektor

$$f'_t = f'(t) = \begin{pmatrix} f'_1(t) \\ f'_2(t) \end{pmatrix} \quad \text{beziehungsweise} \quad f'_t = f'(t) = \begin{pmatrix} f'_1(t) \\ f'_2(t) \\ f'_3(t) \end{pmatrix}$$

als Tangentialvektor an die Kurve deuten. Auf eine mögliche Schwierigkeit, die auch in höheren Dimensionen zu bedenken ist, ist aber hinzuweisen: Es können nämlich durchaus auch bei unendlich oft differenzierbarem f **Singularitäten** auftreten. Darunter versteht man Stellen t mit $f'(t) = \mathbf{o}$, wo die Interpretation der Ableitung als Tangentialvektor also nicht sinnvoll ist. Geometrisch kann eine Singularität Unterschiedliches bedeuten. Das sei anhand der drei Beispiele

$$f : t \mapsto \begin{pmatrix} t^2 \\ t^2 \end{pmatrix}, \quad g : t \mapsto \begin{pmatrix} t^3 \\ t^3 \end{pmatrix} \quad \text{und} \quad h : t \mapsto \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}$$

illustriert, die wegen

$$f'(t) = \begin{pmatrix} 2t \\ 2t \end{pmatrix}, \quad g'(t) = \begin{pmatrix} 3t^2 \\ 3t^2 \end{pmatrix} \quad \text{und} \quad h'(t) = \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix}$$

alle bei $t = 0$ eine Singularität aufweisen.

Übungsaufgabe 141. (P) Skizzieren Sie die Kurven f , g und h .

Der Bildbereich von f ist die Menge aller $(s, s) \in \mathbb{R}^2$ mit $s = t^2 \geq 0$, $t \in \mathbb{R}$, also die Halbgerade, bestehend aus der „positiven“ Hälfte der Diagonale inklusive Nullpunkt $f(0) = \mathbf{o}$ als Endpunkt. An einem Endpunkt existiert keine Tangente im herkömmlichen Sinn. Im Gegensatz dazu stellt g die volle Diagonale dar, die anschaulich sehr wohl auch im Ursprung eine Tangente hat, nämlich sich selbst. Trotzdem gilt $g'(0) = \mathbf{o}$ nicht als Tangentialvektor. Sehr typisch für Singularitäten ist die Situation bei h , das sich geometrisch als die Vereinigung der beiden Kurven $y = \sqrt{x^3}$ und $y = -\sqrt{x^3}$ mit $x \geq 0$ auffassen lässt. Bei ihrem Berührungspunkt \mathbf{o} bilden diese beiden Äste eine Spitze. Dieselbe Kurve ergibt sich auch als Grundriss (die ersten beiden Koordinaten betrachten) der Kurve

$$k : t \mapsto \begin{pmatrix} t^2 \\ t^3 \\ t \end{pmatrix}.$$

im Raum, die wegen

$$k'(t) = \begin{pmatrix} 2t \\ 3t^2 \\ 1 \end{pmatrix} \neq \mathbf{o}.$$

keine Singularität aufweist. Für den Parameterwert $t = 0$ ist $k'(t) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ ein parallel zur z -Achse senkrecht nach oben weisender Tangentialvektor.

Übungsaufgabe 142. (P) Man stelle die Kurve k in Grund-, Auf-, Kreuz und Schrägriss dar.

Ebene Koordinatentransformationen ($n = m = 2$): So wie im soeben behandelten Fall der krummen Fläche im Raum ($n = 2, m = 3$) brauchen wir für eine sinnvolle Visualisierung wieder zwei Koordinatensysteme. Für $n = m = 2$ wird eine achsenparallele Rasterung des Bereichs D durch f in einen gekrümmten, nun aber ebenen Bereich transformiert mit ganz ähnlicher Interpretation wie zuvor. Die Ableitung (Funktionalmatrix) $f'_{\mathbf{x}_0}$ von f im Punkt \mathbf{x}_0 verzerrt ein kleines Quadrat um den Punkt \mathbf{x}_0 in ein kleines Parallelogramm im Bildbereich. Die Funktionaldeterminante $\det f'_{\mathbf{x}_0}$ ist der Flächenvergrößerungsfaktor von f in der Umgebung von \mathbf{x}_0 . Als Beispiel behandeln wir den Zusammenhang von kartesischen und Polarkoordinaten:

$$\begin{aligned} r &= \sqrt{x^2 + y^2} & x &= r \cos(\varphi) \\ \cos \varphi &= \frac{x}{\sqrt{x^2 + y^2}} & y &= r \sin(\varphi). \\ \sin \varphi &= \frac{y}{\sqrt{x^2 + y^2}} \end{aligned}$$

2 Differentialrechnung in mehreren Variablen

Wir betrachten die Abbildung $f : (r, \varphi) \mapsto (x, y)$ mit $x = x(r, \varphi) = r \cos(\varphi)$ und $y = y(r, \varphi) = r \sin(\varphi)$. Die Funktionalmatrix von f hat die Form

$$f'_{(r,\varphi)} = \begin{pmatrix} x_r & x_\varphi \\ y_r & y_\varphi \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix}.$$

Folglich ist die Funktionaldeterminante gegeben durch

$$\det f'_{(r,\varphi)} = r \cos^2(\varphi) + r \sin^2(\varphi) = r.$$

Flächen ($m = 3, n = 2$): Eine Funktion $f : D \rightarrow \mathbb{R}^3$, $D \subseteq \mathbb{R}^2$, stellen wir uns vorzugsweise als eine Parametrisierung einer Fläche vor, wobei f den ebenen Bereich D in ein typischerweise gekrümmtes Flächenstück im Raum überführt. Stellt man sich D hinreichend fein achsenparallel gerastert vor und ist f an einer Stelle $\mathbf{x}_0 \in D$ differenzierbar, so gehen kleine Quadrate in D über in annähernde Parallelelogramme. Die Ableitung $f'_{\mathbf{x}_0}$ ist die lineare Approximation von f in \mathbf{x}_0 , die die Quadrate in exakte und periodische Parallelelogramme innerhalb der Tangentialebene der Fläche F im \mathbb{R}^3 überführt. Der Unterschied zur geometrischen Deutung im Fall $n = 2$ und $m = 1$ besteht darin, dass zur Veranschaulichung jetzt zwei Koordinatensysteme notwendig sind – ein zweidimensionales zur Darstellung des Definitionsbereichs D und ein dreidimensionales für den Zielbereich –, während $n + m = 2 + 1 = 3$ Dimensionen in einem einzigen räumlichen Modell dargestellt werden können, weil man die eindimensionalen Funktionswerte entlang einer dritten Koordinatenachse senkrecht über dem zweidimensionalen Definitionsbereich auftragen kann.

Räumliche Koordinatentransformationen ($m = n = 3$): Ganz Analoges wie bei der Transformation ebener Koordinaten gilt für $m = n = 3$, nur dass es um Quadrate geht, die durch f in Parallelepipede transformiert werden. Die Funktionaldeterminante ist dann der Volumsvergrößerungsfaktor. Beispielsweise für Kugelkoordinaten sind ganz ähnliche Rechnungen wie oben für Polarkoordinaten durchzuführen. Man geht aus von der Transformation

$$f : \begin{pmatrix} r \\ \alpha \\ \beta \end{pmatrix} \mapsto \begin{pmatrix} x(r, \alpha, \beta) \\ y(r, \alpha, \beta) \\ z(r, \alpha, \beta) \end{pmatrix} := \begin{pmatrix} r \cos(\alpha) \cos(\beta) \\ r \sin(\alpha) \cos(\beta) \\ r \sin(\beta) \end{pmatrix}.$$

Die zugehörige Funktionalmatrix an einer bestimmten Stelle (r, α, β) ist daher

$$f'_{(r,\alpha,\beta)} = \begin{pmatrix} x_r & x_\alpha & x_\beta \\ y_r & y_\alpha & y_\beta \\ z_r & z_\alpha & z_\beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \cos(\beta) & -r \sin(\alpha) \cos(\beta) & -r \cos(\alpha) \sin(\beta) \\ \sin(\alpha) \cos(\beta) & r \cos(\alpha) \cos(\beta) & -r \sin(\alpha) \sin(\beta) \\ \sin(\beta) & 0 & r \cos(\beta) \end{pmatrix}.$$

Verwendet man zur Berechnung der Funktionaldeterminante die Regel von Sarrus, so bleiben wegen der Eintragung 0 in der Mitte der letzten Zeile von den sechs Summanden nur vier übrig, nämlich

$$\det f'_{(r,\alpha,\beta)} = A + B - C - D$$

mit:

$$\begin{aligned}
 A &= (\cos(\alpha) \cos(\beta)) \cdot (r \cos(\alpha) \cos(\beta)) \cdot (r \cos(\beta)) &= r^2 \cos^2(\alpha) \cos^3(\beta) \\
 B &= (-r \sin(\alpha) \cos(\beta)) \cdot (-r \sin(\alpha) \sin(\beta)) \cdot (\sin(\beta)) &= r^2 \sin^2(\alpha) \cos(\beta) \sin^2(\beta) \\
 C &= (\sin(\beta)) \cdot (r \cos(\alpha) \cos(\beta)) \cdot (-r \cos(\alpha) \sin(\beta)) &= -r^2 \cos^2(\alpha) \cos(\beta) \sin^2(\beta) \\
 D &= (r \cos(\beta)) \cdot (\sin(\alpha) \cos(\beta)) \cdot (-r \sin(\alpha) \cos(\beta)) &= -r^2 \sin^2(\alpha) \cos^3(\beta)
 \end{aligned}$$

Wegen

$$\begin{aligned}
 A - C &= r^2 \cos^2(\alpha) \cos(\beta)(\cos^2(\beta) + \sin^2(\beta)) &= r^2 \cos^2(\alpha) \cos(\beta) &\text{ und} \\
 B - D &= r^2 \sin^2(\alpha) \cos(\beta)(\cos^2(\beta) + \sin^2(\beta)) &= r^2 \sin^2(\alpha) \cos(\beta)
 \end{aligned}$$

folgt insgesamt

$$\det f'_{(r,\alpha,\beta)} = (A - C) + (B - D) = r^2(\cos^2(\alpha) + \sin^2(\alpha)) \cos(\beta) = r^2 \cos(\beta).$$

Diese Formel wird uns im Zusammenhang mit der Substitutionsregel für die höherdimensionalen Integration noch wertvolle Dienste leisten.

Vektorfelder ($m = n = 2, 3$): Bei Vektorfeldern erweisen sich gewisse Größen als besonders interessant, die sich aus den partiellen Ableitungen der Komponentenfunktionen zusammensetzen. Speziell zu nennen sind Rotation und Divergenz. Ihre Bedeutung wird im Zusammenhang mit den Integralsätzen von Green, Stokes und Gauß, siehe Abschnitt 4.4, deutlich werden, worauf an dieser Stelle verwiesen sei.

Wie angekündigt wenden wir uns noch etwas ausführlicher den Potential- oder Gradientenfeldern zu.

2.2.6 Potentialfelder (Gradientenfelder)

Inhalt in Kurzfassung: Aus dem Satz von Schwarz über die Vertauschbarkeit partieller Differentiation nach verschiedenen Variablen folgt, dass stetig differenzierbare Gradientenfelder (Potentialfelder) notwendig die sogenannten Integrabilitätsbedingungen erfüllen. Ein nichttrivialer Satz besagt: Zusammen mit einer geometrischen Bedingung an den Definitionsbereich (einfach zusammenhängendes Gebiet) sind diese Bedingungen aber auch hinreichend. Unmittelbar ergibt sich eine Methode zur Berechnung des Potentials eines gegebenen Gradientenfeldes. Diese wieder anhand eines Beispiels erläutert.

Zur Erinnerung: Ein n -dimensionales Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$, heißt **Potential-** oder **Gradientenfeld**, wenn es von der Gestalt $\mathbf{v}(\mathbf{x}) = \text{grad}_{\mathbf{x}}(f)$ mit einer Funktion $f : D \rightarrow \mathbb{R}$ ist. So ein f heißt **Potential** von \mathbf{v} und wird deshalb auch gerne mit dem Buchstaben P bezeichnet.

Im Fall $n = 3$ und

$$\mathbf{v} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} u(x, y, z) \\ v(x, y, z) \\ w(x, y, z) \end{pmatrix}$$

2 Differentialrechnung in mehreren Variablen

mit Komponentenfunktionen $u, v, w : D \rightarrow \mathbb{R}$, ist also $u = f_x, v = f_y$ und $w = f_z$. Ist f zweimal stetig differenzierbar, so folgt aus dem Satz von Schwarz (2.2.4.3):

$$\begin{aligned}u_y &= f_{xy} = f_{yx} = v_x \\u_z &= f_{xz} = f_{zx} = w_x \\v_z &= f_{yz} = f_{zy} = w_y\end{aligned}$$

Diese drei sogenannten **Integrabilitätsbedingungen** sind bei einem stetig differenzierbaren 3-dimensionalen Vektorfeld \mathbf{v} notwendig dafür, dass es sich um ein Gradientenfeld handelt. Im 2-dimensionalen Fall

$$\mathbf{v} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$$

erhält man analog eine einzige Integrabilitätsbedingung, nämlich

$$u_y = f_{xy} = f_{yx} = v_x.$$

Es stellt sich die Frage, ob umgekehrt jedes stetig differenzierbare Vektorfeld \mathbf{v} , das die Integrabilitätsbedingungen erfüllt, ein Gradientenfeld ist, d.h. ob man eine reellwertige Funktion f mit $f' = \text{grad } f = \mathbf{v}$ finden kann. Es geht also um eine höherdimensionale Variante des Problems, zu einer gegebenen reellen Funktion eine Stammfunktion zu finden.

Wir wollen uns nun überlegen, dass wir unser Problem dann tatsächlich auf den eindimensionalen Fall zurückführen können, wenn der Definitionsbereich D von \mathbf{v} von geeigneter Art ist, beispielsweise ein achsenparalleler Quader bzw. ein achsenparalleles Rechteck. Gibt man sich nämlich (hier im Fall $n = 2$) irgendeinen Punkt $\mathbf{x}_0 = (x_0, y_0) \in D$ vor, so muss die reelle Funktion $t \mapsto f(t, y_0)$ eine Stammfunktion der partiellen Ableitung $t \mapsto f_x(t, y_0) = u(t, y_0)$ bezüglich der ersten Variablen sein. Analoges gilt für die zweite Variable. Gibt man sich den Wert $f(\mathbf{x}_0)$ vor, so ist f also eindeutig bestimmt, weil man je zwei Punkte in D durch zwei solche achsenparallele Linien – eine parallel zur x -, die andere parallel zur y -Achse – verbinden kann. Mit anderen Worten: Wenn \mathbf{v} ein Gradientenfeld auf D ist, dann ist ein zugehöriges Potential f von \mathbf{v} bis auf eine additive Konstante eindeutig bestimmt und lässt sich durch iterierte Integration der beiden partiellen Ableitungen erhalten. Weil es zwei mögliche Reihenfolgen für die Integration gibt, wäre es a priori denkbar, dass man bei den beiden Möglichkeiten zu widersprüchlichen Ergebnissen kommt. Verantwortlich dafür, dass dies nicht der Fall ist, ist gerade die Integrabilitätsbedingung. Ein strenger Beweis übersteigt die uns hier zur Verfügung stehenden Mittel. Wir werden diesen Themenkreis aber später noch mehrmals berühren, vor allem im Zusammenhang mit Kurvenintegralen. Hier ist der folgende Satz die zentrale Aussage:

Satz 2.2.6.1. *Jedes stetig differenzierbare Gradientenfeld erfüllt die Integrabilitätsbedingungen. Umgekehrt ist jedes stetig differenzierbare Vektorfeld $\mathbf{v} : D \rightarrow \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$, das die Integrabilitätsbedingungen erfüllt ein Gradientenfeld, sofern D ein einfach zusammenhängendes Gebiet ist.*

Zur Erläuterung der in Satz 2.2.6.1 vorkommenden Begriffe: Unter einem n -dimensionalen **Gebiet** versteht man eine offene und zusammenhängende Teilmenge D des \mathbb{R}^n . In diesem Zusammenhang (nämlich wenn D als offen vorausgesetzt ist) lässt sich **zusammenhängend** auch dadurch charakterisieren, dass je zwei Punkte in D durch eine Kurve innerhalb von D verbunden werden kann.

Ein Gebiet lässt sich auch als offene Menge charakterisieren, innerhalb derer je zwei Punkte durch eine Kurve verbunden werden können. **Einfach zusammenhängend** wiederum heißt eine Teilmenge D des \mathbb{R}^n , wenn sie, anschaulich gesprochen, keine Löcher hat. Etwas genauer formuliert: Wenn man jede Schleife (d.h. Kurve, deren Anfangspunkt mit ihrem Endpunkt übereinstimmt) in D , die an einem Punkt fixiert ist, innerhalb von D stetig an diesen Punkt zusammengezogen werden kann. Zum Beispiel ist jede konvexe und auch jede sternförmige Menge einfach zusammenhängend. Nicht einfach zusammenhängend ist im Gegensatz dazu beispielsweise die Menge $\mathbb{R}^2 \setminus \{\mathbf{o}\}$ oder allgemein jede Teilmenge der Ebene, aus der man eine nichtleere und beschränkte Menge herausnimmt. Man kann auch so formulieren: Ein Gebiet in \mathbb{R}^2 ist einfach zusammenhängend, wenn sein Komplement zusammenhängend ist.

Nahe liegt die Frage, wie weit eine Funktion $f : D \rightarrow \mathbb{R}$ durch ihre partiellen Ableitungen bestimmt ist. Der nachfolgende Satz gibt die Antwort: bis auf eine additive Konstante.

Satz 2.2.6.2. *Es sei $D \subset \mathbb{R}^2$ ein Gebiet und die Funktionen $f, g : D \rightarrow \mathbb{R}$ haben stetige partielle Ableitungen. Gilt*

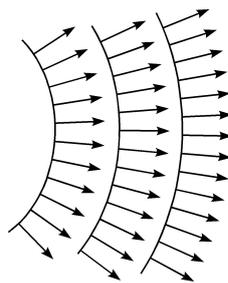
$$\begin{aligned} f_x &= g_x, \\ f_y &= g_y, \end{aligned}$$

dann gibt es eine Konstante $c \in \mathbb{R}$ mit $f(x) = g(x) + c$ für alle $x \in D$.

Sei $\mathbf{v} : G \rightarrow \mathbb{R}^2$ oder \mathbb{R}^3 ein Vektorfeld und $f = P : G \rightarrow \mathbb{R}$ ein Potential mit stetig differenzierbaren partiellen Ableitungen. Die Kurven (für $G \subset \mathbb{R}^2$) bzw. die Flächen (für $G \subset \mathbb{R}^3$)

$$P(\mathbf{x}) = P(x, y, z) = c \quad (c \in \mathbb{R})$$

heißen **Potentiallinien** bzw. **Potentialflächen**. Dann lässt sich zeigen, dass in jedem Punkt \mathbf{a} der Vektor $\mathbf{v}(\mathbf{a}) = \text{grad}P(\mathbf{a})$ senkrecht auf die Potentialfläche durch \mathbf{a} steht.



Zur Illustration folgen nun einige Beispiele.

Zentrales Kraftfeld: Wir betrachten ein zentrales Kraftfeld im $\mathbf{v} : \mathbb{R}^3 \setminus \{\mathbf{o}\} \rightarrow \mathbb{R}^3$, d.h. ein solches, wo in jedem Punkt $\mathbf{x} \in \mathbb{R}^3 \setminus \{\mathbf{o}\}$ eine Kraft in Richtung des Ursprungs

2 Differentialrechnung in mehreren Variablen

\mathbf{o} (oder umgekehrt) weist. Die Stärke der Kraft hänge dabei lediglich von $\|\mathbf{x}\|$, also von der Entfernung von \mathbf{x} zum Ursprung, ab und zwar auf stetige Weise. Es gebe also eine stetige reelle Funktion $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ mit

$$\mathbf{v}(\mathbf{x}) = f(\|\mathbf{x}\|) \frac{\mathbf{x}}{\|\mathbf{x}\|} \quad (f: \mathbb{R}^+ \rightarrow \mathbb{R} \text{ stetig}).$$

Wenig überraschend überzeugt kann man sich nämlich davon überzeugen, dass \mathbf{v} ein Potentialfeld ist. Ein Potential $P: \mathbb{R}^3 \setminus \{\mathbf{o}\} \rightarrow \mathbb{R}$ von \mathbf{v} ist nämlich gegeben durch

$$P(\mathbf{x}) := \int_1^{\|\mathbf{x}\|} f(s) ds.$$

Zur Überprüfung bemüht man die Formel von Leibniz zur Differentiation von Parameterintegralen (eine Kombination aus Hauptsatz der Differential- und Integralrechnung sowie der Kettenregel). Ihr zufolge gilt

$$P_x = \frac{\partial}{\partial x} \int_1^{(x^2+y^2+z^2)^{1/2}} f(s) ds = f((x^2+y^2+z^2)^{1/2}) \cdot \frac{1}{2} \frac{2x}{(x^2+y^2+z^2)^{1/2}} = f(\|\mathbf{x}\|) \frac{x}{\|\mathbf{x}\|},$$

$$P_y = \dots = f(\|\mathbf{x}\|) \frac{y}{\|\mathbf{x}\|} \quad \text{und} \quad P_z = \dots = f(\|\mathbf{x}\|) \frac{z}{\|\mathbf{x}\|},$$

womit

$$\text{grad}_{\mathbf{x}}(P) = f(\|\mathbf{x}\|) \frac{\mathbf{x}}{\|\mathbf{x}\|} = \mathbf{v}(\mathbf{x})$$

gezeigt ist.

Ein prominenter Spezialfall davon (diesmal allerdings mit dem Zentrum nicht in \mathbf{o} , sondern in einem beliebigen Punkt $\mathbf{u} \in \mathbb{R}^3$) ist das **Gravitationsfeld eines Massenpunktes**: Nach dem Newtonschen Gravitationsgesetz erzeugt ein Punkt der Masse m an der Stelle \mathbf{u} das Kraftfeld

$$\mathbf{f}(\mathbf{x}) = \frac{\gamma m}{\|\mathbf{u} - \mathbf{x}\|^2} \cdot \frac{\mathbf{u} - \mathbf{x}}{\|\mathbf{u} - \mathbf{x}\|} = \gamma m \frac{\mathbf{u} - \mathbf{x}}{\|\mathbf{u} - \mathbf{x}\|^3} \quad \text{für } \mathbf{x} \in \mathbb{R}^3 \setminus \{\mathbf{u}\}$$

mit einer naturgegebenen Gravitationskonstanten $\gamma \in \mathbb{R}$. Daher ist

$$-\gamma m \int_1^{\|\mathbf{u} - \mathbf{x}\|} \frac{ds}{s^2} = \frac{\gamma m}{\|\mathbf{u} - \mathbf{x}\|} - \gamma m$$

eine Potentialfunktion von \mathbf{f} und damit auch

$$\frac{\gamma m}{\|\mathbf{u} - \mathbf{x}\|}.$$

Denkt man sich die Masse nicht in einem Punkt konzentriert, sondern auf einen Körper $K \subseteq \mathbb{R}^3$ mit gleichmäßiger Dichte ρ verteilt, so erhält man für das **Gravitationsfeld dieses Körpers** das Gravitationspotential

$$\iiint_K \frac{\gamma \rho(\mathbf{u})}{\|\mathbf{u} - \mathbf{x}\|} du dv dw.$$

(Achtung Vorgriff: Dreifachintegrale wie dieses werden in Kapitel 4 behandelt.)

Zur Illustration noch ein **Rechenbeispiel** in zwei Variablen: Sei das Vektorfeld \mathbf{v} auf \mathbb{R}^2 durch die beiden Komponentenfunktionen

$$\begin{aligned}u(x, y) &= 2xy - 2y^2 + 1 \quad \text{und} \\v(x, y) &= x^2 - 4xy - 2y\end{aligned}$$

gegeben und ein Potential f mit $f(0, 0) = 3$ gesucht. Die Integrabilitätsbedingung

$$u_y(x, y) = 2x - 4y = v_x(x, y)$$

ist erfüllt, außerdem ist \mathbb{R}^2 einfach zusammenhängend, und beide Komponentenfunktionen sind stetig differenzierbar. Also gibt es nach Satz 2.2.6.1 ein f mit $f_x = u$ und $f_y = v$. Aus $f_x = u(x, y)$ folgt

$$f(x, y) = \int u(x, y) dx + g(y) = \int (2xy - 2y^2 + 1) dx + g(y) = x^2y - 2xy^2 + x + g(y)$$

mit einer nur von y abhängigen Funktion g . Man beachte, dass bei dieser Integration nach x die andere Variable y als Konstante zu behandeln ist. Weil f aber sehr wohl auch von y abhängt, hängt die additive Konstante beim unbestimmten Integral, die wir sonst meist mit c bezeichnen, hier von y ab, was durch den additiven Term $+g(y)$ berücksichtigt ist. Das erhaltene Zwischenergebnis setzen wir nun in die zweite Bedingung $f_y = v$ ein:

$$x^2 - 4xy + g'(y) = f_y(x, y) = v(x, y) = x^2 - 4xy - 2y,$$

wobei die Ableitung g' bezüglich der Variablen y zu bilden ist. Wir schließen aus der erhaltenen Gleichung weiter $g'(y) = -2y$, also

$$g(y) = \int -2y dy + c = -y^2 + c,$$

folglich $f(x, y) = x^2y - 2xy^2 + x - y^2 + c$. Aus der Anfangsbedingung folgt schließlich auch noch $c = 3$. Als einzig mögliche Lösung haben wir also die Funktion

$$f(x, y) := x^2y - 2xy^2 + x - y^2 + 3$$

erhalten, die tatsächlich $f_x = u$, $f_y = v$ und $f(0, 0) = 3$ erfüllt. Dieselbe Funktion f hätten wir erhalten, wenn wir die Reihenfolge der Variablen vertauscht, d.h. zuerst $f(x, y) = \int v(x, y) dy + g(x)$ angesetzt hätten und symmetrisch vorgegangen wären.

Dass die so erhaltene Funktion f tatsächlich eine Lösung des Problems ist, kann nur garantiert werden, wenn die Voraussetzungen von Satz 2.2.6.1 erfüllt sind, d.h. wenn es sich bei der Definitionsmenge um ein einfach zusammenhängendes Gebiet handelt. Zur Illustration noch eine Übungsaufgabe:

Übungsaufgabe 143. (T^*)

2 Differentialrechnung in mehreren Variablen

1. Führen Sie in obigem Beispiel die analoge Rechnung mit vertauschter Reihenfolge der Koordinaten durch.
2. Sei weiterhin $u(x, y) = 2xy - 2y^2 + 1$, jedoch $v(x, y) = x^2 - 4xy - 2y$. Wenden Sie dieselbe Rechenmethode wie vorher an. An welcher Stelle treten Probleme auf? Welche Voraussetzung in Satz 2.2.6.1 ist verletzt?
3. Untersuchen Sie das auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ definierte Vektorfeld mit den Komponentenfunktionen $u(x, y) = -\frac{y}{x^2+y^2}$ und $v(x, y) = \frac{x}{x^2+y^2}$ mit derselben Methode. Welche Probleme treten dabei auf, und wie lassen sich diese im Lichte von Satz 2.2.6.1 erklären?
4. Übertragen Sie die Methode auf 3-dimensionale Vektorfelder, indem Sie ein (nicht-triviales) 3-dimensionales Gradientenfeld Ihrer Wahl behandeln.

2.3 Nichtlineare Gleichungssysteme

In diesem Abschnitt werden aus der Linearen Algebra (Kapitel 1) bekannte Tatsachen über lineare Abbildungen und Gleichungssysteme auf den nichtlinearen Fall übertragen, wobei allerdings globale Aussagen zu lokalen abgeschwächt werden müssen. Die Übertragung gelingt durchwegs gemäß dem generellen Paradigma der Differentialrechnung, nichtlineare Funktionen lokal durch lineare zu approximieren und die Probleme damit möglichst auf die Lineare Algebra zurückzuspielen. Entsprechend treten als entscheidende Voraussetzungen stets die stetige Differenzierbarkeit der involvierten Abbildung sowie die Regularität der Funktionalmatrix oder eines Teils davon auf. In 2.3.1 wird ein Überblick über die drei Schwerpunkte des Abschnitts gegeben. Diese werden danach ausführlicher behandelt: das mehrdimensionalen Newtonverfahren in 2.3.2, der Hauptsatz über Umkehrfunktionen in 2.3.3 und der Hauptsatz über implizite Funktionen in 2.3.4.

2.3.1 Übersicht

Inhalt in Kurzfassung: In diesem Überblick über den Abschnitt wird versucht, einen roten Faden durch die Hauptthemen zu legen, die Lösung nichtlinearer Systeme von n Gleichungen in n Variablen mit Hilfe des Newton-Verfahrens, die Interpretation dieser Aufgabe aus Auswertung einer lokalen Umkehrfunktion und die Deutung impliziter Gleichungen in mehreren Variablen als implizite Funktionen.

In der Linearen Algebra (Kapitel 1) haben wir eine Lösungstheorie linearer Gleichungssysteme kennen gelernt, die kaum Wünsche offen lässt. Denn basierend auf theoretischen Überlegungen zeigt das Eliminationsverfahren von Gauß, wie man ein beliebig vorgegebenes System mit klar überschaubarem Rechenaufwand lösen kann. Bei nichtlinearen Gleichungen liegen die Dinge natürlich komplizierter. Dennoch lassen sich, wenn man dem allgemeinen Paradigma der Differentialrechnung, nichtlineare aber (stetig) differenzierbare Funktionen durch ihre linearen Approximationen zu ersetzen, folgt, sehr weit

reichende Analogien zur Linearen Algebra herstellen und daraus wirkungsvolle Methoden entwickeln.

Für den eindimensionalen Fall haben wir in Mathematik 1 mit dem Newton-Verfahren bereits ein äußerst effizientes Näherungsverfahren für die Lösung von Gleichungen der Form $f(x) = 0$ kennen gelernt. Dieses Verfahren beruht auf der Idee, die Funktion f durch ihre lineare Approximation im Sinne der Differentialrechnung – d.h. durch das erste Taylorpolynom t_{f,x_0} an einer Näherungsstelle x_0 für die gesuchte Nullstelle – zu ersetzen. Die resultierende Gleichung $t_{f,x_0}(x) = 0$ kann man leicht lösen. Mit der erhaltenen Lösung x_1 fährt man dann genauso fort. Unter geeigneten Bedingungen konvergiert die resultierende Iterationsfolge sehr rasch gegen eine Nullstelle von f , d.h. gegen die gesuchte Lösung der Gleichung.

Da die in diesem Kapitel entwickelte Differentialrechnung in mehreren Variablen genau diesem Paradigma (lineare Approximation von Funktionen) folgt, liegt es nahe, nach einer Verallgemeinerung des Newton-Verfahrens zu suchen, um Gleichungen der Form $f(\mathbf{x}) = \mathbf{o}$ für Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ näherungsweise zu lösen. Das ist der Inhalt von 2.3.2.

Betrachtet man Gleichungen $f(\mathbf{x}) = \mathbf{y}$ für beliebige (variable) rechte Seiten $\mathbf{y} \in \mathbb{R}^n$, stellt sich die Frage nach einer Umkehrfunktion $f^{(-1)}$ von f , so dass die gesuchte Lösung \mathbf{x} sich aus $\mathbf{x} = f^{(-1)}(\mathbf{y})$ ablesen lässt. Auch wenn so eine Lösung sicher nicht uneingeschränkt möglich ist – zum Beispiel muss f ja nicht injektiv sein – lässt sich unter sehr natürlichen Voraussetzungen die Existenz wenigstens einer lokalen Umkehrfunktion garantieren (siehe 2.3.3).

Eine weitere naheliegende Verallgemeinerung führt in 2.3.4 zu dem für vielfältige Anwendungen nützlichen Hauptsatz über implizite Funktionen. Mit seiner Hilfe lassen sich (natürlich wieder unter geeigneten Voraussetzungen) Systeme von k Gleichungen in $n > k$ Variablen nach gewissen k ausgewählten dieser Variablen auflösen. Das ist so zu verstehen, dass die Lösung als Funktion in den verbleibenden Variablen aufgefasst werden kann. Für den sehr anschaulichen einfachsten der interessanten Fälle, nämlich für $k = 1$ und $n = 2$, lassen sich Gleichungen für Kurven in der Ebene (zum Beispiel die Kreisgleichung $x^2 + y^2 = 1$) zu Funktionen (zum Beispiel zu $y(x) = \pm\sqrt{x^2 - 1}$ oder $x(y) = \pm\sqrt{y^2 - 1}$) umschreiben. Mit Hilfe des Hauptsatzes lässt sich sehr gut verstehen, worauf es dabei ankommt und was darüber hinaus über die resultierenden Funktionen gesagt werden kann.

2.3.2 Mehrdimensionales Newtonverfahren

Inhalt in Kurzfassung: Das aus Mathematik 1 bekannte Newtonverfahren zur Bestimmung von Nullstellen reeller Funktionen mit geeigneten Differenzierbarkeitseigenschaften lässt sich mühelos auf Funktionen mit Definitions- und Wertebereich in \mathbb{R}^n übertragen. Man kann es als Methode zur Lösung von Systemen aus n Gleichungen in n Variablen auffassen.

Wir rekapitulieren aus Mathematik 1 das Newtonverfahren für die Dimension $n = 1$: Gesucht ist eine Nullstelle ξ der reellen Funktion f . Dazu verwendet man Iterationsfolgen

2 Differentialrechnung in mehreren Variablen

$x_{n+1} := T(x_n)$ bezüglich der Funktion

$$T(x) := x - \frac{f(x)}{f'(x)}.$$

Diese Funktion beschreibt, wie man für gegebenes x die Nullstelle $T(x)$ der Tangente an f im Punkt $(x, f(x))$ berechnet. Ist $f'(\xi) \neq 0$ und f zweimal differenzierbar, so gibt es eine Umgebung U von ξ derart, dass für jeden Anfangswert $x_0 \in U$ die Iterationsfolge gegen ξ konvergiert. Damit kann man sehr rasch gute Approximationen für eine Lösung der Gleichung $f(x) = 0$ finden.

Ähnliche Überlegungen kann man anstellen, wenn man es statt mit einer Gleichung $f(x) = 0$ mit einem Gleichungssystem von n Gleichungen

$$f_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, \dots, n,$$

mit (im Gegensatz zu Kapitel 1 eventuell auch nichtlinearen) reellwertigen Funktionen $f_i : D \rightarrow \mathbb{R}$ in n Variablen, d.h. mit $D \subseteq \mathbb{R}^n$ zu tun hat. Man betrachtet dazu die Abbildung

$$f : D \rightarrow \mathbb{R}^n, \quad \mathbf{x} = (x_1, x_2, \dots, x_n) \mapsto \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{pmatrix}.$$

Die Lösung des Gleichungssystems lässt sich als Nullstellensuche für f interpretieren. Analog wie beim eindimensionalen Newtonverfahren betrachtet man für ein beliebiges $\mathbf{x}_0 \in D$ die (inhomogene) lineare Approximation

$$l_{f, \mathbf{x}_0} : \mathbf{x} \mapsto f(\mathbf{x}_0) + f'_{\mathbf{x}_0}(\mathbf{x} - \mathbf{x}_0)$$

und sucht eine Nullstelle \mathbf{x}_1 von l_{f, \mathbf{x}_0} . Ist die lineare Abbildung $f'_{\mathbf{x}_0}$ regulär (invertierbar, was sich mit dem Determinantenkriterium $\det f'_{\mathbf{x}_0} \neq 0$ überprüfen lässt), so lässt sich die Gleichung $l_{f, \mathbf{x}_0}(\mathbf{x}_1) = f(\mathbf{x}_0) + f'_{\mathbf{x}_0}(\mathbf{x}_1 - \mathbf{x}_0) = \mathbf{0}$ nach \mathbf{x}_1 zu

$$\mathbf{x}_1 = \mathbf{x}_0 - (f'_{\mathbf{x}_0})^{(-1)}(f(\mathbf{x}_0))$$

aufösen. Die Newtoniteration $\mathbf{x}_{n+1} := T(\mathbf{x}_n)$ erfolgt also bezüglich der Iterationsfunktion

$$T(\mathbf{x}) := \mathbf{x} - (f'_{\mathbf{x}})^{(-1)}(f(\mathbf{x})).$$

Man beachte, dass diese Formel erwartungsgemäß für $n = 1$ in jene für das eindimensionale Newtonverfahren übergeht. Aus ähnlichen Gründen wie für $n = 1$, wenn auch mit technisch aufwendigerem Beweis lässt sich zeigen:

Satz 2.3.2.1. *Sei $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ stetig differenzierbar und ξ innerer Punkt von D mit $f(\xi) = \mathbf{0}$ und regulärem f'_{ξ} . Dann gibt es eine Umgebung $U \subseteq D$ von ξ derart, dass das Newtonverfahren für jeden Startwert $\mathbf{x}_0 \in U$ gegen ξ konvergiert.*

Übungsaufgabe 144. (P) Finden Sie mit Hilfe des Newtonverfahrens Näherungen für eine oder mehrere Nullstellen der Funktion

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x^2 - y^2 - 2x + 3 \\ 2(x-2)y \end{pmatrix}.$$

Hinweis zur Kontrolle Ihrer Lösung: Fasst man x und y als Real- und Imaginärteil einer komplexen Variablen auf, so lässt sich f mit der komplexen Funktion $z \mapsto z^2 - 2z + 3$ identifizieren.

2.3.3 Lokale Umkehrfunktionen

Inhalt in Kurzfassung: Will man zu einer gegebenen Funktion $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ eine Umkehrfunktion finden, so hat man zu einem gegebenen \mathbf{y} ein \mathbf{x} mit $f(\mathbf{x}) = \mathbf{y}$ zu finden. Ein solches \mathbf{x} ist Nullstelle der Funktion $\mathbf{x} \mapsto f(\mathbf{x}) - \mathbf{y}$ und deshalb unter geeigneten Voraussetzungen zum Beispiel mit Hilfe des Newtonverfahrens aus 2.3.2 zu ermitteln. Untersucht man die Bedingungen sorgfältig, erhält man den Hauptsatz für Umkehrfunktionen, der für stetig differenzierbares f in der Umgebung einer Stelle \mathbf{x}_0 mit regulärer Ableitung $f'_{\mathbf{x}_0}$ eine lokale Umkehrfunktion von f garantiert, die selbst wieder stetig differenzierbar ist.

Jede Gleichung der Form $f(\mathbf{x}) = \mathbf{y}$ mit $f : D \rightarrow \mathbb{R}^n$ und $D \subseteq \mathbb{R}^n$ kann für vorgegebenes \mathbf{y} zu $f(\mathbf{x}) - \mathbf{y} = \mathbf{o}$ umgeschrieben und somit mit dem in 2.3.2 beschriebenen Newtonverfahren behandelt werden. Man erhält damit eine Näherung für die von \mathbf{y} abhängige Lösung $\mathbf{x}_{\mathbf{y}}$. Interessiert man sich auch für die Zuordnung $\mathbf{y} \mapsto \mathbf{x}_{\mathbf{y}}$, so sind darüber hinausgehende Überlegungen nötig, die sich allerdings sehr eng an die bisherigen anlehnen. Wir wollen uns der Problemstellung für allgemeines n über den Fall $n = 1$ annähern.

Viele reelle Funktion $f : D \rightarrow \mathbb{R}$ (wie etwa \cos und \sin) sind mangels Injektivität nicht umkehrbar, sehr wohl aber, wenn man sie auf einen geeigneten Teilbereich $U \rightarrow \mathbb{R}$ einschränkt. Man spricht von **lokaler Umkehrbarkeit** und von einer **lokalen Umkehrfunktion** von f (wie zum Beispiel \arccos von \cos und \arcsin von \sin).

Die Frage, die uns beschäftigt, lautet etwas präziser: Gibt es zu einem Punkt x_0 im Definitionsbereich D einer Funktion f eine Umgebung U von x_0 derart, dass f eingeschränkt auf $D \cap U$ injektiv ist? Für reelle Funktionen auf einem Intervall ist das zum Beispiel schon dann garantiert, wenn f in diesem Bereich streng monoton ist. Für differenzierbares f ist $f'(x_0) \neq 0$ in Verbindung mit stetiger Differenzierbarkeit von f eine hinreichende Bedingung. Denn dann gibt es ein $\varepsilon > 0$, so dass f' auf ganz $U := (x_0 - \varepsilon, x_0 + \varepsilon) \subseteq D$ konstantes Vorzeichen hat, f also auf ganz U streng monoton ist. Weil f dann erst recht stetig ist, folgt aus dem Zwischenwertsatz für stetige Funktionen, dass der Wertebereich $f(U)$ selbst wieder ein (in diesem Fall offenes) Intervall V rund um $f(x_0)$ ist. Diese Überlegung zeigt: Eine stetig differenzierbare reelle Funktion ist in jedem inneren Punkt x_0 ihres Definitionsbereichs lokal umkehrbar, sofern dort $f' \neq 0$ gilt. Die lokale Umkehrfunktion g von f ist dann ebenfalls stetig differenzierbar und gehorcht der Ableitungsregel $g'(f(x)) = f'(x)^{-1}$. Wir fragen, ob es zu dieser Aussage

2 Differentialrechnung in mehreren Variablen

eine mehrdimensionale Verallgemeinerung gibt.

Betrachten wir zunächst lineare Abbildungen $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Für diese ist lokale Bijektivität mit globaler Bijektivität äquivalent, und diese setzt gleiche Dimensionen, also $m = n$ voraus. In diesem Fall, also für lineares $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, ist wiederum $\det f \neq 0$ sowohl notwendig als auch hinreichend.

Die Macht der Differentialrechnung und ihres Paradigmas der Linearisierung zeigt sich darin, was unsere Überlegungen einerseits zu stetig differenzierbaren (eindimensionalen) reellen Funktionen und andererseits zur Umkehrbarkeit (n -dimensionaler) linearer Funktionen nahelegen. Es gilt nämlich der **Hauptsatz über Umkehrfunktionen**:

Satz 2.3.3.1. *Sei $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ stetig differenzierbar, \mathbf{x}_0 ein innerer Punkt von D und die Ableitung (das Differential, die Funktionalmatrix) $f'_{\mathbf{x}_0}$ von f im Punkt \mathbf{x}_0 regulär (d.h. $\det f'_{\mathbf{x}_0} \neq 0$). Dann ist f in \mathbf{x}_0 lokal umkehrbar, explizit:*

Es gibt offene Umgebungen U von \mathbf{x}_0 und V von $f(\mathbf{x}_0)$ derart, dass die Einschränkung f_U von f auf U als Abbildung $f_U : U \rightarrow V$ bijektiv ist.

Überdies ist die Umkehrabbildung $f_U^{(-1)} : V \rightarrow U$ dieser Einschränkung ebenfalls stetig differenzierbar mit

$$\left(f_U^{(-1)}\right)'_{f(\mathbf{x})} = \left(f'_{\mathbf{x}}\right)^{(-1)}$$

für alle $\mathbf{x} \in U$.

Die Aussage von Satz 2.3.3.1 lässt sich griffig auch so zusammenfassen: Stetig differenzierbare Funktionen mit umkehrbarer Ableitung sind selbst lokal umkehrbar, und die Ableitung ihrer Umkehrung ist die Umkehrung ihrer Ableitung. Oder, noch etwas informeller: Funktionen, die sich in umkehrbarer Weise gut linear approximieren lassen, sind selbst lokal umkehrbar, und die lineare Approximation ihrer Umkehrung ist die Umkehrung ihrer linearen Approximation.

Einige Worte zum Beweis von Satz 2.3.3.1: Man kann sich auf eine Umgebung von \mathbf{x}_0 beschränken, wo f' hinreichend wenig von $f'(\mathbf{x}_0)$ abweicht. Die Hauptschwierigkeit besteht darin, zu jedem vorgegebenen und hinreichend nahe bei $f(\mathbf{x}_0)$ gelegenen \mathbf{y} ein \mathbf{x} mit $f(\mathbf{x}) = \mathbf{y}$ zu finden (das wiederum nahe bei \mathbf{x}_0 liegt). Das gelingt, indem man die zu lösende Gleichung $f(\mathbf{x}) = \mathbf{y}$ als Nullstellenaufgabe für $\mathbf{x} \mapsto f(\mathbf{x}) - \mathbf{y}$ auffasst und zeigt, dass das Newtonverfahren (oder eine beweistechnisch leichter zu behandelnde Variante davon) gegen die gesuchte Lösung $\mathbf{x}_{\mathbf{y}}$ konvergiert. Somit ist geklärt, dass sich zu jedem \mathbf{y} ein $\mathbf{x}_{\mathbf{y}}$ finden lässt. Aus der stetigen Differenzierbarkeit lässt sich dann recht leicht ableiten, dass die Zuordnung $\mathbf{y} \mapsto \mathbf{x}_{\mathbf{y}}$ auch die weiteren im Satz behaupteten Eigenschaften der Umkehrfunktion hat.

Übungsaufgabe 145. *(P) Polarkoordinaten lassen sich als Funktionen $f : \mathbb{R}^2 \supseteq D \rightarrow \mathbb{R}^2$, $(r, \alpha) \mapsto (r \cos(\alpha), r \sin(\alpha))$ auffassen. Behandeln Sie diese Funktion unter dem Gesichtspunkt lokaler (oder globaler) Umkehrfunktionen. Stellen Sie analoge Untersuchungen auch für Kugelkoordinaten an.*

2.3.4 Implizite Funktionen

Inhalt in Kurzfassung: Eine naheliegende Folgerung des Hauptsatzes über Umkehrfunktionen ist jener über implizite Funktionen. Enthält ein System aus m Gleichungen $n > m$ Variable, so ist es in vielen Fällen möglich, durch Vorgabe der Werte von $n - m$ dieser Variablen eine eindeutige Lösung für die übrigen m Variablen zu erzwingen. Das entspricht der Situation bei linearen Gleichungssystemen, wo dafür allerdings eine gewisse quadratische Teilmatrix der Systemmatrix regulär ist. Im Hauptsatz über implizite Funktionen bezieht sich die entsprechende Bedingung auf die Funktionalmatrix der durch das Gleichungssystem gegebenen (und im Satz als stetig differenzierbar vorausgesetzten) Abbildung $f : D \rightarrow \mathbb{R}^m$ mit $D \subseteq \mathbb{R}^n$.

Wir beginnen mit einem einfachen Beispiel, an dem bereits das Wichtigste des Themas, um das es jetzt geht, sichtbar wird:

Durch die Kreisgleichung $x^2 + y^2 = 1$ ist eine Teilmenge

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

der Ebene (die Einheitskreislinie) definiert. K ist keine Funktion, weil es zu allen $x \in (-1, 1)$ nicht nur einen, sondern zwei y -Werte mit $(x, y) \in K$ gibt. Schränken wir aber auf den oberen Halbkreis

$$K_0^+ := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1, y \geq 0\},$$

ein, so liegt eine Funktion $y : [-1, 1] \rightarrow \mathbb{R}$ vor. In diesem Bereich können wir leicht die explizite Darstellung

$$y(x) = \sqrt{1 - x^2}$$

geben. Selbstverständlich lassen sich die Rollen von x und y vertauschen und entsprechend $x = x(y) = \sqrt{1 - y^2}$ als Funktion von y auffassen, wenn wir K auf Punkte mit $x \geq 0$ einschränken. Man sagt, die Kreisgleichung lässt sich lokal nach y bzw. nach x auflösen.

Es fällt auf, dass die Punkte $(\pm 1, 0)$ und $(0, \pm 1)$ eine Sonderrolle spielen, weil sie keine Umgebungen haben, in denen die Kreislinie bijektiv ist. An den Punkten $(\pm 1, 0)$ ist keine Auflösung nach y möglich, an den Punkten $(0, \pm 1)$ entsprechend keine nach x . Später, wenn wir dieses einfache Beispiel wieder aufgreifen werden, wird deutlich werden, dass die Nichtauflösbarkeit mit dem Verschwinden der partiellen Ableitung der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \mapsto x^2 + y^2 - 1$, als deren Nullstellenmenge K ja definiert ist, zu tun hat. Um das besser zu verstehen, ist aber ein größerer Kontext wünschenswert, der sich automatisch ergibt, wenn wir uns der Verallgemeinerung auf höhere Dimensionen zuwenden. Dann haben wir es mit der folgenden Situation zu tun.

Gegeben seien eine oder, allgemeiner, m Gleichungen $f_i(x_1, \dots, x_n) = 0$, $i = 1, \dots, m$, mit reellwertigen Funktionen $f_i : D \rightarrow \mathbb{R}$. Ähnlich wie beim Newtonverfahren, nun

2 Differentialrechnung in mehreren Variablen

jedoch mit $n > m$ (statt $m = n$) betrachten wir die Funktion

$$f : D \rightarrow \mathbb{R}^m, \quad \mathbf{x} = (x_1, x_2, \dots, x_n) \mapsto \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

und suchen ihre Nullstellen. Allerdings wollen wir die Werte für gewisse Variablen (etwa x_1, \dots, x_k) vorgeben und daraus die übrigen (wir bezeichnen sie ab nun vorzugsweise mit $y_1 = x_{k+1}, \dots, y_m = x_n$) eindeutig bestimmen.

Denken wir an lineare Gleichungen, so sehen wir, dass in der Regel $k = n - m$ der sinnvolle Wert ist, weil dann aus den m gegebenen Gleichungen bei Vorgabe der ersten $n - m$ Werte x_1, \dots, x_{n-m} gerade die m Unbekannten y_1, \dots, y_m berechnet werden können, sofern eine gewisse Matrix vollen Rang hat. Und zwar geht es um einen Teil der Koeffizientenmatrix A von f (natürlich bezüglich der kanonischen Basis). Zur Präzisierung schreiben wir A als $A = A_1|A_2$ an mit einer $m \times (n - m)$ -Matrix A_1 und einer quadratischen $m \times m$ -Matrix A_2 . Kürzen wir außerdem zu $\mathbf{x} = (x_1, \dots, x_{n-m})$ und $\mathbf{y} = (y_1, \dots, y_m)$ ab, so wird das Systems zu $A_1\mathbf{x} + A_2\mathbf{y} = \mathbf{o}$, woraus wir

$$\mathbf{y} = -A_2^{-1}A_1\mathbf{x}$$

ablesen, sofern A_2 als invertierbar vorausgesetzt wird. Lesen wir diese Gleichung als Zuordnung $\mathbf{x} \mapsto \mathbf{y}$, so haben wir für den linearen Fall unser Ziel erreicht, nämlich \mathbf{y} als Funktion von \mathbf{x} darzustellen. Weil wir aus der Linearen Algebra wissen, wie Matrizen zu invertieren sind, lässt sich für invertierbares A_2 die implizite Darstellung durch das ursprünglichen System also sogar in algorithmisch befriedigender Weise in eine explizite Darstellung von \mathbf{y} als Funktion von \mathbf{x} umwandeln.

Für die Übertragung auf den nichtlinearen Fall legt der Satz über Umkehrfunktionen (2.3.3.1) nahe, dass wir uns auch bei impliziten Funktionen mit lokalen Lösungen bescheiden müssen, die an einem bestimmten Punkt ansetzen. Ähnlich wie beim linearen Fall schreiben wir Vektoren $(x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n)$ auch als (\mathbf{x}, \mathbf{y}) mit $\mathbf{x} = (x_1, x_2, \dots, x_{n-m}) \in \mathbb{R}^{n-m}$ und $\mathbf{y} = (y_1, \dots, y_m) = (x_{m+1}, \dots, x_n) \in \mathbb{R}^m$, also $y_i = x_{m+i}$ für $i = 1, \dots, m$. Ein Vektor in \mathbb{R}^n ist dann gleichzeitig ein Paar $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$. Für eine Lösung (\mathbf{x}, \mathbf{y}) des Systems gilt in dieser Notation $f(\mathbf{x}, \mathbf{y}) = \mathbf{o}$.

Unterwerfen wir alle unsere Erkenntnisse über Umkehrfunktionen im allgemeinen Fall und unsere obigen Überlegungen zu impliziten Funktionen im linearen Fall dem generellen Linearisierungsparadigma der Differentialrechnung, so lässt sich die korrekte Formulierung des **Hauptsatzes über implizite Funktionen** (wir verwenden dabei die oben eingeführten Schreibweisen) geradezu erraten. Er lautet:

Satz 2.3.4.1. *Sei $D \subseteq \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}^m$ mit $m < n$ stetig differenzierbar, $\mathbf{x}_0 \in \mathbb{R}^{n-m}$, $\mathbf{y}_0 \in \mathbb{R}^m$ mit $(\mathbf{x}_0, \mathbf{y}_0) \in D$ und $f(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{o}$. Sei $A = A_1|A_2$ die Matrixdarstellung von $f'_{(\mathbf{x}_0, \mathbf{y}_0)}$ (bezüglich der kanonischen Basis) und A_2 regulär (d.h. $\det A_2 \neq 0$). Dann*

definiert f in einer Umgebung von $(\mathbf{x}_0, \mathbf{y}_0)$ implizit eine stetig differenzierbare Funktion $\mathbf{x} \mapsto \mathbf{y}$, genauer:

Es gibt Umgebungen $U \subseteq \mathbb{R}^{n-m}$ von \mathbf{x}_0 und $V \subseteq \mathbb{R}^m$ von \mathbf{y}_0 mit folgenden Eigenschaften: Zu jedem $\mathbf{x} \in U$ gibt es genau ein \mathbf{y} in V mit $f(\mathbf{x}, \mathbf{y}) = \mathbf{o}$. Schreiben wir für dieses \mathbf{y} auch $g(\mathbf{x})$, so ist die Funktion $g : U \rightarrow V$ stetig differenzierbar. Sei außerdem $B = B_1 | B_2$ die Matrixdarstellung von $f'_{(\mathbf{x}, \mathbf{y})}$ für irgendein $(\mathbf{x}, \mathbf{y}) \in U \times V$. Dann ist B_2 regulär, und die Ableitung g'_x an einer Stelle $\mathbf{x} \in U$ wird durch die Matrix $-B_2^{-1} B_1$ dargestellt.

Ein formaler Beweis dieses Satzes gelingt durch Anwendung des Satzes 2.3.3.1 auf eine geschickt zu wählende Funktion, die sich aus f ergibt. Sonst verwendet der Beweis aber keine nennenswerten zusätzlichen Ideen, weshalb wir ihn in die folgende Übungsaufgabe mit Hinweis auslagern.

Übungsaufgabe 146. (E) Man beweise Satz 2.3.4.1. Hinweis: Man verwende die Funktion $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}, f(\mathbf{x}, \mathbf{y}))$, und wende darauf in geeigneter Weise Satz 2.3.3.1 an.

Es sei ausdrücklich auf die Analogie zur Lösung eines linearen Systems aus m Gleichungen in n Variablen hingewiesen. Wenn man mit Hilfe des Eliminationsverfahrens von Gauß eine Parameterdarstellung der Lösungsmenge ermittelt, so bedeutet das: Gewisse Variablen (in der Notation von Satz 2.3.4.1 wären es x_1, \dots, x_{n-m}) werden mit Parametern identifiziert, so dass sich jede Wahl der Parameter (d.h. jedes $\mathbf{x} = (x_1, \dots, x_{n-m}) \in \mathbb{R}^{n-m}$) zu genau einer Lösung $(x_1, \dots, x_{n-m}, x_{n-m+1}, \dots, x_n)$ des Gleichungssystems vervollständigen lässt. Die dadurch definierte Zuordnung $\mathbf{x} \mapsto \mathbf{y} := (x_{n-m+1}, \dots, x_n)$ ist die implizite Funktion g aus Satz 2.3.4.1. Man beachte, dass lediglich aus notationellen Gründen bei den linearen Gleichungssystemen in 1.5.5 die Reihenfolge der Variablen so gewählt war, dass sich die ersten m Variablen als Funktion der letzten $n - m$ Variablen aufgefasst wurden und nicht umgekehrt, die letzten m aus den ersten $n - m$.

Obwohl die Aussage des Satzes über implizite Funktionen also äußerst natürlich ist, mag sie auf den ersten Blick einigermaßen kompliziert anmuten. Deshalb wollen wir nun nochmals das einfache Beispiel der Kreislinie aufgreifen und studieren, was der Satz und die darin vorkommende Regularitätsbedingung an A_2 in diesem Fall besagt.

In der Notation von Satz 2.3.4.1 ist jetzt $n = 2$, $m = 1$, $\mathbf{x} = x$, $\mathbf{y} = y$ und $f(x, y) = x^2 + y^2 - 1$. Die partiellen Ableitungen von f sind $\frac{\partial f}{\partial x} = 2x$ und $\frac{\partial f}{\partial y} = 2y$. Folglich ist $A = \text{grad } f = (2x \ 2y) = A_1 | A_2$ mit den beiden 1×1 -Matrizen $A_1 = (2x)$ und $A_2 = (2y)$. Wollen wir nach y auflösen, ist Invertierbarkeit von A_2 zu garantieren, was bei Skalaren schlicht Nichtverschwinden, also $y \neq 0$ bedeutet. Nur zwei Lösungen (x_0, y_0) der Kreisgleichung haben y -Koordinate 0, nämlich die Punkte $(1, 0)$ und $(-1, 0)$. Genau diese Punkte haben wir schon eingangs als die problematischen identifiziert. Analoges gilt für Auflösung nach x und die problematischen Punkte $(0, 1)$ und $(0, -1)$. Eine genauere Analyse ist Teil der folgenden Übungsaufgabe.

Übungsaufgabe 147. (E)

2 Differentialrechnung in mehreren Variablen

1. Es ist durchaus möglich, dass die Regularitätsbedingung an A_2 in Satz 2.3.4.1 verletzt ist, aber trotzdem Auflösbarkeit vorliegt. Man illustriere das anhand der Gleichung $f(x, y) = x^3 - y^3 = 0$ am Punkt $(0, 0)$.
2. Will man argumentieren, dass die Kreisgleichung zum Beispiel am Punkt $(1, 0)$ nicht lokal nach y auflösbar ist, genügt es – wie Teil (a) zeigt – nicht, das Verschwinden der partiellen Ableitung nach y am Punkt $(1, 0)$ festzustellen, sondern es ist folgendes zu zeigen: Zu beliebigen Umgebungen U der ersten Koordinate 1 und V der zweiten Koordinate 0 gibt es $x \in U$, sodass $f(x, y) = 0$ für mehr als ein $y \in V$ gilt. Erklären Sie, warum das tatsächlich der Fall ist.

Das behandelte Beispiel der Kreislinie ist eine sehr einfache Anwendung von Satz 2.3.4.1, nicht zuletzt deshalb, weil die Dimensionen $m = 1$ und $n = 2$ die kleinst möglichen sind, für die der Satz von Interesse ist.

Übungsaufgabe 148. (E) Behandeln Sie Gleichungen $f(x, y) = 0$ für interessante stetig differenzierbare Funktionen $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^2$, Ihrer Wahl im Lichte von Satz 2.3.4.1. Skizzieren Sie insbesondere die durch f definierte ebene Kurve.

Ein Beispiel für $n = 3$ und $m = 1$ ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 149. (E) Untersuchen Sie die durch die Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) := x^2 + y^2 + z^2 - 1$, gegebene Gleichung für die Sphäre der Kugel unter dem Gesichtspunkt von Satz 2.3.4.1.

Übungsaufgabe 150. (E) Finden Sie ein interessantes Beispiel zu Satz 2.3.4.1 mit $n = 3$ und $m = 2$.

Auch für $n > 3$ gibt es wichtige Anwendungen, wenn auch nicht mehr unbedingt geometrischer Natur. Ein Beispiel bilden Extremwertaufgaben mit Nebenbedingungen und die Methode der sogenannten Lagrange-Multiplikatoren, siehe 2.4.4.

Weitere interessante Anwendung impliziter Funktionen sind auch gewisse Differentialgleichungen, deren Lösung sich nur implizit (also als ebene Kurve) angeben lassen, siehe 3.3.2.

2.4 Extremwertbestimmung in mehreren Variablen

Analog zu reellen Funktionen in einer Variablen spielen Extremwertaufgaben auch für Funktionen in n Variablen, also $f : D \rightarrow \mathbb{R}$, $\mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x})$ mit $D \subseteq \mathbb{R}^n$ (z.B. Optimierungsaufgaben) eine wichtige Rolle, insbesondere in Anwendungen. In 2.4.1 werden Begriffe und Problemstellung auf diesen allgemeineren Fall übertragen sowie Komplikationen besprochen, die sich dabei ergeben. Ist f differenzierbar und \mathbf{x}_0 ein lokales Extremum von f im Inneren von D , so muss (wie bei einer Variablen) an der Stelle \mathbf{x}_0 die Ableitung, das heißt: alle partiellen Ableitungen von f , in \mathbf{x}_0 verschwinden. Wie auch im Fall $n = 1$ ist diese Bedingung aber nicht hinreichend. Feiner Analysen involvieren die zweiten partiellen Ableitungen und führen in 2.4.2 zum Hauptminorenkriterium.

Die Hintergründe dazu werden in 2.4.3 vertieft. Sie involvieren den Satz von Taylor in mehreren Variablen, quadratische Formen und ihre Definitheit sowie die Hauptachsentransformation. Im Fall von Extremwertaufgaben unter Nebenbedingungen bildet die Methode der Lagrangeschen Multiplikatoren ein unverzichtbares methodischen Hilfsmittel und wird in 2.4.4, mit denen Nebenbedingungen behandelt werden. Abschnitt 2.4.5 bringt einen resümierenden Überblick über das Zusammenspiel der eingesetzten Hilfsmittel und leitet daraus eine systematische Strategie ab. Der Abschnitt schließt in 2.4.6 mit Regressionsgerade und Ausgleichsrechnung als Anwendungen der entwickelten Methoden.

2.4.1 Problemstellung und Lösungsstrategie im Überblick

Inhalt in Kurzfassung: Begriffe und Problemstellung von Extremwertaufgaben können vom Fall einer auf den von n Variablen übertragen werden. Auch die Lösungsstrategie folgt ähnlichen Mustern. Zunächst geht es insbesondere darum, wann das Verschwinden der ersten Ableitung (des Gradienten) eine notwendige Bedingung für das Vorliegen eines Extremums ist.

Der grundlegende Begriff des Extremums sowie seine Varianten und Verwandten lassen sich unmittelbar vom eindimensionalen auf den mehrdimensionalen Fall übertragen:

Definition 2.4.1.1. Sei $f : D \rightarrow \mathbb{R}$, $\mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x})$ mit $D \subseteq \mathbb{R}^n$. Dann heißt $\mathbf{x}_0 \in D$ (globale) **Extremstelle** und $f(\mathbf{x}_0)$ **Extremum** von f , wenn eine der folgenden beiden Aussagen zutrifft:

$$\text{Erste Aussage: } f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad \text{für alle } \mathbf{x} \in D.$$

In diesem Fall heißt \mathbf{x}_0 (globale) **Maximumsstelle** und $f(\mathbf{x}_0)$ (globales) **Maximum** von f .

Oder:

$$\text{Zweite Aussage: } f(\mathbf{x}_0) \leq f(\mathbf{x}) \quad \text{für alle } \mathbf{x} \in D.$$

In diesem Fall heißt \mathbf{x}_0 (globale) **Minimumsstelle** und $f(\mathbf{x}_0)$ (globales) **Minimum** von f .

Gibt es eine Umgebung U von \mathbf{x}_0 , so dass \mathbf{x}_0 Extremstelle der Einschränkung von f auf $U \cap D$ ist, so heißt \mathbf{x}_0 eine **lokale Extremstelle** und $f(\mathbf{x}_0)$ ein **lokales Extremum** von f . Ist \mathbf{x}_0 eine Maximumsstelle der Einschränkung auf $U \cap D$, so spricht man auch von einer **lokalen Maximumsstelle** bzw. von $f(\mathbf{x}_0)$ als einem **lokalen Maximum** von f . Ist \mathbf{x}_0 eine Minimumsstelle der Einschränkung von f auf $U \cap D$, so spricht man auch von einer **lokalen Minimumsstelle** bzw. von $f(\mathbf{x}_0)$ als einem **lokalen Minimum** von f .

Aus Mathematik 1 wissen wir, dass Extrema einer differenzierbaren Funktion an inneren Punkten des Definitionsbereichs nur dort auftreten können, wo die Ableitung den Wert 0 annimmt. Die entsprechende Aussage gilt auch im Mehrdimensionalen. Anschaulich auf ein Funktionsgebirge bezogen entspricht das der Tatsache, dass Tangentialebenen auf Berggipfeln stets waagrecht liegen. Die mathematische Satz dazu lautet wie folgt.

Satz 2.4.1.2. Sei $f : D \rightarrow \mathbb{R}$, $\mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x})$ mit $D \subseteq \mathbb{R}^n$ differenzierbar und \mathbf{x}_0 ein innerer Punkt von D und Extremstelle von f . Dann gilt $f'_{\mathbf{x}_0} = \text{grad}_{\mathbf{x}} f = \mathbf{o}$, d.h.

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = \frac{\partial f}{\partial x_2}(\mathbf{x}_0) = \dots = \frac{\partial f}{\partial x_n}(\mathbf{x}_0) = 0.$$

Beweis. Wir spielen die Situation auf den eindimensionalen Fall zurück. Für ein beliebiges $i \in \{1, \dots, n\}$ müssen wir $\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = 0$ zeigen. Sei $\mathbf{x}_0 = (y_1, \dots, y_n)$, außerdem D_i die Menge aller $t \in \mathbb{R}$ mit $(y_1, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n) \in D$. Dann ist y_i eine lokale Extremstelle der differenzierbaren reellen Funktion

$$f_i : D_i \rightarrow \mathbb{R}, \quad t \mapsto f(y_1, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n).$$

im Inneren von D_i . Nach dem entsprechenden Satz aus Mathematik 1 über das Verschwinden der Ableitung an Extremstellen im Inneren folgt

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) = f'_i(y_i) = 0.$$

□

Die Bedingung $\text{grad}_{\mathbf{x}} f = \mathbf{o}$ besagt für eine differenzierbare Funktion f , dass sich der Wert von f bei einer kleinen Bewegung von \mathbf{x} weg (fast) nicht ändert, egal in welche Richtung man sich bewegt. Deshalb nennt man so ein \mathbf{x} auch einen **stationären Punkt** von f .

Schon im Eindimensionalen war die Extremwertbestimmung mit dem Aufsuchen sämtlicher stationären Punkte noch lange nicht abgeschlossen. Im Höherdimensionalen gilt das umso mehr. Selbst für eine beliebig oft differenzierbare Funktion $f : D \rightarrow \mathbb{R}$ in n Variablen sind nämlich die folgenden möglichen Komplikationen zu bedenken, die über jene für $n = 1$ teilweise beträchtlich hinaus gehen:

1. Das Verschwinden der Ableitung ist keine hinreichende Bedingung für das Vorliegen von Extremstellen. Diese Einschränkung hat bei n Variablen wesentlich größeres Gewicht als für $n = 1$ und wird uns noch ausführlich beschäftigen, siehe vor allem 2.4.3.
2. Ob es sich bei einem Extremum um ein Maximum oder Minimum handelt, konnte im Eindimensionalen in der Regel mit Hilfe der zweiten oder, in Spezialfällen, auch höheren Ableitungen ermittelt werden. Diese Methode fußte auf dem Satz von Taylor. Auch für Funktionen in n -Variablen gibt es eine entsprechende Verallgemeinerung dieses zentralen Satzes der Differentialrechnung. Deshalb ist eine Analyse der Qualität eines Extremums mit Hilfe höherer Ableitungen auch im Mehrdimensionalen grundsätzlich möglich. Allerdings werden die Dinge so kompliziert, dass man sich in der Praxis mit der Untersuchung der zweiten Ableitung, genauer mit der Definitheit der Hesseschen Matrix, zufrieden gibt. Immerhin versteht man die Situation aufgrund derselben mathematischen Hintergründe wie die bereits im ersten Punkt angesprochene Frage nach dem Hinreichen der Bedingung $f'(\mathbf{x}_0) = \mathbf{o}$. Deshalb wird dieses Thema gleichfalls durch 2.4.3 erfasst.

3. Randpunkte des Definitionsbereichs D werden durch Satz 2.4.1.2 nicht erfasst, müssen daher extra untersucht werden. Auch hier stellt sich der n -dimensionale Fall als substantiell komplizierter dar. In der Regel ist nämlich D zusammenhängend. Das bedeutet für $n = 1$, dass es nur höchstens zwei Randpunkte extra zu untersuchen gilt. Für $n \geq 2$ enthält der Rand ∂D von D aber meist unendlich viele Punkte. Man muss deshalb nach anderen Methoden Ausschau halten. Lässt sich ∂D (stückweise) als Lösungsmenge von einer oder mehreren Gleichungen schreiben, so lassen sich diese Gleichungen als sogenannte Nebenbedingungen auffassen, für welche die Methode der Lagrange-Multiplikatoren zur Verfügung steht, siehe ??.

Die Strategie, die durch diese Sachlage für die Lösung von Extremwertaufgaben nahegelegt wird, werden wir in 2.4.5 behandeln. Davor wollen wir jedoch die für die Umsetzung dieser Strategie notwendigen Methoden entwickeln.

2.4.2 Die Rolle der zweiten Ableitungen

Inhalt in Kurzfassung: In Analogie zur Rolle der zweiten Ableitung bei der Extremwertbestimmung für Funktionen in einer Variablen gilt im Fall mehrere Variablen das sogenannte Hauptminorenkriterium. Darin spielt die Hesse-Matrix, bestehend aus sämtlichen partiellen Ableitungen zweiter Ordnung die zentrale Rolle.

Die erste der drei in 2.4.1 beschriebenen Komplikationen, die bei mehreren Variablen hinzukommen, lässt sich am einfachsten durch die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := x^2 - y^2$$

und ihr Verhalten in der Nähe des Koordinatenursprungs $\mathbf{x} = (0, 0)$ illustrieren. Der Gradient ist $\text{grad}_{\mathbf{x}} f = \text{grad}_{(x,y)} f = (f_x, f_y) = (2x, 2y)$, also ist $(0, 0) = \mathbf{o}$ der einzige stationäre Punkt von f . Weil f auf ganz \mathbb{R}^2 differenzierbar ist, kommt nach Satz 2.4.1.2 also nur der Koordinatenursprung \mathbf{o} als lokale Extremstelle von f in Frage. Die zweiten partiellen Ableitungen $f_{xx}(x, y) = 2$ und $f_{yy}(x, y) = -2$ sind dort (wie überall auf \mathbb{R}^2) von 0 verschieden. Trotzdem gibt es in jeder Umgebung U von \mathbf{o} Funktionswerte sowohl kleiner als auch größer $f(\mathbf{o}) = 0$. Zum Beispiel gilt für $\mathbf{x}_1 := (0, \varepsilon)$ und $\mathbf{x}_2 := (\varepsilon, 0)$, ($\varepsilon \neq 0$ beliebig nahe 0)

$$f(\mathbf{x}_1) = -\varepsilon^2 < 0 = f(\mathbf{o}) < \varepsilon^2 = f(\mathbf{x}_2).$$

Also ist \mathbf{o} keine lokale Extremstelle von f . Um das Vorliegen einer Extremstelle an einer stationären Stelle zu garantieren, ist also eine stärkere Bedingung erforderlich als das Nichtverschwinden der zweiten Ableitung.

Bevor wir die gesuchte Bedingung besprechen, wollen wir uns die Situation geometrisch veranschaulichen: Entlang der x -Achse verhält sich f wie eine nach oben offene Parabel, entlang der y -Achse wie eine nach unten offene. Man kann sich so einen Punkt \mathbf{x} vorstellen wie einen Pass (Sattel) in der Landschaft des Funktionsgebirges, weshalb man

in so einer Situation auch von einem **Sattelpunkt** spricht. Entlang der x -Achse verläuft der Bergkamm mit \mathbf{x} als am tiefsten gelegenen Punkt, entlang der y -Achse die Passstraße, die in \mathbf{x} ihren höchsten Punkt hat. Diesen Überlegungen entnehmen wir auch, dass für ein lokales Maximum an der Stelle \mathbf{x} sowohl $f_{xx}(\mathbf{x}) \leq 0$ also auch $f_{yy}(\mathbf{x}) \leq 0$ gelten muss, für ein lokales Minimum sowohl $f_{xx}(\mathbf{x}) \geq 0$ also auch $f_{yy}(\mathbf{x}) \geq 0$. Für eine lokale Extremstelle \mathbf{x} ist also $f_{xx}(\mathbf{x}) \cdot f_{yy}(\mathbf{x}) \geq 0$ notwendig. Hinreichend ist aber auch das nicht, wie das folgende Beispiel zeigt:

Für die Funktion $f(x, y) := x^2 + 4xy + y^2$ haben die ersten partiellen Ableitungen $f_x(x, y) = 2x + 4y$, $f_y(x, y) = 4x + 2y$ in $\mathbf{o} = (0, 0)$ den Wert 0. Außerdem sind die partiellen Ableitungen zweiter Ordnung $f_{xx}(x, y) = 2$, $f_{xy}(x, y) = f_{yx}(x, y) = 4$ und $f_{yy}(x, y) = 2$ alle positiv. Trotzdem liegt bei $\mathbf{x} = (0, 0)$ kein lokales Minimum vor. Denn für alle $\mathbf{x}_\varepsilon := (\varepsilon, -\varepsilon)$ mit $\varepsilon \neq 0$ gilt:

$$f(\mathbf{x}_\varepsilon) = f(\varepsilon, -\varepsilon) = \varepsilon^2 - 4\varepsilon^2 + \varepsilon^2 = -2\varepsilon^2 < 0 = f(0, 0).$$

Weil jede Umgebung U von $(0, 0)$ solche \mathbf{x}_ε mit negativem Funktionswert ebenso enthält wie solche mit positivem (z.B. $(\varepsilon, 0)$) hat auch dieses f in \mathbf{o} keine Extremstelle. Um das Verhalten von f besser zu verstehen, stellen wir folgende Überlegungen an. Entlang der beiden Koordinatenachsen gilt $f(x, 0) = x^2$ und $f(0, y) = y^2$, wie bei einem Paraboloid der Gestalt $x^2 + y^2$. Allerdings ist der Verhalten von f insgesamt komplizierter. Man versteht es besser, wenn man zu Polarkoordinaten übergeht: $x = r \cos(\alpha)$ und $y = r \sin(\alpha)$ mit $r \geq 0$ und $\alpha \in [0, 2\pi)$. Dann erhält man unter Verwendung der Beziehung $\cos^2 + \sin^2 = 1$ und des Additionstheorems für den Sinus für f nämlich die Darstellung

$$f(x, y) = f(r, \alpha) = r^2(\cos^2(\alpha) + 4 \cos(\alpha) \sin(\alpha) + \sin^2(\alpha)) = r^2(1 + 2 \sin(2\alpha)).$$

Für festes α (d.h. bei Einschränkung auf die Gerade bestehend aus allen Punkten $(r \cos(\alpha), r \sin(\alpha))$, r variabel) ist f also in der Regel eine quadratische Parabel, die für $2 \sin(2\alpha) < -1$ (d.h. für $\frac{7}{12}\pi < \alpha < \frac{11}{12}\pi$ und $\frac{19}{12}\pi < \alpha < \frac{23}{12}\pi$) nach unten, für $2 \sin(2\alpha) < -1$ nach oben offen ist. Nur für $\alpha \in \{\frac{7}{12}, \frac{11}{12}, \frac{19}{12}, \frac{23}{12}\}$ ist f die Nullfunktion. Der oben betrachtete Fall mit $(\varepsilon, -\varepsilon)$ entspricht den Winkeln $\alpha = \frac{3}{4}\pi$ und $\alpha = \frac{7}{4}\pi$ mit nach unten offener Parabel.

Übungsaufgabe 151. (P) Man skizziere diese Funktion f in einer Umgebung von $(0, 0)$ und deute die oben durchgeführten Rechnungen geometrisch.

Wie dieses Beispiel im Vergleich mit anderen zeigt, enthalten also die beiden zweiten partiellen Ableitungen f_{xx} und f_{yy} nicht genügend Information, um über Maximum, Minimum oder Sattelpunkt zu entscheiden. Sehr wohl enthält man aber ein sehr brauchbares Kriterium, wenn man auch die gemischten Ableitungen f_{xy} und f_{yx} heranzieht, die nach dem Satz von Schwarz 2.2.4.3 im Falle zweimaliger stetiger Differenzierbarkeit von f (die hier stillschweigend vorausgesetzt sei) übereinstimmen. Und zwar spielt die Determinante

$$\det H = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - f_{xy}^2$$

2.4 Extremwertbestimmung in mehreren Variablen

der Hesseschen Matrix eine wesentliche Rolle. Sind f_{xx} und $\det H$ beide positiv, so liegt ein Minimum vor, im Fall $f_{xx} < 0$ und $\det H > 0$ ein Maximum. Ist $\det H < 0$, so liegt ein Sattelpunkt vor. In obigem Beispiel ist tatsächlich $\det H = 2 \cdot 2 - 4^2 = -12 < 0$.

Bei diesem Zusammenhang handelt es sich um einen Spezialfall des viel allgemeineren, trotzdem aber recht einprägsamen **Hauptminorenkriterium**:

Satz 2.4.2.1. Sei $f : D \rightarrow \mathbb{R}$ auf $D \subseteq \mathbb{R}^n$ zweimal stetig differenzierbar, außerdem \mathbf{x}_0 ein innerer Punkt von D und gleichzeitig ein stationärer Punkt von f (d.h. $\text{grad}_{\mathbf{x}_0} f = \mathbf{0}$). Sei außerdem

$$H(f, \mathbf{x}_0) = (f_{x_i x_j}(\mathbf{x}_0)) = \begin{pmatrix} f_{x_1 x_1}(\mathbf{x}_0) & f_{x_1 x_2}(\mathbf{x}_0) & \cdots & f_{x_1 x_n}(\mathbf{x}_0) \\ f_{x_2 x_1}(\mathbf{x}_0) & f_{x_2 x_2}(\mathbf{x}_0) & \cdots & f_{x_2 x_n}(\mathbf{x}_0) \\ \cdots & \cdots & \cdots & \cdots \\ f_{x_n x_1}(\mathbf{x}_0) & f_{x_n x_2}(\mathbf{x}_0) & \cdots & f_{x_n x_n}(\mathbf{x}_0) \end{pmatrix}$$

die Hesse-Matrix von f an der Stelle \mathbf{x}_0 . Mit H_i bezeichnen wir jene $i \times i$ -Teilmatrix von $H(f, \mathbf{x}_0)$, die aus den ersten i Zeilen und Spalten besteht, mit $d_i := \det H_i$ ihre Determinante. Dann gilt:

1. Ist $d_i > 0$ für alle $i = 1, \dots, n$, so hat f in \mathbf{x}_0 ein lokales Minimum.
2. Haben die d_i , negativ beginnend, abwechselndes Vorzeichen, d.h.: $(-1)^i d_i > 0$ für $i = 1, \dots, n$, so hat f in \mathbf{x}_0 ein lokales Maximum.
3. Sind alle $d_i \neq 0$, aber nicht von einem der beiden erwähnten Typen, dann hat f in \mathbf{x}_0 kein lokales Extremum.

Man beachte, dass für den Fall $d_i = 0$ keine allgemeinen Aussagen gemacht werden können. Das Phänomen ist vom Fall $n = 1$ aus Mathematik 1 bekannt: Verschwindet nicht nur die erste, sondern auch die zweite Ableitung, so muss kein Extremum vorliegen, wie das Beispiel $f(x) = x^3$ an der Stelle $x = 0$ in typischer Weise zeigt. In so einer Situation kann man höhere Ableitungen zu Rate ziehen, wie das in Mathematik 1 auch getan wurde. Für $n > 1$ wird die Bestimmung höherer Ableitungen aber sehr komplex. Man versucht deshalb meist, mit Überlegungen, die ad hoc an die konkret vorgegebene Funktion f angepasst sind, zum Ziel zu gelangen. Außerdem handelt es sich um einen relativ selten auftretenden Ausnahmefall, weshalb die technisch aufwendige allgemeine Theorie dazu auch in der Lehrbuchliteratur selten behandelt wird.

Der Beweis von Satz 2.4.2.1 würde hier zu weit führen. Trotzdem mögen die folgenden Erklärungen zum Hintergrund das Verständnis vertiefen und dabei helfen, sich dieses Hauptminorenkriterium einzuprägen.

2.4.3 Taylor, quadratische Formen und Hauptachsentransformation

Inhalt in Kurzfassung: Auch für Funktionen in mehreren Variablen gilt ein Analogon zum Satz von Taylor, und auch dort hilft dieser Satz wesentlich für das Verständnis von Extremstellen. Unter den analogen Bedingungen wie im Fall einer Variablen lassen sich

2 Differentialrechnung in mehreren Variablen

nämlich auch Funktionen in mehreren Variablen so gut durch ihr Taylorpolynom zweiten Grades approximieren, dass die Analyse von Extremwertaufgaben auf das Studium solcher Polynome und der durch sie induzierten quadratischen Formen hinausläuft. Daraus ergibt sich ein besseres geometrisches Verständnis dafür, worum es im Hauptminorenkriterium geht. Unterstützt wird dieses Verständnis auch noch durch die Hauptachsentransformation.

Beim Satz von Taylor für Funktionen in n Variablen geht wie schon für $n = 1$ um die Approximation einer hinreichend oft differenzierbaren Funktion f durch Polynome. Der Grad dieser Polynome entspricht dabei der Ordnung, bis zu der die partiellen Ableitungen verwendet werden. Zuerst ist also zu klären, was wir unter einem Polynom eines gewissen Grades in n Variablen verstehen wollen.

Definition 2.4.3.1. *Unter einem (reellen) **Polynom** (oder genauer: einer **Polynomfunktion**) p in n Variablen x_1, \dots, x_n versteht man eine Funktion, die sich in der Form*

$$p(x_1, \dots, x_n) = \sum_{i_1, \dots, i_n=0}^m a_{i_1, \dots, i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$$

mit einem $m \in \mathbb{N}$ und Zahlen $a_{i_1, \dots, i_n} \in \mathbb{R}$, den sogenannten **Koeffizienten** von p , darstellen lässt. Die größte der Zahlen $k = i_1 + \dots + i_n$ mit $a_{i_1, \dots, i_n} \neq 0$ heißt der **Grad** von p .

Analog zum Fall $n = 1$ gibt es wieder einen wichtigen Zusammenhang zwischen den Koeffizienten eines Polynoms p und den Werten der höheren Ableitungen von p an der Stelle $\mathbf{x}_0 = \mathbf{o}$, nämlich:

Proposition 2.4.3.2. *Für ein Polynom p wie in Definition 2.4.3.1 und $k = i_1 + \dots + i_n$ gilt:*

$$\frac{\partial^k f}{(\partial x_1)^{i_1} \dots (\partial x_n)^{i_n}}(\mathbf{o}) = (i_1!) \cdot (i_2!) \cdot \dots \cdot (i_n!) a_{i_1, \dots, i_n}$$

Analog gilt für eine sogenannte *Entwicklung*

$$p(x_1, \dots, x_n) = \sum_{i_1, \dots, i_n=0}^m b_{i_1, \dots, i_n} (x_1 - y_1)^{i_1} (x_2 - y_2)^{i_2} \dots (x_n - y_n)^{i_n}$$

von p um eine beliebige Stelle $\mathbf{x}_0 = (y_1, \dots, y_n) \in \mathbb{R}^n$:

$$\frac{\partial^k f}{(\partial x_1)^{i_1} \dots (\partial x_n)^{i_n}}(\mathbf{x}_0) = (i_1!) \cdot (i_2!) \cdot \dots \cdot (i_n!) b_{i_1, \dots, i_n}$$

Übungsaufgabe 152. (E) *Beweisen Sie Proposition 2.4.3.2. Hinweis: Orientieren Sie sich dabei an der entsprechenden Aussage für $n = 1$ aus Mathematik 1.*

Auch ein Polynom in mehreren Variablen ist daher eindeutig bestimmt durch die Werte seiner höheren partiellen Ableitungen an einer bestimmten Stelle \mathbf{x}_0 . Umgekehrt kann

2.4 Extremwertbestimmung in mehreren Variablen

man einer k -mal differenzierbaren Funktion f eindeutig ihr k -tes **Taylorpolynom** $p_{f,\mathbf{x}_0,k}$ an einer Stelle \mathbf{x}_0 im Inneren des Definitionsbereichs zuordnen, indem man fordert, dass der Grad von p höchstens k ist und alle Ableitungen bis zur Ordnung k von f und $p_{f,\mathbf{x}_0,k}$ an der Stelle \mathbf{x}_0 übereinstimmen. Man hofft, dass dieses Taylorpolynom in der Nähe von \mathbf{x}_0 eine gute Approximation von f ist. Es geht also um Abschätzungen des Restglieds $r(\mathbf{x}) := f(\mathbf{x}) - p_{f,\mathbf{x}_0,k}(\mathbf{x})$. Ähnlich wie für $n = 1$ besagt der **Satz von Taylor** auch für Funktionen f in n Variablen, die $k + 1$ -mal differenzierbar sind:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{r(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^k} = 0.$$

Doch zurück zum Extremwertproblem: Ist \mathbf{x}_0 ein stationärer Punkt von f (d.h. einer mit $\text{grad}_{\mathbf{x}_0}(f) = \mathbf{o}$), so fallen im zweiten Taylorpolynom $p_{f,\mathbf{x}_0,2}$ von f an der Stelle \mathbf{x}_0 die linearen Terme weg. Daraus folgt, dass sich für zweimal stetig differenzierbares f in $p_{f,\mathbf{x}_0,2}(\mathbf{x}) = f(\mathbf{x}_0) + q_0(\mathbf{x})$ der Summand $q_0(\mathbf{x}) = q(\mathbf{x} - \mathbf{x}_0)$ als quadratische Form q in $\mathbf{x} - \mathbf{x}_0$ auffassen lässt (siehe gegen Ende von Abschnitt 1.4.2). Schreiben wir der übersichtlicheren Notation halber $\mathbf{x} - \mathbf{x}_0 = (x_1, \dots, x_n)$, so gilt:

$$q(\mathbf{x} - \mathbf{x}_0) = \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(\mathbf{x} - \mathbf{x}_0) \cdot x_i^2 + \sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} - \mathbf{x}_0) \cdot x_i x_j$$

Wegen des Satzes von Schwarz gilt stets $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x} - \mathbf{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x} - \mathbf{x}_0)$. Also ist die zugrundeliegende Bilinearform symmetrisch, und wir haben es wirklich mit einer quadratischen Form zu tun.

Die Approximation von f durch $p_{f,\mathbf{x}_0,2}$ ist, wie sich unschwer aus dem Satz von Taylor ableiten lässt, hinreichend gut, dass das Extremwertverhalten von q bei \mathbf{o} in der Regel für jenes von f bei \mathbf{x}_0 bestimmt. Wir verwenden dafür die übliche Terminologie für quadratische Formen:

Definition 2.4.3.3. Sei

$$q : \mathbb{R}^n \rightarrow \mathbb{R}, \quad q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j$$

mit $a_{i,j} = a_{j,i}$ für alle $i, j \in \{1, \dots, n\}$ eine quadratische Form. Dann heißt q

- **positiv definit**, wenn für alle $\mathbf{x} \neq \mathbf{o}$ die strikte Ungleichung $q(\mathbf{x}) > 0$ gilt. (In diesem Fall nimmt q nur in \mathbf{o} seinen minimalen Wert 0 an.)
- **negativ definit**, wenn für alle $\mathbf{x} \neq \mathbf{o}$ die strikte Ungleichung $q(\mathbf{x}) < 0$ gilt. (In diesem Fall nimmt q in \mathbf{o} seinen maximalen Wert 0 an.)
- **indefinit**, wenn es sowohl \mathbf{x} mit $q(\mathbf{x}) > 0$ als auch solche mit $q(\mathbf{x}) < 0$ gibt. (In diesem Fall hat q keine Extrema. Der Punkt $\mathbf{x} = \mathbf{o}$ ist ein sogenannter Sattelpunkt.)

2 Differentialrechnung in mehreren Variablen

- **positiv semidefinit**, wenn für alle \mathbf{x} die Ungleichung $q(\mathbf{x}) \geq 0$ gilt. (In diesem Fall nimmt q in \mathbf{o} seinen minimalen Wert 0 an, das kann aber auch an anderen Stellen der Fall sein.)
- **negativ semidefinit**, wenn für alle \mathbf{x} die Ungleichung $q(\mathbf{x}) \leq 0$ gilt. (In diesem Fall nimmt q in \mathbf{o} seinen maximalen Wert 0 an, das kann aber auch an anderen Stellen der Fall sein.)

Analog zur Situation in einer Variablen ergibt sich nun mit Hilfe des Satzes von Taylor:

Proposition 2.4.3.4. *Ist $f : \mathbb{R}^n \supseteq D \rightarrow \mathbb{R}$ zweimal stetig differenzierbar, \mathbf{x}_0 ein innerer Punkt von D , ein stationärer Punkt von f (d.h. $\text{grad}_{\mathbf{x}_0}(f) = \mathbf{o}$), und sei q die quadratische Form wie oben (die sich also aus dem zweiten Taylorpolynom von f an der Stelle \mathbf{x}_0 ergibt), dann gilt:*

- Ist q positiv definit, so ist \mathbf{x}_0 Stelle eines lokalen Minimums von f .
- Ist q negativ definit, so ist \mathbf{x}_0 Stelle eines lokalen Maximums von f .
- Ist q indefinit, so ist \mathbf{x}_0 nicht Stelle eines lokalen Extremums von f .
- Ist q semidefinit (positiv oder negativ), so kann keine allgemeine Aussage gemacht werden.

Hinter dem Hauptminorenkriterium 2.4.2.1 steckt nichts anderes, als eine Aussage über die Definitheit der quadratischen Form q , die sich aus dem zweiten Taylorpolynom ergibt.

Zum besseren Verständnis der Definitheit von quadratischen Formen und des darauf aufbauenden Hauptminorenkriteriums wollen wir die möglichen Situationen noch an einfachen, aber charakteristischen Beispielen illustrieren und den Fall $n = 1$ mit dem allgemeinen in Verbindung bringen.

Zunächst halten wir fest, dass das Hauptminorenkriterium für $n = 1$ mit dem bekannten Kriterium über die Rolle der zweiten Ableitung reeller Funktionen zusammenfällt: Alles hängt vom Vorzeichen der zweiten Ableitung $f''(\mathbf{x}_0) = f_{x_1x_1}(\mathbf{x}_0) = d_1$ ab. Ist $d_1 > 0$ (positive Definitheit), so liegt der erste Fall in Satz 2.4.2.1 und somit ein lokales Minimum von f vor. Für $d_1 < 0$ (negative Definitheit) sind wir im zweiten Fall, Maximum. Der indefinite Fall aus Satz 2.4.2.1 kommt bei einer Variablen nicht vor. Semidefinitheit entspräche der Nullform mit $q(\mathbf{x}) = 0$ für alle \mathbf{x} .

Für die schwierigere Situation mit $n \geq 2$ geht es um quadratische Formen in n Variablen, also um Funktionen $q : \mathbb{R}^n \rightarrow \mathbb{R}$ von der Gestalt

$$q(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{i,j} x_i x_j$$

mit $a_{i,j} = a_{j,i}$. Besonders interessiert uns die Stelle $\mathbf{x}_0 = \mathbf{o} = (0, \dots, 0)$, wo offenbar $q(\mathbf{o}) = q(0, \dots, 0) = 0$ gilt. Für $n = 2$ lassen sich alle Fälle von Definitheit sehr leicht diskutieren. Um Indizes zu sparen, schreiben wir $x = x_1$ und $y = x_2$.

2.4 Extremwertbestimmung in mehreren Variablen

- Das einfachste Beispiel einer positiv definiten Form ist $q(x, y) := x^2 + y^2$, d.h. $a_{1,1} = 1$, $a_{1,2} = a_{2,1} = 0$ und $a_{2,2} = 1$. Denn offenbar gilt für alle $\mathbf{x} = (x, y) \neq \mathbf{o}$ die strikte Ungleichung $q(\mathbf{x}) > 0$. Geometrisch hat man ein nach oben offenes Paraboloid vor Augen mit Scheitel (= tiefstem Punkt, Minimum) bei $\mathbf{x} = \mathbf{o}$. Jede zweimal stetig differenzierbare Funktion f mit dieser quadratischen Form als zweitem Taylorpolynom in $\mathbf{x}_\mathbf{o}$ wird durch q so gut approximiert, so dass auch f an dieser Stelle ein Minimum hat.
- Ganz analog ist $q(x, y) := -x^2 - y^2$ negativ definit. Diesmal ist das Paraboloid nach unten geöffnet, der Scheitel ist ein Maximum. Sonst ist alles analog.
- Das einfachste Beispiel für den indefiniten Fall, der gegenüber $n = 1$ substantiell Neues bringt, ist $f(x, y) := x^2 - y^2$. Hier ist also $a_{1,1} = 1$, $a_{1,2} = a_{2,1} = 0$ und $a_{2,2} = -1$. Man führe sich die sattelförmige Gestalt des Funktionsgebirges vor Augen. Entlang der x -Achse haben wir es mit einer nach oben hin offenen Parabel zu tun, entlang der y -Achse mit einer nach unten hin offenen. Vergleicht man das mit einer Passhöhe (daher auch die Bezeichnung **Sattelpunkt**), so verläuft die x -Achse in die Richtung des Bergkamms und die y -Achse in die der Passstraße.
- Semidefinit ist, wie schon weiter oben erwähnt, die Nullform $q(\mathbf{x}) = q(x, y) := 0$ (für alle $x, y \in \mathbb{R}$). Beispiele von Funktionen f mit diesem q als zweitem Taylorpolynom wären z.B. Potenzen höheren Grades. In Hinblick auf ein Extremum bei $\mathbf{x} = \mathbf{o}$ gibt es unterschiedliche Möglichkeiten, ganz analog dem Fall $n = 1$ beim Verschwinden der zweiten Ableitung: $f(x, y) = x^3$ hat in \mathbf{o} kein Extremum, $f(x, y) := x^4 + y^4$ ein striktes Minimum (d.h. ein Minimum nur in diesem Punkt), $f(x, y) := -x^4 - y^4$ ein striktes Maximum, $f(x, y) := x^4$ ein Minimum (aber kein striktes, weil an allen Punkten $\mathbf{x} = (0, y)$, $y \in \mathbb{R}$, ebenfalls ein Minimum vorliegt) und analog $f(x, y) := -x^4$ ein Maximum, das nicht strikt ist.

Wir vergegenwärtigen uns nochmals den indefiniten Fall mit der Veranschaulichung durch Bergkamm und Passstraße. Es scheint sehr plausibel, dass beide (wie x - und y -Achse im Beispiel) immer im rechten Winkel aufeinander stehen. Das ist tatsächlich der Fall, und zwar als Folgerung von Satz 1.6.3.4.

In unserem Zusammenhang bedeutet dieser Satz nämlich, dass es eine orthogonale Variablentransformation (d.h. einen Wechsel zu einer orthogonalen Basis in \mathbb{R}^n) gibt, so dass bezüglich der neuen Variablen x'_1, \dots, x'_n mit geeigneten Zahlen $\lambda_1, \dots, \lambda_n$ (den Eigenwerten der symmetrischen Matrix, die zur quadratischen Form gehört) gilt:

$$q(x'_1, \dots, x'_n) = \lambda_1 \cdot x'^2_1 + \dots + \lambda_n \cdot x'^2_n$$

Die oben als typische Beispiele für $n = 2$ behandelten Funktionen liegen durchwegs schon in dieser Gestalt vor. Für $f(x, y) = x^2 - y^2$ beispielsweise kann man $x'_1 = x$, $x'_2 = y$, $\lambda_1 = 1$ und $\lambda_2 = -1$ wählen.

2 Differentialrechnung in mehreren Variablen

Abschließend bringen wir noch ein Beispiel zur Transformation quadratischer Formen auf die Normalform, wie sie sich aus Satz 1.6.3.4 ergibt (der sogenannten **Hauptachsentransformation**) in Verbindung mit einer geometrischen Anwendung.

Es sei also

$$q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = \sum_{i,k=1}^n a_{i,k} x_i x_k \quad (A = A^T)$$

eine quadratische Form, dargestellt bezüglich $\mathbf{e}_1, \dots, \mathbf{e}_n$. Da A symmetrisch ist, kann man ein System von Eigenvektoren finden, $\mathbf{e}'_1, \dots, \mathbf{e}'_n$ etwa, das orthonormal ist. Es sei $S = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$. Dann wird die quadratische Form q in dieser neuen Basis in der Form

$$q(\mathbf{x}') = \mathbf{x}'^T S^T A S \mathbf{x}' = \lambda_1 x_1'^2 + \dots + \lambda_n x_n'^2$$

dargestellt.

Betrachtet man etwa die (verallgemeinerte) Fläche mit der Gleichung

$$q(\mathbf{x}) = a,$$

so hat sie im neuen Koordinatensystem die Gleichung

$$\lambda_1 x_1'^2 + \dots + \lambda_n x_n'^2 = a.$$

Aus dieser Gleichung kann man ablesen, um welchen Flächentyp es sich handelt:

$n=2$ Ellipse oder Hyperbel, Geradenpaar, Doppelgerade;

$n=3$ Ellipsoid, ein- oder zweischaliges Hyperboloid, ell., hyp. Zylinder, Kegel, Ebenenpaar, Doppellebene.

Beispiel. Man untersuche die Fläche

$$-x_1^2 + 2x_2^2 - x_3^2 + 4x_1x_2 + 8x_1x_3 - 4x_2x_3 = \mathbf{x}^{tr} \begin{pmatrix} -1 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & -1 \end{pmatrix} \mathbf{x} = 1.$$

$$\begin{pmatrix} -1 & 2 & 4 \\ 2 & 2 & -2 \\ 4 & -2 & -1 \end{pmatrix} \text{ ist symmetrisch;}$$

Eigenwerte:

$$\begin{vmatrix} -1 - \lambda & 2 & 4 \\ 2 & 2 - \lambda & -2 \\ 4 & -2 & -1 - \lambda \end{vmatrix} = 0 \Rightarrow \lambda_1 = 3, \lambda_2 = 3, \lambda_3 = -6;$$

Eigenvektoren:

$$\lambda_1 = \lambda_2 = 3 : \mathbf{e}'_1 = \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \mathbf{e}'_2 = \frac{1}{3} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \lambda_3 = -6 : \mathbf{e}'_3 = \frac{1}{3} \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}.$$

2.4 Extremwertbestimmung in mehreren Variablen

Bezüglich dieser neuen Orthonormalbasis wird die Fläche durch die folgende Gleichung dargestellt

$$3x_1'^2 + 3x_2'^2 - 6x_3'^2 = 1.$$

Es handelt sich also um ein einschaliges Rotationshyperboloid. Rotationsachse ist die x_3' -Achse. Der Umgang von den alten zu den neuen Koordinaten wird geleistet durch

$$\mathbf{x}' = \left[\frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \right]^{-1} \mathbf{x} = \frac{1}{3} \begin{pmatrix} 1 & -2 & 2 \\ 2 & 2 & 1 \\ -2 & 1 & 2 \end{pmatrix} \mathbf{x}.$$

Übungsaufgabe 153. (T) Geben Sie die Definition einer orthogonalen Matrix und bestimmen Sie eine orthogonale Matrix S , sodass

1.

$$S^{-1} \begin{pmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{pmatrix} S$$

2.

$$S^{-1} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} S$$

eine Diagonalmatrix ist.

Übungsaufgabe 154. (T)

1. Schreiben Sie das Polynom $3x_1^2 + 2x_2^2 + 2x_1x_2 + 3x_1x_3 + 4x_2x_3$ in der Form $\mathbf{x}A\mathbf{x}^T$ mit $\mathbf{x} = (x_1, x_2, x_3)$ und einer passenden 3×3 -Matrix A .
2. Dasselbe, wobei die Matrix A überdies symmetrisch sein soll.

Übungsaufgabe 155. (T*) Die folgenden Kegelschnittlinien sind durch eine Drehung S (d.h. orthogonale Matrix mit $\det S = 1$) des Koordinatensystems auf Hauptachsenform zu transformieren:

1. $x^2 + 3xy + y^2 = 1$
2. $3x^2 + 4xy - 3y^2 = c, \quad c \in \mathbb{R}.$

Um welche geometrischen Objekte handelt es sich dabei?

Übungsaufgabe 156. (T*) Die folgenden Flächen 2.Ordnung sind auf Hauptachsenform zu bringen und zu diskutieren:

1. $2xy + z^2 = 1$ (Drehhyperboloid),
2. $2xy + z^2 = 0$ (Doppelkegel).

Fertigen Sie eine Skizze an!

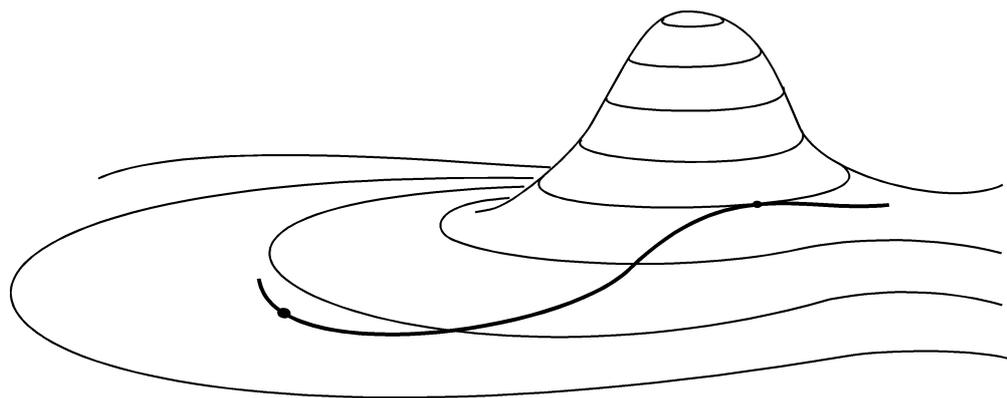
2.4.4 Extrema unter Nebenbedingungen

Inhalt in Kurzfassung: Eine Extremwertaufgabe mit Nebenbedingungen ist eine solche, wo man sich nur für einen Teilbereich des n -dimensionalen Definitionsbereich einer Funktion interessiert, der durch eine oder mehrere Gleichungen (Nebenbedingungen) beschrieben wird. Für solche Aufgaben steht die Methode der sogenannten Lagrangeschen Multiplikatoren zur Verfügung. Für den einfachsten Fall von 2 Variablen und einer Nebenbedingung wird diese Methode sowohl anschaulich motiviert als auch in ihrer präzisen Fassung präsentiert.

Oft sind Extrema einer Funktion $f : D \rightarrow \mathbb{R}$ auf einem Bereich $D \subseteq \mathbb{R}^n$ gesucht, dessen Rand ∂D (oder Teile von ∂D) durch eine Gleichung gegeben sind. Ein sehr einfaches Beispiel wäre $n = 2$ und ein Rechteck $D = [a, b] \times [c, d]$. Durch Nullsetzen der Ableitung f' würde man nur Extremstellen im Inneren des Rechtecks erfassen, nicht aber auf dem Rand. In dieser einfachen Situation kann man sich sehr leicht behelfen, denn allfällige Extrema auf den Seiten des Rechtecks findet man nach den aus Mathematik 1 bekannten Methoden durch Behandlung der reellen Funktionen $x \mapsto f(x, c)$ und $x \mapsto f(x, d)$ auf dem Intervall $[a, b]$ beziehungsweise $y \mapsto f(a, y)$ und $y \mapsto f(b, y)$ auf dem Intervall $[c, d]$.

Schwieriger wäre die Aufgabe, wenn D beispielsweise die durch $x^2 + y^2 = 1$ begrenzte Einheitskreisscheibe ist. Zwar könnte man aus der Kreisgleichung eine der Variablen durch die andere ausdrücken, den Kreis in Halbkreise zerlegen (etwa $y = g_1(x) := \sqrt{1 - x^2}$ und $y := g_2(x) := -\sqrt{1 - x^2}$), diese Darstellung in f einsetzen und auf diese Weise die Funktion $f(x, y)$ in zwei Variablen auf die Funktionen $f(x, g_1(x))$ und $f(x, g_2(x))$ in nur mehr einer Variablen zurückführen. Rechnerisch ist das aber oft sehr unangenehm. Eine Methode, die dem abhilft, ist die der sogenannten **Lagrange-Multiplikatoren**.

Die geometrische Idee ist die folgende. Gegeben sei ein Funktionsgebirge (eingezeichnet mit Höhenschichtlinien) mit einer Straße (gekennzeichnet durch die dicke Linie), deren höchster Punkt $(\mathbf{a}, f(\mathbf{a}))$ gesucht ist:



Der Grundriss der Höhenschichtlinie, auf der dieser Punkt liegt, ist durch die Gleichung

$$(2) \quad f(x, y) = f(\mathbf{a}).$$

gegeben. Wir wollen annehmen, dass auch der Grundriss der Straße durch eine Gleichung

$$(1) \quad g(x, y) = 0,$$

gegeben ist. Die Anschauung lehrt, dass Höhenschichtlinie und Straße im höchsten Punkt der Straße dieselbe Tangente haben. Damit haben auch die Grundrisse beider Kurven dieselbe Normale, deren Richtung wiederum die Gradienten $\text{grad}_{\mathbf{a}}(f)$ und $\text{grad}_{\mathbf{a}}(g)$ von f und g sind. Also ist einer der beiden Gradientenvektoren ein Vielfaches des anderen. Das wiederum bedeutet, dass es einen Faktor $\lambda \in \mathbb{R}$ gibt mit

$$\text{grad}_{\mathbf{a}}(f) + \lambda \text{grad}_{\mathbf{a}}(g) = \mathbf{o},$$

oder

$$\begin{aligned} f_x(\mathbf{a}) + \lambda g_x(\mathbf{a}) &= 0, \\ f_y(\mathbf{a}) + \lambda g_y(\mathbf{a}) &= 0. \end{aligned}$$

Mit Hilfe des Hauptsatzes über implizite Funktionen 2.3.4.1 lässt sich diese Überlegung zu folgendem Satz präzisieren:

Satz 2.4.4.1. *Angenommen, $f, g : D \rightarrow \mathbb{R}$ seien auf $D \subseteq \mathbb{R}^2$ stetig differenzierbar. Ist $\mathbf{a} \in D$ eine relative Extremstelle von f unter der Nebenbedingung $g = 0$ mit $\text{grad}_{\mathbf{a}} g \neq \mathbf{o}$, dann gibt es ein $\lambda \in \mathbb{R}$ mit*

$$(3) \quad \begin{aligned} f_x(\mathbf{a}) + \lambda g_x(\mathbf{a}) &= 0, \\ f_y(\mathbf{a}) + \lambda g_y(\mathbf{a}) &= 0, \\ g(\mathbf{a}) &= 0. \end{aligned}$$

Die Voraussetzung $\text{grad}_{\mathbf{a}} g \neq \mathbf{o}$ versteht man, wenn man bedenkt, dass man im Beweis von Satz 2.4.4.1 den Hauptsatz über implizite Funktionen auf die Nebenbedingung $g(x, y) = 0$ anwendet. Damit eine der beiden Variablen als Funktion der anderen dargestellt werden kann, ist erforderlich, dass in \mathbf{a} wenigstens eine der beiden partiellen Ableitungen von g verschieden ist, sprich dass der Gradient von g in \mathbf{a} nicht verschwindet.

Die Anwendung von Satz 2.4.4.1 zwecks Extremwertsuche läuft also in der Praxis auf die Lösung des (in der Regel nichtlinearen) Systems (3) von drei Gleichungen in den drei Unbekannten x, y (für $\mathbf{a} = (x, y)$) und λ hinaus. Oft gelingt das durch geschickte Variablenelimination, oder aber man kann es mit dem Newtonverfahren (siehe 2.3.2) versuchen. Hat man die Lösungen des Gleichungssystems gefunden, hat man aber noch zu untersuchen, welche davon tatsächlich Extremstellen sind.

Satz 2.4.4.1 ist die einfachste Variante der Methode der sogenannten **Lagrangeschen Multiplikatoren**. Es gibt eine allgemeinere Version für Funktionen in mehr als 2 Variablen und eventuell auch mehreren Nebenbedingungen. Beim Beweis wird der Hauptsatz über implizite Funktionen 2.3.4.1 in seiner allgemeinen Form wirksam, die Bedingung $\text{grad}_{\mathbf{a}} g \neq \mathbf{o}$ aus Satz 2.4.4.1 ist daher auch an die etwas kompliziertere des Hauptsatzes anzupassen.

Übungsaufgabe 157. (T) Finden Sie drei positive Zahlen x , y und z , deren Summe 48 ist und deren Produkt maximal ist.

Hinweis: Dieses Beispiel kann man auch direkt, d.h. ohne Lagrange-Multiplikatoren lösen. Die Bedingung an die Summe der drei Zahlen erlaubt es, eine der Zahlen explizit auszudrücken–Einsetzen dieses Ausdrucks in die zu maximierende Funktion liefert ein Extremalproblem in zwei Variablen.

Übungsaufgabe 158. (T) Sei $f(x, y) = x^2 + 4y^3$. Bestimmen Sie mithilfe von Lagrange-Multiplikatoren die Extrema auf der Ellipse $x^2 + 2y^2 = 1$. Begründen Sie, warum die gefundenen Punkte tatsächlich Minima bzw. Maxima sind.

Übungsaufgabe 159. (T) Gegeben sei die durch die Gleichung $3(x-2)^2 + 2y^2 - z^2 - 1 = 0$ bestimmte Fläche im \mathbb{R}^3 . Bestimmen Sie mit Hilfe Lagrange'scher Multiplikatoren, welche Punkte dieser Fläche den geringsten Abstand vom Ursprung haben.

Übungsaufgabe 160. (T) Berechnen Sie unter Verwendung der Lagrange'schen Multiplikatoren den minimalen Abstand des Punktes $P = (2, 3)$ zu der Geraden $y = 2x - 4$. Begründen Sie warum der gefundene Punkt tatsächlich ein Minimum darstellt.

Übungsaufgabe 161. (T) Bestimmen Sie mit der Methode der Lagrange-Multiplikatoren einen Zylinder maximalen Volumens, der in eine Kugel vom Radius R eingeschrieben werden kann.

Übungsaufgabe 162. (T) Man bestimme diejenigen Dreiecke, für welche das Produkt der Sinuswerte der Winkel maximal ist, d.h. für welche $f(x, y) = \sin x \cdot \sin y \cdot \sin(x + y)$ in $0 \leq x \leq \pi$, $0 \leq y \leq \pi$, $0 \leq x + y \leq \pi$ ein Maximum annimmt.

Übungsaufgabe 163. (T) Bestimmen Sie die Extrema der folgenden Funktionen unter den angegebenen Nebenbedingungen.

1. $f(x, y) = x^2 + y^3$, $x + y = 1$
2. $f(x, y) = x^3 + y^3 + z^3$, $x + y + z = 1$
3. $f(x, y) = 3x^2 + 4xy + 3y^2$, $x^2 + y^2 = 1$
4. $f(x, y) = x^2 + y^2$, $3x^2 + 4xy + 3y^2 = 2$
5. $f(x, y, z) = 2x^2 + 3y^2 + z^2$, $x^2 + y^2 + z^2 = 1$
6. $f(x, y, z) = 2x^2 + 3y^2 + z^2$, $2x + 3y + z = 0$

2.4.5 Strategie zur Lösung von Extremwertaufgaben

Inhalt in Kurzfassung: Aus den bisher entwickelten Methoden wird eine Strategie entwickelt, mit der man viele Extremwertaufgaben lösen kann.

2.4 Extremwertbestimmung in mehreren Variablen

Sei wieder $f : D \rightarrow \mathbb{R}$, $\mathbf{x} = (x_1, \dots, x_n) \mapsto f(\mathbf{x})$ mit einem nichtleeren Definitionsbereich $D \subseteq \mathbb{R}^n$ gegeben. In jedem Fall wollen wir für f Stetigkeit, oft auch Differenzierbarkeitseigenschaften voraussetzen. Bei der Extremstellensuche macht es einen Unterschied, ob – wie etwa bei einer typischen Optimierungsaufgabe – nur globale oder – eher im Sinne einer Kurvendiskussion – auch lokale Extrema gesucht sind.

Zunächst empfiehlt es sich zu klären, ob f auf D überhaupt ein Maximum und/oder Minimum annimmt. Jedenfalls ist das nach dem Satz vom Maximum der Fall, wenn f stetig und D kompakt (d.h. abgeschlossen und beschränkt) ist. Mit den bisher entwickelten Mitteln der Differentialrechnung ist es dann sehr häufig möglich, eine endliche Menge E zu ermitteln, die alle potentiellen lokalen Extremstellen enthält. Sucht man lediglich nach globalen Extrema, so genügt es, $f(\mathbf{x})$ für alle $\mathbf{x} \in E$ auszuwerten und jene $\mathbf{x} \in E$ zu identifizieren, wo dieser Wert am größten bzw. am kleinsten ist.

Ist D nicht kompakt, so muss auch ein stetiges f keine Extrema annehmen. Meist erkennt man schnell, ob f nach unten und/oder oben unbeschränkt ist, in welchem Fall die Suche nach einem globalen Minimum bzw. Maximum obsolet wird. Ist f sehr wohl beschränkt, sind eigene, an das jeweilige f angepasste Überlegungen erforderlich, um zu entscheiden, ob globale Extrema existieren.

Sind auch sämtliche lokalen Extrema gesucht sind, ist zu jedem Wert $\mathbf{x} \in E$ zu entscheiden, ob es sich um ein lokales Maximum, ein lokales Minimum oder um kein lokales Extremum handelt. Für Punkte im Inneren von D ist die wirkungsvollste Methode das Hauptminorenkriterium 2.4.2.1. Man beachte, dass sein Einsatz in den meisten Fällen, wo nur globale Extrema gesucht sind, nicht erforderlich ist. Überdies ist zu bedenken, dass das Hauptminorenkriterium für Punkte am Rand von D nicht wirksam ist.

Zusammenfassend kann für Extremwertaufgaben in mehreren Variablen folgende Strategie empfohlen werden:

1. Kontrolle der Regularität der Aufgabenstellung, genauer:
 - a) Ist f auf ganz D stetig, differenzierbar, zweimal differenzierbar? Punkte im Inneren von D , wo f nicht differenzierbar ist, müssen extra behandelt werden. Häufig ist f aber unendlich oft differenzierbar, und man hat diesbezüglich keine Probleme.
 - b) Lässt sich der Rand ∂D von D in endlich viele Teile mit Dimensionen $< n$ zerlegen, so dass jeder Teilbereich als stetig differenzierbare Nebenbedingung zu verstehen ist? Für $n = 2$ geht es dabei nur um berandende Kurvenstücke, die jeweils durch eine Gleichung beschrieben werden. Für $n = 3$ können sowohl berandende Seitenflächen (eine Gleichung in drei Variablen) als auch, als deren Schnittmengen, Seitenkanten (zwei Gleichungen in drei Variablen) auftreten. Für noch größeres n kann die Situation entsprechend noch komplexer werden. Sind von vornherein Nebenbedingungen gegeben, ist damit sinngemäß umzugehen, während im Folgenden jene Schritte, die sich auf das Innere von D beziehen wegfallen.

Für die weitere Vorgangsweise sei vorausgesetzt, dass a) und b) positiv beantwortet

2 Differentialrechnung in mehreren Variablen

werden können. Um das zu entscheiden, genügt oft ein kurzer Blick.

2. Ist D zusätzlich kompakt (d.h. abgeschlossen und beschränkt), so nimmt f sowohl ein Maximum als auch ein Minimum an. Dies zu wissen, kann im Folgenden Arbeit ersparen und sollte deshalb früh festgehalten werden. Doch auch bei nicht kompaktem D steht den nächsten Schritten nichts im Wege.
3. Zerlegung des Randes ∂D von D im Sinne von 1.b) in Teilbereiche D_i der folgenden drei Typen, die man mit der jeweils angemessenen Methode behandelt:
 - a) Inneres D° von D : Hier sind keine Nebenbedingungen wirksam.
 - b) Teilmengen des Randes ∂D von D der Dimension 1 bis $n - 1$, von denen sich jede als (Teil der) Nullstellenmenge von einer oder mehreren Gleichungen (als Teil der Lösungsmenge von Nebenbedingungen) darstellen lässt. Für $n = 2$ geht es nur um Kurven, die D begrenzen.
 - c) Einzelne (Eck-)Punkte von D .

Diese drei Typen werden in den folgenden drei Schritten behandelt.

4. Auf dem n -dimensionalen Bereich D° sucht man alle Nullstellen der Ableitung von f auf. Das entspricht der Lösung eines Systems bestehend aus den n Gleichungen $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = 0$, $i = 1, \dots, n$, in den n Variablen x_1, \dots, x_n . Die Menge all dieser Nullstellen sei E_1 . (Bei Extremwertaufgaben mit Nebenbedingungen entfällt dieser Punkt.)
5. Für jede der Teilmengen von ∂D aus 3.b) löst man das entsprechende Extremwertproblem unter Nebenbedingungen wie in 2.4.4 beschrieben. Die Menge aller so erhaltenen Lösungen sei E_2 .
6. Sei E_3 die Menge aller Eckpunkte laut 3.c). Alle lokalen Extrema von f sind in der Menge $E := E_1 \cup E_2 \cup E_3$ enthalten. Für die Bestimmung der globalen Extrema hat man also nur alle Werte $f(\mathbf{x})$ auszuwerten und zu vergleichen, welche am größten bzw. am kleinsten sind.
7. Interessiert man sich auch für lokale Extrema, wendet man auf die Punkte in E_1 das Hauptminorenkriterium 2.4.2.1 an. Für die Punkte aus E_2 und E_3 muss man zu ad-hoc-Methoden greifen.

Übungsaufgabe 164. (T) Konstruieren Sie für die Funktion

$$f(x, y) = 2x^3 - 5x^2 + 3xy - 2y^2 + 9x - 9y - 9$$

das Taylor-Polynom 2. Ordnung um den Punkt $(1, -1)$. Geben Sie außerdem die Tangentialebene bzw. das Schmiegeparaboloid an den Graphen von f im Punkt $(1, -1, 1)$ an!

Übungsaufgabe 165. (T) Entwickeln Sie die Funktion $f(x, y) = x^2y - x^2 + xy + 2x - 3y + 9$ mit Hilfe der Taylorformel nach Potenzen von $x + 1$ und $y - 2$.

Übungsaufgabe 166. (T) Konstruieren Sie für die Funktion

$$f(x, y) = \cos x \sin y e^{x-y}$$

das Taylor-Polynom 2. Ordnung um den Punkt $(0, 0)$. Geben Sie weiters die Tangentialebene bzw. das Schmiegeparaboloid an den Graphen von f im Punkt $(0, 0, 0)$ an!

Übungsaufgabe 167. (T) Sei $f(x, y) = x^2 + xy + 17x - 3y + 4$. Berechnen Sie die vollständige Taylorreihe von $f(x, y)$ im Punkt $P = (1, 2)$. Lesen Sie aus dieser Taylorformel die Gleichung für die Tangentialebene im Punkt P ab. Geben Sie den Normalvektor dieser Tangentialebene an. Von welcher Art ist der Punkt P (elliptisch, hyperbolisch bzw. parabolisch)?

Übungsaufgabe 168. (T) Durch $z = f(x, y) = \sin(xy)$ ist eine Fläche gegeben. Bestimmen Sie die Tangentialebene bei $(x_0, y_0) = (\sqrt{\frac{\pi}{2}}, \sqrt{\frac{\pi}{2}})$.

Übungsaufgabe 169. (T) Untersuchen Sie die Fläche $z = f(x, y) := e^{-r^2}$ mit $r = \sqrt{x^2 + y^2}$:

1. Von welchem Typ sind die Punkte $(-\frac{1}{2}, 0, z)$, $(\frac{3}{5}, \frac{4}{5}, z)$ und $(-\frac{1}{\sqrt{2}}, 0, z)$.
2. In welchen Bereichen der x, y -Ebene besteht die Fläche aus elliptischen bzw. hyperbolischen bzw. parabolischen Punkten? (Skizzieren Sie die Fläche!)

Übungsaufgabe 170. (T*) Betrachten Sie die Funktion

$$f(x, y) = y - x \tan\left(\frac{\pi}{2}(x^2 + y^2)\right).$$

Durch die implizite Gleichung $f(x, g(x)) = 0$ wird eine reelle Funktion $g(x)$ definiert. Geben Sie das Taylorpolynom 2. Ordnung von $g(x)$ um $x = -\sqrt{2}$ an. Wie lautet die Schmiegeparabel an $g(x)$ an $x = -\sqrt{2}$?

Übungsaufgabe 171. (T) Sei $f(x, y) = xy - x^2$. Berechnen Sie alle lokalen und globalen Extrema im Einheitsquadrat $R = \{(x, y) : 0 \leq x, y \leq 1\}$. Beachten Sie speziell auch mögliche globale Extremstellen am Rand. Welche Extrema sind Minima bzw. Maxima?

Übungsaufgabe 172. (T) Berechnen Sie alle lokalen und globalen Extrema folgender auf \mathbb{R}^2 definierter Funktionen

$$f(x, y) = y^2 + 2x^2 - 6y + 8x - 1, \quad g(x, y) = x \cos(y).$$

Übungsaufgabe 173. (T) Bestimmen Sie alle lokalen und globalen Extrema der folgenden Funktionen $f(x, y)$ im angegebenen Bereich D . Beachten Sie insbesondere auch mögliche Extrema am Rand von D .

1. $f(x, y) = x^2 - y^2$, $D = \{(x, y) : x^2 + y^2 \leq 1\}$
2. $f(x, y) = e^{x^2 - y^2}$, $D = \{(x, y) : \frac{1}{2} \leq x^2 + y^2 \leq 2\}$

2 Differentialrechnung in mehreren Variablen

minimal ist. Die gesuchte Gerade approximiert also die gegebene Punktmenge in diesem speziellen Sinn bestmöglich.

Bei $|a| + |b| \rightarrow \infty$ geht $f(a, b) \rightarrow +\infty$. Die stetige Funktion f muss also mindestens eine Minimalstelle haben. Nach unserem Satz kommen als Kandidaten für eine solche Stelle die Punkte $\mathbf{a} = (a, b)$ in Frage, für die gilt

$$\begin{aligned}f_a(\mathbf{a}) = 0 &= -2 \sum (y_i - a - bx_i), \\f_b(\mathbf{a}) = 0 &= -2 \sum x_i (y_i - a - bx_i).\end{aligned}$$

D.h.

$$\begin{aligned}an + b \sum x_i &= \sum y_i, \\a \sum x_i + b \sum x_i^2 &= \sum x_i y_i.\end{aligned}$$

Dieses Gleichungssystem für a, b ist, wie man leicht zeigen kann, wegen unserer Voraussetzung, dass die x_i 's nicht alle gleich sind, eindeutig nach a, b auflösbar. Da mindestens eine Minimalstelle vorhanden ist und nur ein Kandidat dafür existiert, stellt dieser tatsächlich die Minimalstelle dar. Die zugehörige Gerade heißt **Regressionsgerade**.

3 Differentialgleichungen (DGLen)

Differentialgleichungen entstammen vor allem den zahlreichen Anwendungsgebieten der Mathematik, allen voran und in klassischer Weise der Physik, aber auch der Biologie, den Wirtschaftswissenschaften etc. Aufgrund ihrer außermathematischen Bedeutung hat sich auch innerhalb der Mathematik eine weit verzweigte Theorie entwickelt. Im Rahmen einer einführenden Mathematikvorlesung für angehende Ingenieure ist es unmöglich, einen repräsentativen Querschnitt der Theorie auch nur jener DGLen zu geben, die in den spezielleren Fächern der jeweiligen Disziplin Bedeutung erlangen (können). Deshalb erscheint es vor allem wichtig, jene Begriffe kennen zu lernen, die dem sehr weitläufigen Gebiet eine sinnvolle Struktur aufgeprägt und somit als Orientierungshilfe für weitere Vertiefungen dienen können.

In diesem Sinne behandeln wir im ersten Abschnitt des Kapitels (3.1) vor allem allgemeine Aspekte. Konkreter wird es in ??, wo anhand einfacher linearer DGLen einige Standardmethoden zur Lösung präsentiert werden. Von großer Bedeutung, gleichzeitig aber auch von deutlich höherer Komplexität sind gekoppelte lineare DGLen, die in ?? behandelt werden. Generell spielen lineare DGLen deshalb eine wichtige Rolle, weil es einerseits eine vergleichsweise befriedigende Theorie und auch relativ wirkungsvolle Lösungsmethoden gibt und andererseits nichtlineare DGLen oft näherungsweise durch lineare ersetzt werden können. In 3.3, dem letzten Abschnitt des Kapitels, werden ein paar Herangehensweisen dargestellt, mit denen auch einige wichtige Typen nichtlinearer Differentialgleichungen behandelt werden können.

3.1 Allgemeine Theorie

In diesem Abschnitt besprechen wir zuerst (3.1.1) den Begriff der DGL als Spezialfall einer Funktionalgleichung. In 3.1.2 halten wir einige Notationen fest. Sodann erläutern wir in 3.1.3 die Unterscheidung zwischen gewöhnlichen (auf die wir uns später weitgehend beschränken werden) und den komplizierteren partiellen DGLen (die in der Vorlesung Mathematik 3 ausführlicher behandelt werden). Der wichtige Begriff der Ordnung einer DGL wird in 3.1.4 thematisiert. Weil die Ordnung einer DGL reduziert werden kann, sofern man eine höhere Dimension in Kauf nimmt, werden diese Aspekte gemeinsam behandelt. Der wichtigste theoretische Satz aus der Theorie der DGLen, der Existenz- und Eindeutigkeitssatz für explizite DGLen, steht im Zentrum von 3.1.5. Am Beispiel autonomer DGLen gehen wir in 3.1.6 auf qualitative Untersuchungen von autonomen DGLen ein, die in vielen Fällen auch dann möglich sind, wenn keine expliziten Lösungsverfahren zur Verfügung stehen. Vor der eigentlichen Lösungstheorie für lineare DGLen in den Abschnitten ?? und ?? wird der Aspekt der Linearität in 3.2.1 noch unter allgemeinen Gesichtspunkten behandelt.

3.1.1 DGLen als Funktionalgleichungen

Inhalt in Kurzfassung: Eine DGL ist eine Gleichungen, deren Unbekannte (oft mit dem Buchstaben y bezeichnet) nicht eine Zahl ist, sondern eine Funktion. Allgemein nennt man so etwas eine Funktionalgleichung. Bei einer DGL kommt hinzu, dass neben y selbst auch Ableitungen von y vorkommen dürfen. Bereits bekannt ist die DGL $y' = y$ mit den Lösungen $y(x) = ce^x$, $c \in \mathbb{R}$. Verlangt man die Anfangsbedingung $y(0) = 1$, so wird die Lösung eindeutig: $y = \exp$.

Eine **Funktionalgleichung** ist eine Gleichung, deren (gesuchte) Lösungen nicht Zahlen oder Vektoren sondern Funktionen sind. Wichtige Beispiele, die wir bereits in Mathematik 1 kennen gelernt haben, sind (gefordert jeweils für alle x, y aus dem Definitionsbereich):

- $f(x + y) = f(x) + f(y)$, Lösungen: homogene lineare Funktionen $f = \text{lin}_k : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto kx$, mit einer Steigung $k \in \mathbb{R}$.
- $f(x + y) = f(x) \cdot f(y)$, Lösungen: Exponentialfunktionen $f = \exp_a : \mathbb{R} \rightarrow \mathbb{R}^+$, $x \mapsto a^x$, zu einer Basis $a > 0$.
- $f(x \cdot y) = f(x) + f(y)$, Lösungen: Logarithmen $f = \log_a : \mathbb{R}^+ \rightarrow \mathbb{R}$, $x \mapsto \log_a x$, zu einer Basis $a > 0$, $a \neq 1$.
- $f(x \cdot y) = f(x) \cdot f(y)$, Lösungen: Potenzfunktionen $f = \text{pot}_\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, $x \mapsto x^\alpha$, mit einem Exponenten $\alpha \in \mathbb{R}$.

Unter sehr schwachen Zusatzvoraussetzungen (z.B. Stetigkeit an nur einer Stelle oder Monotonie) sind die angegebenen Funktionen die einzigen, die die angegebenen Gleichungen für alle x, y aus ihrem Definitionsbereich erfüllen. Auch die Parameter k, a bzw. α ergeben sich eindeutig, wenn man nur einen einzigen Funktionswert an einer geeigneten Stelle vorgibt. Das ist Inhalt von sogenannten Existenz- und Eindeutigkeitsätzen für die angegebenen Funktionstypen.

Das gemeinsame Merkmal von Funktionalgleichungen wie den oben angegebenen ist, dass sie Funktionswerte an mehreren Stellen, nämlich an den Stellen x, y und $x + y$ und/oder xy , zueinander in Beziehung setzen. Man beachte den Unterschied zum Beispiel zu Gleichungen und Gleichungssystemen, wie wir sie in der Linearen Algebra behandelt haben, wo die gesuchten Objekte einzelne Zahlen oder Vektoren waren, nicht Funktionen, die auf einer unendlichen Menge definiert sind.

Noch andere, einfachere Beispiele von Funktionalgleichungen haben wir bereits kennengelernt, nämlich rekursive Folgen: Ist eine Transformation $T : X \rightarrow X$ auf einer Menge X gegeben, so lässt sich eine Rekursionsgleichung $a_{n+1} = T(a_n)$ auch als Funktionalgleichung $f(n+1) = T(f(n))$ für Funktionen $f : \mathbb{N} \rightarrow X$ auffassen, deren Lösungen genau die rekursiven Folgen zu beliebigem Anfangswert x_0 sind. Ist auch x_0 vorgegeben, so gibt es, weil die Rekursionsgleichung ja für alle $n \in \mathbb{N}$ gelten soll, genau eine Lösung. Das ist ein Spezialfall der sogenannten **Rekursionsatzes**.

Differentialgleichungen sind spezielle Funktionalgleichungen. Bei ihnen kommen Ableitungen der gesuchten Funktion vor. Ableitungen involvieren als Differentialquotienten Grenzwerte und somit sogar Funktionswerte an unendlich vielen Stellen. Deshalb sind sie tatsächlich Beispiele von Funktionalgleichungen, auch wenn sie selten so bezeichnet werden und man beim Wort *Funktionalgleichung* nicht unbedingt an DGLen denkt.

Differentialgleichungen sind sehr häufig durch ursprünglich außermathematische Fragestellungen motiviert, die sich aber sehr präzise mathematisch modellieren lassen. Klassische Beispiele stammen aus der Physik. Denn physikalische Systeme lassen sich sehr häufig durch eine Funktion f beschreiben, deren Werte den Zustand des Systems in Abhängigkeit beispielsweise der Zeit t und/oder des Ortes x angeben. Die zeitliche und räumliche Änderung des Zustandes (die Ableitung von f nach der Zeit oder nach Raumkoordinaten) steht oft in einem naturgesetzlichen Zusammenhang mit dem Zustand, also mit dem Wert der Funktion selbst. Wenn sich dieser Zusammenhang durch eine Gleichung ausdrücken lässt, liegt eine Differentialgleichung vor. Wie die für die moderne Physik so wichtige **Schrödingergleichung** zeigt, sind sogar komplexwertige Funktionen als Lösungen von DGLen von Interesse. Wir werden uns aber ausschließlich mit reellwertigen befassen.

Das wichtigste einfache Beispiel einer interessanten Differentialgleichung kennen wir bereits: Die Ableitungsregel $\exp' = \exp$ für die Exponentialfunktion \exp zur Basis e besagt, dass die Exponentialfunktion Lösung der Differentialgleichung $y' = y$ ist. (Diese DGL tritt typischerweise bei Wachstums- oder Zerfallsprozessen auf, wo die Zu- bzw. Abnahme proportional zum Bestand der durch y beschriebenen Quantität sind.) Dabei ist y eine Variable für reelle Funktionen $y : x \mapsto y(x)$. Manchmal schreibt man auch lieber $x(t)$ für eine gesuchte reelle Funktion $x : t \mapsto x(t)$. Die Variable t verwendet man vorzugsweise, wenn sie als Zeit interpretiert wird. Für die Ableitung x' von x nach t schreibt man dann auch gerne \dot{x} .

Doch zurück zur Differentialgleichung $y' = y$. Neben \exp sind auch alle Vielfachen $c \exp$ mit $c \in \mathbb{R}$ Lösungen, weil ja auch $(c \exp)' = c \exp' = c \exp$ gilt. Gibt man sich für die gesuchte Lösung einen Funktionswert, zum Beispiel $y(0) := y_0 \in \mathbb{R}$ als sogenannten **Anfangswert** vor, so erweist sich $c \exp$ mit $c = y_0$ sogar als einzige Lösung des sogenannten **Anfangswertproblems**. Für $y_0 = 1$ lässt sich das sehr leicht rechnerisch begründen. Dafür hat man zu zeigen, dass eine differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f' = f$ und $f(0) = 1$ automatisch $f = \exp$ erfüllt. Dazu betrachten wir die Funktion $g := \frac{f}{\exp} : \mathbb{R} \rightarrow \mathbb{R}$. Mit der Quotientenregel berechnen wir die Ableitung

$$g' = \left(\frac{f}{\exp} \right)' = \frac{f' \exp - f \exp'}{\exp^2} = \frac{f \exp - f \exp}{\exp^2} = 0.$$

Daraus folgt, dass $g = \frac{f}{\exp}$ konstant ist, also $\frac{f(x)}{e^x} = c$ und somit $f(x) = ce^x$ für alle $x \in \mathbb{R}$ mit einem $c \in \mathbb{R}$. Setzen wir speziell $x = 0$, so erhalten wir $c = ce^0 = f(0) = y_0 = 1$. Damit ist $f(x) = e^x$ für alle $x \in \mathbb{R}$, also $f = \exp$ bewiesen.

Dieser Beweis für die Aussage, dass es genau eine Lösung des Anfangswertproblems $y' = y$, $y_0 = y(0) = 1$ gibt, gelang ad hoc aufgrund der sehr einfachen Bauart der gegebe-

3 Differentialgleichungen (DGLen)

ne DGL. Ziel der Theorie sind vor allem allgemeinere Existenz- und Eindeutigkeitsätze für DGLen. Für eine weite Klasse (in die auch unser Beispiel fällt), wird der Beweis eines solchen Satzes in 3.1.5 gelingen. Leider werden damit aber noch lange nicht alle interessanten DGLen erfasst. Oft muss man sich mit mühsam errungenen schwächeren Aussagen begnügen. Noch gar nicht berührt ist damit die Frage nach Rechenverfahren, mit denen allfällig existierende Lösungen auch explizit gefunden werden können. Die damit angerissenen Fragen sollen nun präzisiert und behandelt werden.

3.1.2 Bemerkungen zur Notation

Inhalt in Kurzfassung: Die Notation für DGLen ist in der Literatur generell und auch innerhalb dieses Skriptums nicht einheitlich. Einige weit verbreitete Möglichkeiten sollen nun zusammengestellt werden.

Die Theorie der DGLen geht auf die Gründerväter der Differentialrechnung zurück, Schon zur Zeit von Newton und Leibniz gab es Uneinigkeit in der Notation (und nicht nur darüber). Vor allem durch Euler haben sich im 18. Jahrhundert zwar manche Konventionen breit durchgesetzt, weil sie sich im praktischen Umgang mit DGLen bewährt haben. Allerdings sind nicht alle diese Notationen mit den Sichtweisen der modernen Mathematik verträglich, wie sie sich vor allem im 20. Jahrhundert durchgesetzt haben. Aus diesen (und wahrscheinlich manch anderen Gründen) herrscht in den üblichen Notationen nach wie vor eine gewisse Uneinheitlichkeit, die auch im vorliegenden Text durchschlägt. Um Verwirrungen möglichst hintan zu halten, folgt nun eine kurze Zusammenstellung der für uns wichtigsten Schreibweisen.

Schon für die (unabhängigen) Variablen und die davon abhängigen Funktionen, sowie deren Ableitungen sind verschiedene Notationen üblich. Meistens bezeichnet y eine Funktion (abhängige Variable), die von einer unabhängigen Variable x oder t abhängt. Die Variable t hat sich aufgrund physikalischer Anwendungen eingebürgert, in denen Systemveränderungen im Laufe der Zeit beschrieben werden. Wegen lateinisch *tempus*=Zeit (oder auch englisch *time* und französisch *temps*) wird als Zeitvariable gerne der Buchstabe t verwendet. Ableitungen nach der Zeit t werden oft nicht mit y' , sondern mit \dot{y} bezeichnet.

Eine andere und universellere Möglichkeit, die Variable, nach der abgeleitet wird, explizit sichtbar zu machen, bietet die auf Leibniz zurückgehende Differentialschreibweise: $y' = \frac{dy}{dt}$ oder $y' = \frac{dy}{dx}$ – je nachdem, ob y von t oder von x abhängt. Leibniz dachte bei seinen sogenannten **Differentialen** dt, dx, dy an „infinitesimal kleine Größen“, die kleiner sind als jede feste positive Zahl, aber doch nicht 0, damit man eine durch die andere dividieren kann. Der Buchstabe d kommt von *Differenz*, weil man an Änderungen (Differenzen) von Werten der Variablen t, x, y denkt. So werden Quotienten der Differentiale scheinbar tatsächlich zu Differentialquotienten, in Übereinstimmung mit der gegebenen Schreibweise – nur *scheinbar* deshalb, weil dieser vage Begriff von Differential den Exaktheitsansprüchen der modernen Mathematik noch nicht genügt. Man hat

deshalb unterschiedliche Begriffssysteme und Theorien (Schlagwort Differentialformen) entwickelt, um dem abzuhelfen. Für uns würde diese Theorien jedoch viel zu weit führen. Dort, wo sich die Differentialschreibweise praktisch bewährt (etwa in 3.3.2), werden wir sie verwenden und auch rechtfertigen. Übrigens haben wir das auch schon einmal in Mathematik 1 getan, und zwar im Zusammenhang mit der Substitutionsregel der Integralrechnung.

Die Darstellung $y' = \frac{dy}{dx}$ der Ableitung von y nach x passt offenbar auch sehr gut mit der (sehr wohl exakt begründeten) für partielle Ableitungen wie $\frac{\partial y}{\partial x}$ (y nach x) zusammen. Von partiellen Ableitungen spricht man in der Regel jedoch nur dann, wenn y auch noch von anderen Variablen als x abhängt. Weitere gebräuchliche und oft praktische Bezeichnungen für die partiellen Ableitungen, diesmal einer Funktion f , nach den Variablen x, y, \dots sind f_x, f_y, \dots

Weil viele dieser Notationen aber auch in anderen Zusammenhängen mit entsprechend anderer Bedeutung verwendet werden, die hin und wieder auch mit der Differentialrechnung in Berührung kommen, erleichtert Flexibilität in der Notation das Leben ganz wesentlich. Im vorliegenden Kapitel werden wir uns deshalb diesbezüglich mancherlei Freiheiten erlauben.

3.1.3 Gewöhnliche und partielle Differentialgleichungen

Inhalt in Kurzfassung: In den nachfolgenden Abschnitten werden wir uns auf gewöhnliche DGLen konzentrieren. Das sind solche, wo Ableitungen der gesuchten Funktion nur nach einer Variablen vorkommen. Dennoch spielen vor allem in der Physik auch solche DGLen eine Rolle, wo partielle Ableitungen nach verschiedenen Variablen auftreten und die man partielle DGLen nennt. Anhand von nur zwei Beispielen (schwingende Saite, Wärmeleitungsgleichung) soll angedeutet werden, warum auch sie von Interesse sein können.

Die oben behandelte (sehr einfache) DGL $y' = y$ ist Beispiel einer sogenannten **gewöhnlichen DGL**,¹ weil die gesuchte Funktion y nur von einer Variablen abhängt, nach der abgeleitet wird. Die allgemeine Definition:

Definition 3.1.3.1. *Unter einer gewöhnlichen Differentialgleichung der Ordnung n versteht man eine Gleichung der Bauart*

$$F(y, y', \dots, y^{(n)}, t) = 0.$$

*Dabei steht die Variable y für eine n -mal differenzierbare Funktion in der reellen Variablen t mit Werten in \mathbb{R} (oder auch in \mathbb{R}^m) und $F : D_F \rightarrow \mathbb{R}$ (oder auch $F : D_F \rightarrow \mathbb{R}^k$ im Falle eines Systems von eindimensionalen Differentialgleichungen) mit $D_F \subseteq \mathbb{R}^{n+2}$. Hängt F nicht von t ab, kann also auch $F(y, y', \dots, y^{(n)}) = 0$ mit $D_F \subseteq \mathbb{R}^{n+1}$ geschrieben werden, so heißt die Differentialgleichung **autonom**.*

¹abgekürzt oft ODE für englisch *ordinary differential equation*

3 Differentialgleichungen (DGLen)

Eine n -mal differenzierbare Funktion $f : D_f \rightarrow \mathbb{R}$ (oder $f : D_f \rightarrow \mathbb{R}^m$) heißt **Lösung** der Differentialgleichung auf D_f , wenn

$$F(f(t), f'(t), \dots, f^{(n)}(t), t) = 0$$

für alle $t \in D_f$ gilt.

Meist werden wir, anders als in Definition 3.1.3.1, zwischen der Variablen y für Funktionen und den Funktionen f selbst notationell nicht sorgfältig unterscheiden, sondern nur ein Symbol für beide verwenden, häufig y .

Ein kurze Bemerkung zu autonomen Differentialgleichungen: Sie treten in physikalischen und auch anderen Anwendungen sehr häufig auf. Wird nämlich t als Zeitvariable interpretiert und drückt eine DGL ein unveränderliches Naturgesetz aus, das einen Zusammenhang zwischen verschiedenen Größen beschreibt, das sich aber selbst nicht mit der Zeit ändert, dann hängt die Funktion F in Definition 3.1.3.1 nicht von t ab, sondern nur von den anderen Größen.

Wir werden uns fast ausschließlich mit gewöhnlichen DGLen beschäftigen und stets auch solche meinen, sofern nicht ausdrücklich von **partiellen DGLen**² die Rede ist. Um den Unterschied zu verdeutlichen seien hier nur wenige Bemerkungen dazu eingestreut. Ausführlicher behandelt werden sie in der Lehrveranstaltung Mathematik 3.

In einer partiellen DGL hängt die gesuchte Funktion von mehreren Variablen ab. Gewisse partielle Ableitungen (reine oder gemischte, erster oder höherer Ordnung) nach verschiedenen Variablen werden durch die DGL zueinander in Beziehung gesetzt. Zur Illustration seien hier zwei Beispiele erwähnt.

Die Gleichung der schwingenden Saite als Beispiel: Ein Saite der Länge 2π sei entlang der x -Achse eines Koordinatensystems zwischen dem Ursprung und dem Punkt $(0, 2\pi)$ eingespannt. Die Funktion $s : [0, 2\pi] \times [0, \infty) \rightarrow \mathbb{R}$, $(x, t) \rightarrow s(x, t)$ gebe die Auslenkung der Saite an der Stelle x zum Zeitpunkt t an. Physikalische Überlegungen führen dazu, dass die Kraft, mit der ein kleiner Abschnitt der Saite in einer Umgebung der Stelle x in sehr guter Näherung proportional ist zur Krümmung der Saite, was wieder proportional ist zur zweiten partiellen Ableitung $\frac{\partial^2 s}{\partial x^2}(x, t)$ von s nach der Ortsvariablen x . Nach den Newtonschen Gesetzen ist diese Kraft wiederum proportional zur Beschleunigung, also zur zweiten partiellen Ableitung $\frac{\partial^2 s}{\partial t^2}(x, t)$ von s nach der Zeitvariablen t . Wenn man die Einheiten geeignet wählt führt das zur Gleichung

$$\frac{\partial^2 s}{\partial x^2}(x, t) = \frac{\partial^2 s}{\partial t^2}(x, t)$$

der schwingenden Saite, die offensichtlich keine gewöhnliche, sondern eine partielle DGL ist. Für die Funktionen $s_n(x, t) = \sin(nx) \sin(nt)$, $n \in \mathbb{N}$, gilt

$$\frac{\partial^2 s_n}{\partial x^2}(x, t) = -n^2 \sin(nx) \sin(nt) = \frac{\partial^2 s_n}{\partial t^2}(x, t).$$

²abgekürzt oft PDE für englisch *partial differential equation*

Also sind alle s_n Lösungen dieser partiellen DGL. Weil (auch zweimaliges) Differenzieren linear ist, sind sogar alle Linearkombinationen der s_n Lösungen. Tiefer gehende Analysen zeigen, dass sogar „unendliche Linearkombinationen“

$$\sum_{n=1}^{\infty} c_n \sin(nx) \sin(nt)$$

Lösungen der Gleichungen sind, sofern die Koeffizienten c_n (die sogenannten **Fourierkoeffizienten**) mit wachsendem n schnell klein werden. Sehr tief liegende Theorien garantieren, dass tatsächlich alle Lösungen der Gleichung der schwingenden Saite diese Gestalt haben. (In diesem Zusammenhang sei auf die Theorie der sogenannten **Fourierreihen** verwiesen.)

Die Interpretation dieses Ergebnisses ist sehr interessant: Man hat sich lediglich klar zu machen, was die Funktion $s_n(x, t) = \sin(nx) \sin(nt)$ für festes n beschreibt. Der erste Faktor $\sin(nx)$ beschreibt die räumliche Auslenkung der Saite in Form des Graphen der Sinusfunktion, allerdings mit n Perioden innerhalb des Intervalls $[0, \pi]$, in dem die Saite eingespannt ist. Der zweite Faktor $\sin(nt)$ beschreibt die Veränderung im Laufe der Zeit und besagt, dass die Saite innerhalb eines Zeitintervalls der Länge 2π genau n volle Schwingungen erfährt. Denkt man an die Saite eines Musikinstruments, so bedeutet dies, dass ein Oberton erklingt. Seine Frequenz ist die n -fache Grundfrequenz der Saite. Die Fourierkoeffizienten geben also an, wie stark jeder der Obertöne erklingt, was wir als charakteristische Klangfarbe wahrnehmen.

Beim Beispiel der schwingenden, an ihren Endpunkte aber fest eingespannten Saite kamen nur solche Lösungen s in Frage, die an den Randpunkten 0 und 2π des Intervalls den Wert 0 annehmen, also $s(0, t) = s(2\pi, t) = 0$. Man spricht auch von einem **Randwertproblem**. Ist nur für eine Stelle („Anfangspunkt“) ein Wert vorgegeben, so spricht man von einer Anfangsbedingung bzw. von einem **Anfangswertproblem**.

Als ein zweites Beispiel einer partiellen DGL sei, allerdings ohne ihre Lösungen zu besprechen, die **Wärmeleitungsgleichung**

$$\frac{\partial h}{\partial t} = \frac{\partial^2 h}{\partial x^2}$$

erwähnt. Wie bei der schwingenden Saite denken wir an einen idealisiert als x -Achse verlaufenden Wärmeleiter, für den $h(x, t)$ die Temperatur an der Stelle x zum Zeitpunkt t bezeichnet. (Wieder sind die Einheiten geeignet zu wählen, andernfalls wäre noch eine multiplikative Konstante zu ergänzen.) Gesucht ist eine Funktion h in einer Orts- und einer Zeitvariablen, die obige Gleichung löst. Man mache sich die physikalische Interpretation klar: Die linke Seite entspricht der Änderung der Temperatur pro Zeiteinheit. Diese wird umso stärker sein, je stärker die Temperatur sich nach beiden Richtungen in dieselbe Richtung verändert, was wiederum proportional zur zweiten Ableitung von h nach der Ortskoordinate x ist, also zur rechten Seite der Gleichung. Diese rechte Seite ist hingegen $= 0$, wenn die Temperatur nach der einen Seite im selben Maß abfällt, wie

3 Differentialgleichungen (DGLen)

sie nach der anderen Seite zunimmt, was insgesamt ein Gleichbleiben der Temperatur, also ein Verschwinden der Ableitung nach der Zeit (linke Seite) bewirkt.

In jedem der beiden Beispiele war eine Funktion gesucht, die sowohl von einer Ortskoordinate (hier bezeichnet mit x) und einer Zeitkoordinate (hier bezeichnet mit t) abhängt. Das ist typisch für partielle DGLen. Natürlich sind auch andere Situationen denkbar, wie beispielsweise mehrere Ortskoordinaten. Aufgrund der Invarianz physikalischer Gesetze gegenüber orthogonalen Transformationen kann man sehr oft Symmetrien ausnutzen, insbesondere, wenn man zum Beispiel Kugelkoordinaten (r, α, β) statt kartesische Koordinaten (x, y, z) verwendet. Auf diese Weise gelingt es oft, die Abhängigkeiten von α und β zu eliminieren, so dass r als einzige Variable übrig bleibt und die partielle auf eine gewöhnliche DGL zurückgeführt werden kann, die in der Regel leichter zu lösen ist.

Übungsaufgabe 179. (E) Geben Sie eine formale Definition, was man unter einer partiellen Differentialgleichung und einer Lösung einer partiellen Differentialgleichung versteht. Orientieren Sie sich dabei an Definition 3.1.3.1.

3.1.4 Ordnung versus Dimension – gekoppelte Systeme

Inhalt in Kurzfassung: Treten von der in einer DGL gesuchten Funktion y höhere Ableitungen bis hin zu $y^{(n)}$, so heißt n die Ordnung der DGL. Unter der Dimension verstehen wir die Dimension m des Wertebereichs von y . Will man eine DGL lösen, so machen große Werte von n und auch m die Aufgabe in der Regel schwieriger. Es zeigt sich aber, dass sozusagen ein Abtausch möglich ist: Durch Erhöhung der Dimension m lässt sich die Ordnung n auf 1 drücken. In vielen Fällen erweist sich dieser Abtausch als ein entscheidender Schritt in Richtung Lösung der DGL.

DGLen höherer Ordnung sind natürlich komplizierter als solche erster Ordnung. Es ist möglich und oft auch zielführend, die Ordnung zu reduzieren, indem man dafür eine höhere Dimension des Systems in Kauf nimmt.

Sucht man eine reelle Funktion y als Lösung einer DGL der Ordnung n , so bedeutet dies, dass sowohl y als auch alle Ableitungen $y', y'', \dots, y^{(n)}$ in der DGL auftreten können. Man kann sich auf den Standpunkt stellen, dass zunächst einmal $n + 1$ verschiedene Funktionen y_0, y_1, \dots, y_n auf demselben Definitionsbereich gesucht sind. Dabei denken wir an $y_i = y^{(i)}$ für $i = 1, \dots, n$. Deshalb unterliegen diese Funktionen den zusätzlichen Bedingungen $y'_i = y_{i+1}$ für $i = 0, \dots, n - 1$.

Lässt sich die gegebene DGL in der Form

$$y^{(n)} = F(y, y', \dots, y^{(n-1)}, t),$$

also die höchste vorkommende Ableitung $y^{(n)}$ explizit durch die niedrigeren ausdrücken, so hat man es mit einem Spezialfall eines sogenannten **gekoppelten Systems von DGLen** zu tun, allerdings nicht mehr von der Ordnung n , sondern erster Ordnung. Bei

obiger Umbenennung liegt nämlich das System

$$\begin{aligned} y_0'(t) &= y_1(t) \\ y_1'(t) &= y_2(t) \\ &\vdots \\ y_{n-2}'(t) &= y_{n-1}(t) \\ y_{n-1}'(t) &= F(y_0(t), y_1(t), \dots, y_{n-1}(t), t) \end{aligned}$$

vor. Wir fassen die gesuchten Funktionen $y_i : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}$, $t \mapsto y_i(t)$ zu einer Funktion $\mathbf{y} = (y_0, \dots, y_{n-1}) : D \rightarrow \mathbb{R}^n$, $t \mapsto (y_0(t), \dots, y_{n-1}(t))$ zusammen. Damit ist unser System von der allgemeineren Form

$$\mathbf{y}' = \mathbf{F}(\mathbf{y}, t) := \begin{pmatrix} F_1(y_0(t), y_1(t), \dots, y_{n-1}(t), t) \\ F_2(y_0(t), y_1(t), \dots, y_{n-1}(t), t) \\ \vdots \\ F_n(y_0(t), y_1(t), \dots, y_{n-1}(t), t) \end{pmatrix}$$

mit gewissen $F_i : D_i \rightarrow \mathbb{R}$, $D_i \subseteq \mathbb{R}^{n+1}$, hier mit $F_i(y_1(t), y_2(t), \dots, y_n(t), t) = y_i(t)$ für $i = 1, \dots, n-1$ und $F_n = F$. Weil hierin \mathbf{y}' explizit dargestellt wird, spricht man von einer expliziten DGL, der wir im nächsten Unterabschnitt noch besondere Aufmerksamkeit schenken werden. Die Dimension (die Anzahl der eindimensionalen y_i) ist zwar auf n gestiegen, dafür ist die Ordnung nur mehr 1. Offensichtlich ist dieses Verfahren auch auf DGLen anwendbar, wo sowohl Ordnung als auch Dimension größer als 1 sind.

Die Beschäftigung mit höherdimensionalen DGLen ist in den Anwendungen unvermeidlich, wenn man zum Beispiel Vorgänge im dreidimensionalen Raum beschreiben möchte. Auch historisch hat die Differentialrechnung bei Newton einen Ursprung bei der mathematischen Beschreibung von Planetenbewegungen (Keplersche Gesetze), die nicht eindimensional zu erfassen sind. Wie wir sehen werden, gelten manche sehr wichtige Überlegungen und auch Resultate höherdimensional ähnlich wie eindimensional. In den vertiefenden Untersuchungen in Abschnitt ?? über lineare DGLen wird sich der Nutzen dieses Zugangs zeigen.

Übungsaufgabe 180. (P) Schreiben Sie, wenn möglich, in äquivalente Systeme erster Ordnung um (die Gleichungen selbst brauchen nicht gelöst werden):

- | | |
|-----------------------------------|-------------------------|
| 1. $xy''' - x^2y'' + xy' - y = 0$ | 2. $y^2 + 1 = 0$ |
| 3. $y^{(4)} + x^2y' - 9y = 0$ | 4. $y' + 5y = 0$ |
| 5. $8y'' - 9y = 0$ | 6. $y'' - 2y' + 5y = 0$ |
| 7. $y^{(4)} + 8y'' - 9y = e^x$ | 8. $y + 8y'' - 9y = 0$ |

3.1.5 Existenz und Eindeutigkeit von Lösungen expliziter DGLen

Inhalt in Kurzfassung: Will man eine DGL lösen, so lautet die vorgelagerte Frage: Gibt es überhaupt eine Lösung und, wenn ja, ist diese Lösung eindeutig? Im Fall von expliziten

3 Differentialgleichungen (DGLen)

DGLen (das sind solche der Form $y'(t) = f(y, t)$) mit einer Anfangsbedingung $y(t_0) = y_0$ können sowohl Existenz als auch Eindeutigkeit einer lokalen Lösung (das heißt in einer Umgebung von t_0) garantiert werden, sofern folgende Zusatzbedingung erfüllt ist: Die in der DGL auf der rechten Seite auftretende Funktion f ist stetig und in der Variablen y sogar Lipschitz-stetig. Dieser Satz ist der mit Abstand wichtigste in der allgemeinen Theorie der DGLen – jedenfalls so weit wir sie werden kennen lernen. Der Beweis fußt auf dem Kontraktionsprinzip (dem Banachschen Fixpunktsatz) aus Mathematik 1.

In der folgenden Diskussion konzentrieren wir uns auf DGLen erster Ordnung. Wie man in vielen wichtigen Fällen bei höherer Ordnung vorgehen kann, wenn man höhere Dimension in Kauf nimmt, haben wir in 3.1.4 gesehen. Trotzdem besprechen wir der einfacheren Notation halber den Fall der Dimension 1. Hat man ihn verstanden, so fällt es nicht schwer, die Ergebnisse auf höhere Dimensionen zu übertragen.

Definition 3.1.5.1. *Unter einer **expliziten Differentialgleichung** (erster Ordnung) versteht man eine gewöhnliche Differentialgleichung, die in der Form*

$$y' = f(t, y)$$

mit $f : D_f \rightarrow \mathbb{R}$, $D_f \subseteq \mathbb{R}^2$ dargestellt werden kann.

Derselbe Begriff wird analog für Funktionen $\mathbf{y} = (y_1, \dots, y_n)$, $y_i : D_f \rightarrow \mathbb{R}$, mit n -dimensionalem Wertebereich von \mathbf{y} geprägt.

Häufig treten explizite Differentialgleichungen mit Anfangsbedingungen der Gestalt $y(t_0) = y_0$ auf, wobei (t_0, y_0) ein innerer Punkt von D_f ist.

Setzt man Stetigkeit von f voraus, so lässt sich zeigen, dass eine explizite Differentialgleichung mit Anfangsbedingung stets mindestens eine Lösung hat. Beweisen wollen wir jenen verwandten Satz, der unter einer etwas stärkeren Bedingung an f sogar die Eindeutigkeit der Lösung garantiert. Diese Bedingung ist neben der Stetigkeit von f als Funktion in zwei Variablen die sogenannte **Lipschitz-Stetigkeit** von f in der zweiten Komponente, die folgendes bedeutet: Es gibt eine reelle Zahl λ mit der Eigenschaft, dass für alle $(t, y_1), (t, y_2) \in D_f$ die Ungleichung $|f(t, y_1) - f(t, y_2)| \leq \lambda|y_1 - y_2|$ gilt. Diese sogenannte **Lipschitz-Bedingung** ist sehr häufig erfüllt, jedoch keineswegs in allen interessanten Fällen. Ein sehr einfaches Beispiel: Die DGL $y' = y$ für die Exponentialfunktion entspricht der Funktion $f(t, y) = y$, und erfüllt die Lipschitzbedingung mit $\lambda = 1$. Allgemeiner folgt die Lipschitzbedingung in einem kompakten (d.h. sowohl abgeschlossen als auch beschränkten) Bereich D_f , wenn f dort stetig differenzierbar ist. Denn nach dem Satz vom Maximum muss dann auch die partielle Ableitung $\frac{\partial f}{\partial y}$ von f nach der zweiten Variablen betragsmäßig ein Maximum λ annehmen. Nach dem Mittelwertsatz gibt es eine Stelle η zwischen y_1 und y_2 mit

$$f(t, y_1) - f(t, y_2) = \frac{\partial f}{\partial y}(t, \eta)(y_2 - y_1),$$

woraus

$$|f(t, y_1) - f(t, y_2)| \leq \lambda|y_1 - y_2|,$$

also die Lipschitz-Bedingung folgt. Ein typisches Beispiel einer reellen Funktion f , die auf einem abgeschlossenen Intervall stetig, aber nicht Lipschitz-stetig ist, ist $f : [0, 1] \rightarrow \mathbb{R}$, $f(x) := \sqrt{x}$. Diese Funktion hat zwar auf $(0, 1]$ die stetige Ableitung $f'(x) = \frac{1}{2\sqrt{x}}$, ist aber am Randpunkt $x = 0$ nicht differenzierbar.

Der angekündigte **Existenz- und Eindeigkeitsatz**, der auch als **Satz von Picard-Lindelöf** bekannt ist, lautet:

Satz 3.1.5.2. *Sei die explizite DGL*

$$y' = f(t, y)$$

mit der Anfangsbedingung $y(t_0) = y_0$ gegeben. Ist $f : D_f \rightarrow \mathbb{R}$, $D_f \subseteq \mathbb{R}^2$, stetig und in der zweiten Komponente Lipschitz-stetig, so gibt es eine Umgebung von t_0 , auf der die Differentialgleichung genau eine Lösung hat. Diese Lösung kann man als Grenzwert geeigneter Iterationsfolgen erhalten.

Der Satz gilt analog für Gleichungen für Funktionen $\mathbf{y} = (y_1, \dots, y_n)$, $y_i : D_f \rightarrow \mathbb{R}$ mit n -dimensionalem Wertebereich.

Der Beweis verwendet das **Kontraktionsprinzip** (den **Banachschen Fixpunktsatz**) aus Mathematik 1. Anhand eines Beispiels haben wir in Mathematik 1 die dem Beweis von Satz 3.1.5.2 zugrundeliegende Methode bereits kennengelernt. Und zwar sind wir auf die Exponentialfunktion als Lösung der DGL $y' = y$ mit der Anfangsbedingung $y(0) = 1$ gestoßen, indem wir mit der konstanten Funktion $y_0(t) := 1$ für alle $t \in \mathbb{R}$ begonnen haben und rekursiv $y_{n+1}(t) := 1 + \int_0^t y_n(x) dx$ gesetzt haben. Wir haben also eine Transformation T verwendet, die einer stetigen Funktion y die Funktion $T_y(t) := 1 + \int_0^t y(x) dx$ zugeordnet und $y_{n+1} := T_{y_n}$ gesetzt haben. Damals hat man sehr schnell $y_n(t) = \sum_{k=0}^n \frac{t^k}{k!}$ gesehen, also

$$\lim_{n \rightarrow \infty} y_n(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} = \exp(t).$$

Bei weniger einfach gebauten DGLen kann man nicht hoffen, dass man die y_n so leicht explizit bestimmen und daraus ihren Grenzwert ablesen kann. Wenn man konzeptionell vorgeht, führt dieselbe Grundidee aber auch beim allgemeinen Satz 3.1.5.2 zum Ziel:

Beweis. (Idee) Beim Beweis von Satz 3.1.5.2 betrachtet man statt Folgen reeller Zahlen Folgen von Funktionen aus einer geeigneten Menge C (für *continuous* = stetig) stetiger reeller Funktionen. Man definiert eine Transformation T , die jeder Funktion $y \in C$ eine andere Funktion $T_y \in C$ zuordnet, so dass folgendes gilt:

1. Auf der Menge C ist eine Metrik d definiert, so dass C zu einem vollständigen metrischen Raum wird, was bedeutet, dass jede Cauchyfolge konvergiert. Damit gilt das Kontraktionsprinzip auf C aufgrund des völlig analogen Beweises wie für \mathbb{R} .

3 Differentialgleichungen (DGLen)

2. C, d und T sind so definiert, dass $T : C \rightarrow C, y \mapsto T_y$ eine Kontraktion ist. Nach dem Kontraktionsprinzip konvergiert daher jede Iterationsfolge bezüglich T gegen ein $y \in C$, das Fixpunkt von T ist.
3. T ist überdies so definiert, dass die Fixpunkte von T genau die Lösungen der gegebenen Differentialgleichung mit Anfangsbedingung sind. Folglich gibt es tatsächlich genau eine Lösung wie behauptet, die überdies Grenzwert beliebiger Iterationsfolgen in C bezüglich T ist.

Es sind also lediglich C, d und T geeignet zu definieren und die behaupteten Eigenschaften nachzuweisen.

Definition von C : C enthalte als Elemente alle stetigen Funktionen

$$y : [t_0 - \delta, t_0 + \delta] \rightarrow \mathbb{R}.$$

Dabei ist $\delta > 0$ hinreichend klein zu wählen. Wie klein, das wird sich aus dem Beweis ergeben. Wir schreiben ab nun $I := [t_0 - \delta, t_0 + \delta]$.

Definition von d und Vollständigkeit: Für $y_1, y_2 \in C$ nimmt die stetige Funktion $t \mapsto |y_1(t) - y_2(t)|$ auf dem abgeschlossenen Intervall I ein Maximum an, welches wir mit $d(y_1, y_2)$ bezeichnen. Die Funktion $d : (y_1, y_2) \mapsto d(y_1, y_2)$ ist offenbar eine Metrik, denn es gilt: $d(y_1, y_2) \geq 0$ und $d(y_1, y_2) = 0$ nur für $y_1 = y_2$; $d(y_1, y_2) = d(y_2, y_1)$; $d(y_1, y_3) \leq d(y_1, y_2) + d(y_2, y_3)$. Wir beobachten, dass eine Folge von Funktionen $y_n \in C$ genau dann bezüglich dieser Metrik d gegen $y \in C$ konvergiert, wenn die y_n als reelle Funktionen gleichmäßig gegen y konvergieren. Aus Mathematik 1 ist bekannt, dass der gleichmäßige Grenzwert stetiger Funktionen wieder stetig ist. Folglich liegt y ebenfalls in C . Für jede Cauchyfolge von Funktionen $y_n \in C$ bezüglich d bilden für alle $t \in I$ die $y_n(t)$ eine Cauchyfolge in \mathbb{R} , konvergieren daher gegen ein $y(t)$. Das sich daraus ergebende $y : t \mapsto y(t)$ ist (Übung) sogar gleichmäßiger Grenzwert der y_n , nach obiger Überlegung also selbst in C . Somit ist der gesuchte Grenzwert gefunden, C ist bezüglich d also tatsächlich ein vollständiger metrischer Raum.

Definition und Eigenschaften von $T : y \mapsto T_y$: Für $y \in C$ definieren wir die Funktion T_y durch:

$$T_y : t \mapsto y_0 + \int_{t_0}^t f(x, y(x)) dx$$

Nach dem Hauptsatz der Differential- und Integralrechnung ist $T_y'(t) = f(t, y(t))$. Ist die Funktion $y \in C$ ein Fixpunkt von T , so bedeutet das $T_y = y$, ausführlich:

$$y(t) = y_0 + \int_{t_0}^t f(x, y(x)) dx$$

Leitet man diese Gleichung ab und verwendet man rechts den Hauptsatz der Differential- und Integralrechnung, so erhält man $y'(t) = f(t, y(t))$, außerdem $y(t_0) = T_y(t_0) = y_0 + \int_{t_0}^{t_0} f(x, y(x)) dx = y_0$. Somit ist y eine Lösung der gegebenen DGL. Weil offenbar umgekehrt auch jede Lösung der DGL ein Fixpunkt von T ist, sind tatsächlich die Lösungen der DGL genau die Fixpunkte von T . Um die Konvergenz der Iterationsfolge

$y_{n+1} := T_{y_n}$ zu garantieren, verbleibt der Nachweis der Kontraktionseigenschaft und die Bestimmung von δ . Hierzu verwenden wir die vorausgesetzte Lipschitz-Stetigkeit von f in der zweiten Komponente: Es gibt also ein $\lambda_0 \in \mathbb{R}$ mit

$$|f(x, y_1(x)) - f(x, y_2(x))| \leq \lambda_0 |y_1(x) - y_2(x)| \leq \lambda_0 d(y_1, y_2)$$

für alle $x \in I = [t_0 - \delta, t_0 + \delta]$. Daraus folgt für beliebiges $t \in I$:

$$\begin{aligned} |T_{y_1}(t) - T_{y_2}(t)| &= \left| \left(y_0 + \int_{t_0}^t f(x, y_1(x)) dx \right) - \left(y_0 + \int_{t_0}^t f(x, y_2(x)) dx \right) \right| \\ &\leq \int_{t_0}^t \lambda_0 d(y_1, y_2) dx \leq \delta \lambda_0 d(y_1, y_2). \end{aligned}$$

Weil dies für alle $t \in I$ gilt, bedeutet das

$$d(T_{y_1}, T_{y_2}) \leq \lambda d(y_1, y_2)$$

für $\lambda := \delta \lambda_0$. Wir haben also lediglich $\delta < \lambda_0^{-1}$ zu wählen, um garantieren zu können, dass T eine Kontraktion auf C ist, wie behauptet wurde. \square

Übungsaufgabe 181. (E) *Vervollständigen Sie den Beweis von Satz 3.1.5.2 in zwei Schritten:*

1. *Rekapitulieren Sie den Beweis und halten Sie jene Behauptungen fest, die dort nicht im Detail bewiesen wurden.*
2. *Ergänzen Sie diese Details.*

Satz 3.1.5.2 garantiert (unter den angegebenen Voraussetzungen) eine eindeutige Lösung nur in einem kleinen Intervall der Länge $< 2\lambda^{-1}$. In vielen Fällen lässt sich die Eindeutigkeitsaussage aber recht leicht ausdehnen. Das sei kurz am Beispiel der Gleichung $y' = y$ (also $f(x) = x$), $y(0) = 1$, für die Exponentialfunktion illustriert.

Hat man eine Lösung gefunden (in unserem Fall $y = \exp$), so folgt aus Satz 3.1.5.2, dass diese Lösung wenigstens auf einem Intervall $(-\delta, \delta)$ eindeutig ist. Im Beispiel kann $\delta = 1$ gewählt werden, weil $\lambda = 1$ Lipschitz-Konstante für f ist. Also ist \exp auf $(-1, 1)$ die einzige Lösung. Diese lässt sich auch auf die Randpunkte -1 und 1 stetig fortsetzen. Nun kann mit der DGL $y = y'$ und der neuen Anfangsbedingung $y(1) = \exp(1)$ (bzw. nach links hin: $y(-1) = \exp(-1)$) weitergearbeitet werden. Wieder ergibt sich eine eindeutige Lösung, diesmal auf $(0, 2)$ bzw. auf $(-2, 0)$. Insgesamt ist die Lösung \exp also auch auf $(-2, 2)$ eindeutig etc. In diesem Fall kann global mit derselben Lipschitz-Konstanten gearbeitet werden. Für viele interessante explizite DGLen $y' = f(y)$ mit Anfangsbedingung $y(t_0) = y_0$ gibt es für f zwar keine globale Lipschitz-Konstante λ , es genügt aber lokale Lipschitz-Stetigkeit, um Eindeutigkeit wenigstens in einem kleinen Intervall zu gewährleisten. Damit ist folgende schwächere Bedingung gemeint: Um jeden Punkt x gibt es ein Intervall $[x - \varepsilon, x + \varepsilon]$, auf dem f Lipschitz-stetig (mit einer von diesem Intervall abhängigen Lipschitz-Konstanten) ist. Dann kann man folgendermaßen argumentieren: Angenommen y_1 und y_2 seien zwei Lösungen der DGL auf einer offenen

3 Differentialgleichungen (DGLen)

Menge $O \subseteq \mathbb{R}$. Nach Satz 3.1.5.2 stimmen diese beiden Lösungen wenigstens auf einem Intervall um t_0 überein. Sei s das Supremum über alle $x > t_0$, so dass y_1 und y_2 auf ganz $[t_0, x)$ übereinstimmen. Als differenzierbare Funktionen sind y_1 und y_2 erst recht stetig. Liegt s in O , folgt daraus $y_1(s) = y_2(s)$. Dann lässt sich Satz 3.1.5.2 auf die DGL $y' = f(y)$ mit der neuen Anfangsbedingung $y(s) = y_1(s) = y_2(s)$ anwenden, woraus folgt, dass die Funktionen y_1 und y_2 auch noch auf einem etwas größeren Intervall $[t_0, s + \delta]$ übereinstimmen, was der Definition von s widerspricht.

Die Eindeutigkeit der Lösungen kann also nur dann verloren gehen, wenn die offene Menge O verlassen wird. Typischerweise tritt so eine Situation dann auf, wenn eine Lösung y an einer Stelle s einen Pol hat und entsprechend s ein Randpunkt von O ist.

Verzichtet man in Satz 3.1.5.2 auf die Lipschitz-Bedingung an f , so kann die Eindeutigkeit der Lösung auch lokal verloren gehen. Ein einfaches Beispiel ist die an der Stelle 0 nicht Lipschitz-stetige Funktion $f(y) = \sqrt{y}$. Und tatsächlich zeigt das Beispiel der DGL $y' = \sqrt{y}$ mit der Anfangsbedingung $y(0) = 0$, dass dann auch die Eindeutigkeitsaussage von Satz 3.1.5.2 falsch wird. Denn jede der Funktionen y_α mit $\alpha \geq 0$, definiert durch

$$y_\alpha(t) := 0 \text{ für } t \leq \alpha \quad \text{und} \quad y_\alpha(t) := \frac{1}{4}(t - \alpha)^2 \text{ für } t > \alpha$$

ist (wie übrigens auch die Nullfunktion) Lösung der DGL.

Übungsaufgabe 182. (P) Überprüfen Sie, dass die y_α tatsächlich für alle $\alpha \geq 0$ Lösungen der DGL $y' = \sqrt{y}$ sind.

Die Erklärung der Mehrdeutigkeit der Lösung muss angesichts des Satzes 3.1.5.2 in einer Verletzung der Lipschitzbedingung liegen:

Übungsaufgabe 183. (E) Zeigen Sie direkt, dass die Funktion $x \mapsto \sqrt{x}$, $x \geq 0$, in der Nähe von $x = 0$ nicht Lipschitz-stetig ist.

Wie bereits in der Formulierung von Satz 3.1.5.2 erwähnt, gilt er auch für höhere Dimensionen.

Übungsaufgabe 184. (E) Formulieren Sie Satz 3.1.5.2 entsprechend für höhere Dimensionen. Erklären Sie dabei sorgfältig, welche Modifikationen nötig sind.

Diese Übertragung auf höhere Dimensionen ist von großer Bedeutung. Zum Beispiel führt die klassische Newtonsche Mechanik sehr oft zu DGLen, auf die der Existenz- und Eindeutigkeitssatz anwendbar ist. Geht es beispielsweise um eine Funktion $y(t)$, die den Ort eines Teilchens (Körpers, Planeten) im dreidimensionalen Raum zum Zeitpunkt t beschreibt, dann sind $y'(t)$ und $y''(t)$ die (vektorwertige) Geschwindigkeit bzw. Beschleunigung. Besagen die physikalischen Gesetze, dass y Lösung einer DGL ist, welche die Voraussetzungen von Satz 3.1.5.2 erfüllt, so besagt die klassische Physik, dass die weitere Bewegung durch die Anfangsbedingungen, d.h. durch die aktuellen Messwerte der relevanten Größen (zum Beispiel Ort und Geschwindigkeit) vollständig determiniert ist. Diese Erkenntnisse aus Physik und Mathematik ließen im 18. und 19. Jahrhundert viele

führende Köpfe in der Wissenschaft (allen voran ist in diesem Zusammenhang der Name Laplace zu nennen) philosophisch einem sehr weitreichenden Determinismus anhängen. Erst durch die moderne Quantenphysik erfuhr diese Sichtweise wieder beträchtliche Einschränkungen.

3.1.6 Qualitative Untersuchungen autonomer DGLen

Inhalt in Kurzfassung: Eine explizite DGL $y' = f(y, t)$, in der f nicht von t abhängt, also $y' = f(y)$ heißt autonom. Interpretiert man t als Zeitvariable, so beschreibt die gesuchte Lösung $y = y(t)$ einen Ablauf in der Zeit, wobei die Gesetzmäßigkeit, der dieser Ablauf folgt, durch f beschrieben wird. Im Fall eine autonomen DGL beschreibt dieses Gesetz, wie sich die momentane Änderung y' aus dem momentanen Systemzustand y ergibt, der Mechanismus (das Naturgesetz) jedoch ändert sich mit der Zeit nicht. Diese Zeitunabhängigkeit ermöglicht qualitative Untersuchungen von Lösungen autonomer DGLen, die im nicht autonomen Fall unmöglich wären.

Ist es zu aufwendig oder gar unmöglich, Lösungen von DGLen explizit zu berechnen, geben oft schon qualitative Überlegungen interessante Einsichten. Insbesondere ist das bei autonomen DGLen der Fall. Für eine DGL der Gestalt $y' = f(y)$ gibt in typischen Fällen nämlich bereits das Vorzeichenverhalten von f wertvolle Hinweise. Interpretieren wir die unabhängige Variable t als Zeitvariable, so markiert der Funktionsverlauf $y(t)$ eine (reellwertige) Bewegung, die wir auf der Zahlengeraden andeuten werden, indem wir solche Bereiche, wo Lösungen der DGL gewissermaßen eine Bewegung nach links vollführen unterscheiden von solchen, wo sich $y(t)$ nach rechts bewegt. Die resultierenden Darstellungen nennt man auch **Phasendiagramme**. Die Funktion f sei durchwegs als stetig vorausgesetzt, der Definitionsbereich $D \subseteq \mathbb{R}$ von $f : D \rightarrow \mathbb{R}$ als zusammenhängend.

Übungsaufgabe 185. Überlegen Sie, was es bedeuten würde, wenn f unstetig wäre.

Wechselt f auf ganz D nicht das Vorzeichen, so muss jede Lösung y der DGL wegen $y'(t) = f(y(t))$ monoton sein, wachsend für $f \geq 0$, fallend für $f \leq 0$. Im ersten Fall bewegt sich $y(t)$ auf ganz D nach rechts, im zweiten nach links. Nicht jedoch lässt sich aus dieser Information allein schließen, ob y für wachsendes t beschränkt oder unbeschränkt ist. Diese Unbestimmtheit wird durch die folgenden beiden einfachen Beispiele belegt.

Beispiel 1: Die Funktion $y(t) := -\frac{1}{t}$ ist eine auf $D = \mathbb{R}^+$ definierte und nach oben beschränkte Lösung der DGL $y' = f(y) := y^2 \geq 0$.

Beispiel 2: Die Funktion $y(t) := t$ ist eine (sogar die eindeutige) auf $D = \mathbb{R}$ definierte unbeschränkte Lösung der (ziemlich trivialen) DGL $y' = f(y) := 1$ mit der Anfangsbedingung $y(0) = 0$.

Nehmen wir nun an, f wechsele bei y_0 das Vorzeichen, genauer: Es gibt ein $\delta > 0$ mit $f(y_1) \cdot f(y_2) < 0$ für alle y_1, y_2 mit $y_0 - \delta < y_1 < y_0 < y_2 < y_0 + \delta$. Weil f stetig ist, ist das nur für $f(y_0) = 0$ möglich. Die beiden denkbaren Fälle unterscheiden sich qualitativ. Sei y eine Lösung der DGL $y' = f(y)$.

3 Differentialgleichungen (DGLen)

Fall 1, $f(y_1) < 0$ und $f(y_2) > 0$: Ist $y(t_1) = y_1$, so folgt $y'(t_1) = f(y(t_1)) = f(y_1) < 0$. In der Nähe von t_1 ist y also monoton fallend, oder anders formuliert: In eine Umgebung links von y_0 bewegt sich $y(t)$ mit wachsendem t von y_0 weg. Analog bewegt sich $y(t)$ unmittelbar rechts von y_0 nach rechts, also gleichfalls von y_0 weg. Der Punkt y_0 selbst ist ein sogenannter **stationärer Punkt**: Die konstante Funktion $y(t) = y_0$ wäre wegen $0 = y'(t) = f(y(t)) = f(y_0) = 0$ Lösung der DGL. Man beachte, dass es – so wie bei früheren Beispielen – auch andere Lösungen geben kann, sofern f nicht Lipschitz-stetig ist.

Übungsaufgabe 186. Finden Sie ein solches Beispiel zu Fall 1, d.h. eine DGL der Form $y' = f(y)$ und Werte t_0 und x_0 mit mehr als einer Lösung der DGL $y' = f(y)$, die zusätzlich $y(t_0) = x_0$ erfüllt.

Fall 2, $f(y_1) > 0$ und $f(y_2) < 0$: Ganz analoge Überlegungen wie zu Fall 1 zeigen, dass y_0 auch in Fall 2 ein stationärer Punkt ist, allerdings mit einer Bewegung auf y_0 zu, und zwar von beiden Seiten.

Im Lichte der beiden behandelten Fälle unterscheidet man stationäre Punkte unterschiedlicher Art:

Definition 3.1.6.1. Für eine explizite und autonome DGL $y' = f(y)$ mit stetigem $f : D \rightarrow \mathbb{R}$, $[0, \infty) \subseteq D \subseteq \mathbb{R}$ offen und zusammenhängend, heißt $y_0 \in D$ ein **stationärer Punkt**, wenn $f(y_0) = 0$ oder, äquivalent, wenn die konstante Funktion $y = y_0$ eine Lösung der DGL ist. Der stationäre Punkt y_0 heißt:

anziehend, wenn es ein $\delta > 0$ gibt derart, dass $\lim_{t \rightarrow \infty} y(t) = y_0$ für jede Lösung y der DGL mit $|y(0) - y_0| < \delta$.

stabil, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass für jede Lösung y der DGL mit $|y(0) - y_0| < \delta$ für alle $t \geq 0$ die Beziehung $|y(t) - y_0| < \varepsilon$ folgt.

instabil, wenn y_0 nicht stabil ist, d.h. wenn es ein $\varepsilon > 0$ gibt derart, dass es zu jedem $\delta > 0$ eine Lösung y der DGL mit $|y(0) - y_0| < \delta$ und ein $t \geq 0$ mit $|y(t) - y_0| \geq \varepsilon$ gibt.

Klarerweise ist jeder anziehende stationäre Punkt auch stabil.

Übungsaufgabe 187. Begründen Sie diese Behauptung.

Anschaulich plausibel ist auch die Umkehrung. Tatsächlich ist ein stationärer Punkt x_0 wie in Fall 2 stets anziehend. Der Beweis ist allerdings einigermaßen anspruchsvoll. Sogar falsch wird die Umkehrung jedoch, wenn man die Konzepte sinnvoll in höhere Dimensionen überträgt. Wie das möglich ist, soll nun kurz angedeutet werden.

Der sehr grundsätzliche Unterschied zwischen der Dimension 1 und allen höheren Dimensionen besteht darin, dass es im Eindimensionalen nur zwei Richtungen gibt – eine positive (monoton wachsend) und eine negative (monoton fallend). Im Gegensatz dazu, kann ein höherdimensionaler Vektor in unendlich viele Richtungen zeigen. Für eine diagrammartige Darstellung einer Differentialgleichung $\mathbf{y}' = f(\mathbf{y}) \in \mathbb{R}^n$ kann man f als Vektorfeld darstellen. Dann ordnet f einem n -dimensionalen Vektor \mathbf{x} den ebenfalls n -dimensionalen Vektor $f(\mathbf{x})$ zu. Eine Lösung $\mathbf{y} : t \mapsto \mathbf{y}(t)$ der gegebenen DGL kann

man sich als eine durch den Zeitparameter t parametrisierte Kurve im \mathbb{R}^n vorstellen, für die in jedem Punkt $\mathbf{y}(t)$ der Vektor $f(\mathbf{y}(t))$ tangential ist. Weil es bei der qualitativen Analyse in erster Linie auf die Richtung ankommt, spricht man in diesem Zusammenhang auch von einem **Richtungsfeld**. Stationäre Punkte \mathbf{x} sind bei $n > 1$ auf sehr unterschiedliche Arten denkbar. Weiterhin sind sowohl anziehende stationäre Punkte denkbar (mit Pfeilen, die in der Nähe von \mathbf{x}_0 annähernd in die Richtung von \mathbf{x}_0 weisen) als auch abstoßende (wo die Pfeile sternförmig von \mathbf{x}_0 weggehen). Darüber hinaus kann sich, bildhaft gesprochen, das Vektorfeld f aber auch spiralförmig an \mathbf{x}_0 annähern oder entfernen, kreisförmig um \mathbf{x}_0 verlaufen etc. Je größer n , desto kompliziertere Situationen sind denkbar.

Übungsaufgabe 188. (T) Skizzieren Sie für die folgenden Differentialgleichungen die zugehörigen Richtungsfelder mit einigen Lösungskurven:

1. $y' = x^2 + y^2 - 1$
2. $y' = 1 - y^2$
3. $y' = 5y - x$

Als Richtungsfeld bezeichnet man jenes Vektorfeld, welches einem Punkt (x, y) einen ebenen Einheitsvektor mit Steigung $y'(x)$ zuordnet.

Übungsaufgabe 189. (T*) Überprüfen Sie durch Einsetzen, dass $y(x) = 2cx - c^2$ eine Lösung der Differentialgleichung $y'^2 - 4xy' + 4y = 0$ ist. Rechnen Sie nach, dass die Geraden $y(x) = 2cx - c^2$ die Tangenten an die Parabel $y = x^2$ sind, und dass diese ebenfalls eine Lösung der obigen Differentialgleichung darstellt.

Übungsaufgabe 190. (T) Skizzieren Sie das zu der Differentialgleichung $y' = ky$ gehörige Richtungsfeld. Unterscheiden Sie dabei die Fälle $k > 0$, $k = 0$, $k < 0$. Lösen Sie die Differentialgleichung $y' = ky$, und vergewissern Sie sich, dass die erhaltene Schar von Lösungskurven in das jeweilige Richtungsfeld passt. Wie wirkt sich eine geringfügige Änderung von k in jedem dieser Fälle auf das qualitative Verhalten der Lösungskurven aus?

Übungsaufgabe 191. (T) Skizzieren Sie das Richtungsfeld für die Differentialgleichung $y' = xy$ und lösen Sie diese Differentialgleichung.

3.2 Lineare Differentialgleichungen

Ähnlich wie schon in der Linearen Algebra (Kapitel 1) erweist sich auch bei Differentialgleichungen Linearität als eine sehr nützliche Eigenschaft, mit der viele Aufgaben wesentlich effektiver gelöst werden können. In einem ersten Schritt wird in Abschnitt 3.2.1 der Existenz- und Eindeutigkeitssatz für explizite Differentialgleichungen auf lineare Differentialgleichungen angewandt, um bereits daraus starke Schlussfolgerungen zu ziehen. Sodann geht es um Unterscheidungen innerhalb Klasse linearer Differentialgleichungen, die gewisse einfache Fälle mit besonders starker Lösungstheorie identifizieren. In 3.2.2 werden die verschiedenen Varianten überblicksartig besprochen, wobei stets eine simple als Spezialfall einer allgemeineren, komplizierteren gegenübergestellt

3 Differentialgleichungen (DGLen)

wird. In 3.2.3 beginnen wir mit den simpelsten Fällen von DGLen (homogen, konstante Koeffizienten), deren Lösungen sich auf sehr einfache Weise aus Exponentialfunktionen, Sinus- und Cosinustermen sowie Polynomen zusammensetzen. Komplikationen, die beim Übergang zu inhomogenen DGLen auftreten, sind Gegenstand von 3.2.4. Lässt man auch nichtkonstante Koeffizienten zu, müssen andere Wege eingeschlagen werden, wobei wir nur DGLen erster Ordnung besprechen, siehe 3.2.5. Ab 3.2.6 kommt auch der mehrdimensionale Fall „gekoppelter“ lineare DGLen zur Sprache (mehrdimensionale Systeme linearer DGLen). Für den wesentlich spezielleren Situation konstanter Koeffizienten steht mit der Eigenwert-Eigenvektormethode ein sehr mächtiges Lösungsverfahren zur Verfügung. Als Ergänzungen und Illustrationen folgen noch einige weitere Unterabschnitte. In den ersten beiden davon (3.2.6 und 3.2.7) geht es um homogene Systeme erster und höherer Ordnung, sodann um inhomogene Systeme (3.2.8), um die Anpassung an Anfangswerte (3.2.9) und schließlich als Wiederholung der Methode aus 3.1.4 darum, wie der Fall einfacher DGLen höherer Ordnung auf den gekoppelten Fall erster Ordnung zurückgeführt wird (3.2.10 und 3.2.11).

3.2.1 Allgemeine Vorbemerkungen

Inhalt in Kurzfassung: Bevor wir uns in den nachfolgenden Abschnitten der expliziten Lösung gewisser linearer DGLen zuwenden werden, soll hier noch unter einem allgemeinen Gesichtspunkt angedeutet werden, warum die linearen unter sämtlichen DGLen einen Sonderstatus besitzen.

Lineare DGLen sind solche Differentialgleichungen, deren Lösungen genau jene Funktionen y aus einem Vektorraum V_1 von Funktionen sind, die durch eine lineare Abbildung $\Delta : V_1 \rightarrow V_2$ (die in diesem Kontext auch **Differentialoperator** heißt) auf ein bestimmtes $a \in V_2$ des Vektorraums V_2 (der auch $= V_1$ sein darf) abgebildet werden. Weil es sich um eine DGL handelt, setzt sich Δ neben y selbst aus (gewöhnlichen oder, im Falle partieller DGLen, aus partiellen) Ableitungen der gesuchten Funktionen zusammen. Weil das sehr abstrakt klingt, soll es an einem einfachen, bereits bekannten Beispiel illustriert werden.

Für die (explizite) Differentialgleichung $y' = y$ können wir als V_1 den Vektorraum aller differenzierbaren Funktionen auf \mathbb{R} wählen, als V_2 den Vektorraum aller reellen Funktionen. Außerdem setzen wir $\Delta : y \mapsto y - y'$ und $a = 0$ (Nullfunktion in V_2). V ist tatsächlich ein Vektorraum und Δ ist linear (Übung). Deshalb ist $\Delta(y) = 0$ sogar eine homogene lineare Gleichung im Sinne der Linearen Algebra, und ihre Lösungsmenge ein Unterraum von V . Man spricht deshalb von einer **homogenen linearen Differentialgleichung**.

Übungsaufgabe 192. (P) Begründen Sie für dieses Beispiel, warum tatsächlich V_1 und V_2 Vektorräume sind, Δ eine lineare Abbildung von V_1 nach V_2 und f Lipschitz-stetig ist.

Es stellt sich die Frage nach der Dimension des Lösungsraums. Im vorliegenden Beispiel lässt sich diese leicht bestimmen. Die Funktion $f : x \mapsto x$, für die unsere DGL die Gestalt $y' = f(y) = y$ annimmt, ist Lipschitz-stetig (mit der Lipschitz-Konstanten $\lambda = 1$), also

gibt es zu jeder vorgegebenen Anfangsbedingung $y(0) = c$, $c \in \mathbb{R}$, eine eindeutige Lösung y_c . Diese ist sofort gefunden, denn $y_c := c \cdot \exp$ erfüllt sowohl $y'_c = (c \cdot \exp)' = c \cdot \exp' = c \cdot \exp = y_c$ als auch die Anfangsbedingung $y_c(0) = c \cdot \exp(0) = c$. Die Lösungsmenge L_0 der homogenen DGL $y' = y$ besteht also genau aus allen Vielfachen $c \cdot \exp$, $c \in \mathbb{R}$, von \exp . Der Lösungsraum ist also eindimensional mit $\{\exp\}$ als Basis.

Nehmen wir statt $a = 0$ als rechte Seite der Gleichung $y - y' = a$ eine von der Nullfunktion verschiedene reelle Funktion $a : t \mapsto a(t)$, so liegt eine **inhomogene lineare Differentialgleichung** der Gestalt $\Delta(y) = a$ vor. Ist y_p irgendeine (sogenannte **partikuläre**) Lösung und L_0 der Lösungsraum der zugehörigen homogenen DGL $\Delta(y) = 0$, so ist nach den Regeln der Linearen Algebra $y_p + L_0 = \{y_p + y_h : y_h \in L_0\}$ die Lösungsmenge der inhomogenen Gleichung.

In der Praxis haben einfache (d.h. eindimensionale) gewöhnliche lineare Differentialgleichungen n -ter Ordnung die Gestalt

$$y(t)a_0(t) + y'(t)a_1(t) + \dots + y^{(n)}(t)a_n(t) = f(t),$$

eventuell mit n Anfangsbedingungen $y(0) = y_0, y'(0) = y_1, \dots, y^{(n-1)}(0) = y_{n-1}$. Die Koeffizienten $a_i(t)$ dürfen durchaus (auch nichtlineare) Funktionen in t sein, der Differentialoperator

$$\Delta : V_1 \rightarrow V_2, \quad y \mapsto a_0y + a_1y' + \dots + a_ny^{(n)}$$

als Abbildung zum Beispiel vom Raum V_1 der (auf einer geeigneten Teilmenge D_y von \mathbb{R}) n -mal stetig differenzierbaren Funktionen y in den Raum V_2 der stetigen Funktionen ist trotzdem linear. Die Funktion f stellt den inhomogenen Anteil der DGL dar. Für höherdimensionale Lösungen gilt dasselbe, und man kann

$$\mathbf{y}'(t)a_0(t) + \mathbf{y}''(t)a_1(t) + \dots + \mathbf{y}^{(n)}(t)a_n(t) = \mathbf{f}(t)$$

schreiben. Nach 3.1.4 können solche Gleichungen der Ordnung n durch Erhöhung der Dimension auf Ordnung 1 reduziert werden, und man erhält Gleichungen der Gestalt

$$\mathbf{y}'(t) = A(t)\mathbf{y}(t) + \mathbf{f}(t),$$

wobei $A(t)$ nun eine quadratische Matrix ist, deren Komponenten Funktionen in t sind. Die Theorie dazu wird Inhalt von Abschnitt 3.2 sein. Im hier nicht näher behandelten Fall partieller DGLen treten an die Stelle der gewöhnlichen Ableitungen $y^{(i)}(t)$ natürlich partielle Ableitungen nach verschiedenen Variablen, eventuell auch höherer Ordnung.

Es sollte nun nicht mehr verwundern, dass die Lösungstheorie linearer DGLen extensiv von Linearer Algebra Gebrauch machen kann, was vieles wesentlich einfacher macht. So wie die Differentialrechnung in mehreren Variablen im Wesentlichen darin besteht, möglichst viele auch nichtlineare Funktionen (geeignet sind dafür die differenzierbaren) durch lineare zu approximieren, um dadurch die Möglichkeiten der Linearen Algebra zu nutzen, kann man versuchen, möglichst viele auch nichtlineare DGLen durch lineare zu approximieren. Diesen Ansatz werden wir kaum weiter verfolgen. Die Theorie der linearen Differentialgleichungen selbst werden wir in den folgenden Abschnitten jedoch einigermaßen vertiefen.

3 Differentialgleichungen (DGLen)

An dieser Stelle sei noch ein wichtiges Ergebnis über die Lösungsmenge linearer DGLen festgehalten. Aus der obigen Darstellung $\mathbf{y}'(t) = A(t)\mathbf{y}(t) + \mathbf{f}(t)$ ist ersichtlich, dass wir es mit einer expliziten DGL zu tun haben. Sind die Komponentenfunktionen und somit $A(t)$ stetig (was bei linearen DGLen ab nun stillschweigend vorausgesetzt sei), so erfüllt die DGL die Voraussetzungen des Existenz- und Eindeutigkeitsatzes 3.1.5.2 (insbesondere die Lipschitz-Bedingung, Übung). Geben wir eine Anfangsbedingung $\mathbf{y}(t_0) = \mathbf{y}_0 = (y_1, \dots, y_n)$ vor, so gibt es folglich genau eine Lösung \mathbf{y} .

Übungsaufgabe 193. (E) Überprüfen Sie im Detail diese Argumentation für die Eindeutigkeit, indem Sie Folgendes zeigen:

1. Zeigen Sie, dass jede lineare Abbildung $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ Lipschitz-stetig ist. Anleitung: Alle Vektoren $\mathbf{x} = (x_1, \dots, x_m) \in \mathbb{R}^m$ mit Länge $\|\mathbf{x}\| \leq 1$ erfüllen $|x_i| \leq 1$ für $i = 1, \dots, m$. Für so ein \mathbf{x} folgt

$$\|f(\mathbf{x})\| = \left\| \sum_{i=1}^m x_i F(\mathbf{e}_i) \right\| \leq \sum_{i=1}^m |x_i| \|f(\mathbf{e}_i)\| \leq mr,$$

sofern r den größten der Werte $\|f(\mathbf{e}_i)\|$ bezeichnet. Zeigen Sie, dass $\lambda_F := mr$ eine Lipschitz-Konstante für F ist.

2. In unserer Anwendung hängt die lineare Transformation $A(t)$ von t ab. Im ersten Teil wurde eine Lipschitz-Konstante λ_F aber nur für ein festes lineares F gefunden. Verwenden Sie die dortigen Abschätzungen und den Satz vom Maximum, um zu folgern, dass es eine Umgebung von (t_0, \mathbf{y}_0) gibt, auf der eine gemeinsame Lipschitz-Konstante genommen werden kann.
3. Rekapitulieren Sie die entsprechenden Argumente aus 3.1.5 und begründen Sie damit, dass die Lösung einer linearen DGL mit Anfangsbedingung sogar auf dem maximalen zusammenhängenden Definitionsbereich eindeutig ist.

Die Abbildung $\mathbf{y}_0 = (y_1, \dots, y_n) \mapsto \mathbf{y}$, die jeder Anfangsbedingung die zugehörige Lösung \mathbf{y} der DGL zuordnet, ist linear, injektiv (weil verschiedene Anfangsbedingungen verschiedene Lösungen erzwingen) auf dem n -dimensionalen Raum \mathbb{R}^n definiert. Also ist die Lösungsmenge der DGL ein zu \mathbb{R}^n isomorpher Raum von Funktionen, folglich selbst von der Dimension n . Wir haben damit bewiesen:

Satz 3.2.1.1. Die Lösungen einer homogenen linearen DGL n -ter Ordnung mit stetigen Koeffizientenfunktionen bilden einen n -dimensionalen Vektorraum.

Bei der Lösung solcher DGLen wird es also immer darum gehen, n linear unabhängige Lösungen zu finden. Diese bilden dann automatisch eine Basis des Vektorraums aller Lösungen.

3.2.2 Speziellere und allgemeinere lineare DGLen – Überblick

Inhalt in Kurzfassung: In Bezug auf lineare DGLen behandeln wir die Begriffspaare einfach versus gekoppelt, erste Ordnung versus höhere Ordnung, konstante versus allgemeine Koeffizienten und homogen versus inhomogen.

Folgende Unterscheidungen für gewöhnliche lineare DGLen werden uns leiten:

- einfach (eindimensional) – gekoppelt (höherdimensional)
- erste Ordnung – höhere Ordnung
- homogen – inhomogen
- konstante Koeffizienten – allgemeine Koeffizientenfunktionen

Dabei lässt sich jeweils der erste Typ als einfacher Spezialfall des zweiten Typs auffassen. Bevor wir mit der eigentlichen Theorie beginnen, sind ein paar Bemerkungen zu diesen vier Unterscheidungen am Platz:

Einfach – gekoppelt: Zunächst wir einfache (eindimensionale) lineare DGLen. Sind mehrere unbekannte Funktionen gesucht, die durch lineare DGLen aneinander gebunden sind, so spricht man von gekoppelten DGLen. Die mathematische Modellierung führt ins Mehrdimensionale und zu Verallgemeinerungen, für die zwar durchwegs die Lineare Algebra geeignete Methoden bereit stellt, deren Anwendung aber doch einen wesentlich größeren Aufwand mit sich bringt. Trotzdem werden sich Analogien erkennen lassen, die das Verständnis wesentlich erleichtern. Für ausführlichere Erklärungen muss an dieser Stelle noch auf 3.2.6 verwiesen werden.

Erste Ordnung – höhere Ordnung: Wie wir aus 3.1.4 wissen, lassen sich DGLen höherer Ordnung auf solche erster Ordnung zurückführen, sofern man ein Anwachsen der Dimension in Kauf nimmt. Somit könnten wir die Behandlung dieses Aspektes ebenfalls mit einem Verweis auf ?? abtun. Für den Spezialfall konstanter Koeffizienten lohnt es jedoch auch einen anderen Gesichtspunkt einzunehmen.

Homogen – inhomogen: Als Folgerung der allgemeinen Theorie linearer Gleichungen aus der Linearen Algebra kennen wir bereits die Struktur der Lösungsmenge einer inhomogenen linearen DGL (siehe 3.2.1): Ist y_p irgendeine (sogenannte partikuläre) Lösung der zugehörigen homogenen Gleichung und L_0 die Lösungsmenge der zugehörigen homogenen Gleichung, so enthält

$$L = y_p + L_0 = \{y_p + y_h : y_h \in L_0\}$$

genau die Lösungen der inhomogenen Gleichung. Beherrscht man den homogenen Fall, so kommt es also nur noch darauf an, wenigstens eine partikuläre Lösung y_p der inhomogenen Gleichung zu finden.

Konstante Koeffizienten – allgemeine Koeffizienten: Für DGLen mit konstanten Koeffizienten liegt eine ziemlich befriedigende Lösungstheorie vor. Die allgemeine, höherdimensionale Variante wird ein wesentlicher Inhalt von 3.2 sein. Im eindimensionalen Fall sind aber maßgebliche Vereinfachungen möglich, die schon in diesem Abschnitt,

3 Differentialgleichungen (DGLen)

und zwar in 3.2.4 besprochen werden sollen.

Der Rest des Abschnitts beschäftigt sich mit jenen der angeschnittenen Fragen, die für einfache lineare DGLen von besonderem Interesse sind. Wir beginnen mit dem einfachsten Fall (einfach, konstante Koeffizienten, erste Ordnung, homogen) und verallgemeinern dann schrittweise in Richtung höhere Ordnung, inhomogen und allgemeine Koeffizienten. Dem letzten Schritt, der Verallgemeinerung von einfach zu gekoppelt, wird wie gesagt Abschnitt ?? gewidmet sein.

3.2.3 Einfache homogene lineare DGLen mit konstanten Koeffizienten

Inhalt in Kurzfassung: Homogene lineare DGLen mit konstanten Koeffizienten aber beliebiger Ordnung lassen sich mit einer Standardmethode behandeln. Gelingt eine Faktorisierung des sogenannten charakteristischen Polynoms der Gleichung, so lässt sich unmittelbar die Lösungsmenge angeben. Dies geschieht mit Hilfe von Satz 3.2.3.2, der eine Basis von sogenannten Fundamentallösungen der DGL angibt.

Die Exponentialfunktion \exp zur Basis e , Ausgangspunkt für sehr vieles in der Mathematik, trat auf als Lösung einer einfachen, homogenen linearen DGL erster Ordnung mit konstanten Koeffizienten: $y' = y$ oder $y' - y = 0$. Eine beliebige DGL dieses Typs ist von der Gestalt $a_1 y' + a_0 y = 0$ mit $a_0, a_1 \in \mathbb{R}$, wobei wir $a_0, a_1 \neq 0$ annehmen dürfen. So eine Gleichung lässt sich umformen zu $y' = \lambda y$ (mit $\lambda = -\frac{a_0}{a_1}$). Offenbar ist $y = \exp_{e^\lambda} : t \mapsto e^{\lambda t}$ eine Lösung, ebenso wie alle Vielfachen $c \cdot \exp_{e^\lambda}$ mit $c \in \mathbb{R}$. Diese DGL erfüllt alle Voraussetzungen des Existenz- und Eindeigkeitssatzes 3.1.5.2. Deshalb gibt es zu jedem Anfangswert $y_0 \in \mathbb{R}$ und zu jedem $t_0 \in \mathbb{R}$ genau eine Lösung y mit $y(t_0) = y_0$, welche notgedrungen die Funktion $c_0 \cdot \exp_{e^\lambda}$ sein muss mit jenem c_0 , für welches $y_0 = c_0 e^{t_0 \lambda}$ gilt, also $c_0 = y_0 e^{-t_0 \lambda}$. Den Fall homogener linearer DGLen erster Ordnung mit konstanten Koeffizienten beherrscht man also sehr leicht.

Wenden wir uns dem Fall höherer Ordnung zu, wenn die DGL (eventuell nach Division durch den Koeffizienten bei der höchsten vorkommenden Ableitung $y^{(n)}$) also auf die Gestalt

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = 0$$

(d.h. mit führendem Koeffizienten $a_n = 1$) gebracht werden kann.

Ansatz: Versuchen wir es auch hier mit einer Lösung der Form $y(t) = e^{\lambda t}$. Wegen $y^{(k)} = \lambda^k e^{\lambda t} = \lambda^k y$ (Beweis durch Induktion nach k) erhält die DGL die Gestalt

$$\lambda^n e^{\lambda t} + a_{n-1} \lambda^{n-1} e^{\lambda t} + \dots + a_1 \lambda e^{\lambda t} + a_0 e^{\lambda t} = p(\lambda) e^{\lambda t} = 0$$

mit dem sogenannten **charakteristischen Polynom**

$$p(x) := \sum_{i=0}^n a_i x^i$$

beziehungsweise der **charakteristischen Gleichung**

$$p(x) = \sum_{i=0}^n a_i x^i = 0.$$

Ist λ eine Nullstelle von p , so ist daher $y(t) = e^{\lambda t}$ eine Lösung unserer DGL. Im „Idealfall“ hat p als Polynom n -ten Grades n verschiedene reelle Lösungen $\lambda_1, \dots, \lambda_n$. Weil die Lösungen einer homogenen linearen DGL laut Satz 3.2.1.1 einen n -dimensionalen Unterraum bilden, haben wir mit den Funktionen $y(t) = e^{\lambda_i t}$, $i = 1, \dots, n$, auch schon eine Basis dieses Lösungsraums gefunden, sofern wir ihre lineare Unabhängigkeit nachweisen können. Die Überprüfung kann mit der Wronski-Determinante erfolgen, auf die wir später (siehe 3.2.8) noch werden zu sprechen kommen. Aber auch eine elementare Methode bietet sich an:

Wir gehen mittels Induktion nach n vor. Für $n = 0$ haben wir es mit der leeren Menge von Lösungen zu tun, die linear unabhängig ist. Angenommen, je n solcher Funktionen seien linear unabhängig (Induktionsannahme), es bestünde aber (indirekter Beweis) eine lineare Abhängigkeit der Form

$$\sum_{i=1}^{n+1} r_i e^{\lambda_i t} = 0$$

mit geeigneten Koeffizienten $r_i \in \mathbb{R}$, die nicht alle $= 0$ sind, und für alle $t \in \mathbb{R}$. Dann müssen sogar alle r_i von 0 verschieden sein, weil sonst ja der entsprechende Term aus der Summe gestrichen werden könnte, was einen Widerspruch zur Induktionsannahme ergäbe. Wir dürfen annehmen, dass die λ_i der Größe nach geordnet sind:

$$\lambda_1 < \lambda_2 < \dots < \lambda_n < \lambda_{n+1}$$

Wir dividieren die Gleichung durch $e^{\lambda_{n+1} t}$ und erhalten

$$\sum_{i=1}^{n+1} r_i e^{(\lambda_i - \lambda_{n+1})t} = 0.$$

Wegen $\lambda_i - \lambda_n < 0$ für $i = 1, \dots, n$ konvergieren alle Summanden außer dem letzten für $t \rightarrow \infty$ gegen 0, woraus doch $r_{n+1} = 0$ folgt, Widerspruch. Damit haben wir bewiesen:

Proposition 3.2.3.1. *Hat das charakteristische Polynom*

$$p(x) := \sum_{i=0}^n a_n x^n.$$

der homogenen linearen DGL n -ter Ordnung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0 = 0$$

n verschiedene reelle Nullstellen $\lambda_1, \dots, \lambda_n$, so ist die Menge der Lösungen gegeben durch den n -dimensionalen Raum, der von den n linear unabhängigen Funktionen $y_i(t) = e^{\lambda_i t}$, $i = 1, \dots, n$, aufgespannt wird.

Zerfällt das charakteristische Polynom nicht in ausschließlich reelle, sondern teilweise auch komplexe Linearfaktoren, die aber weiterhin paarweise verschieden sind, so greifen

3 Differentialgleichungen (DGLen)

formal sehr ähnliche Argumente. Allerdings werden für komplexe Nullstellen $\alpha + i\beta$ auch die zugehörigen Funktionen

$$f(t) = e^{(\alpha+i\beta)t} = e^{\alpha t}(\cos(\beta t) + i \sin(\beta t))$$

komplexwertig und nicht unmittelbar brauchbar. Echt komplexe Nullstellen treten in Konjugiertenpaaren auf. Zu obiger Lösung gesellt sich daher auch die der konjugierten Nullstelle $\alpha - i\beta$ zugehörige Lösung

$$g(t) = e^{(\alpha-i\beta)t} = e^{\alpha t}(\cos(\beta t) - i \sin(\beta t)).$$

Da wir im Lösungsraum Linearkombinationen bilden dürfen, sind auch

$$\frac{f(t) + g(t)}{2} = e^{\alpha t} \cos(\beta t)$$

und

$$\frac{f(t) - g(t)}{2i} = e^{\alpha t} \sin(\beta t)$$

Lösungen. Diese beiden Lösungen sind reellwertig und wegen $\beta \neq 0$ linear unabhängig (Übung).

Übungsaufgabe 194. (E) Überprüfen Sie die lineare Unabhängigkeit dieser beiden Funktionen $e^{\alpha t} \cos(\beta t)$ und $e^{\alpha t} \sin(\beta t)$.

Wir wollen unsere Überlegungen am einfachen und besonders interessanten Beispiel der DGL $y'' = -y$, also $y'' + y = 0$ überprüfen. Das charakteristische Polynom dieser Gleichung ist $p(x) = x^2 + 1$ mit den beiden komplexen Lösungen i und $-i$. In den allgemeinen Überlegungen ist also $\alpha = 0$ und $\beta = 1$ zu setzen. Lösungen dieser DGL sollten also $e^{\alpha t} \cos(\beta t) = \cos t$ und $e^{\alpha t} \sin(\beta t) = \sin t$ sein, was wegen $\cos'' = -\cos$ und $\sin'' = -\sin$ tatsächlich der Fall ist. Wieder einmal hat sich der fundamentale Zusammenhang zwischen Exponential- und Winkelfunktionen bewährt. Er hat eine Verbindung hergestellt zwischen – grob gesprochen – DGLen mit exponentiell wachsenden oder abfallenden Lösungen bei reellen Nullstellen des charakteristischen Polynoms und periodisch oszillierenden bei komplexen Nullstellen.

Offen bleibt die delikate Frage, was zu tun ist, wenn das charakteristische Polynom mehrfache Nullstellen hat. Dann haben wir mit den gefundenen Basislösungen $e^{\lambda t}$ für Nullstellen λ des charakteristischen Polynoms einer gegebenen linearen DGL n -ten Grades nämlich nicht genug, um einen n -dimensionalen Lösungsraum aufzuspannen. Es zeigt sich, dass eine Basis des Lösungsraumes gegeben ist durch alle Funktionen der Bauart $t^k e^{\lambda t}$, wobei λ wie bisher alle Nullstellen des charakteristischen Polynoms p der DGL durchläuft und $k = 0, 1, \dots, v(\lambda) - 1$, wenn $v(\lambda)$ die Vielfachheit von λ als Nullstelle von p ist. Für den allgemeinen Beweis dieser Tatsache verweisen wir auf die in 3.2.10 behandelte Methode. Für den Fall $n = 2$ lässt sich unsere Behauptung aber explizit überprüfen:

Ist λ eine doppelte Nullstelle des charakteristischen Polynoms $p(x)$ der DGL $y'' + a_1y' + a_0 = 0$, so bedeutet das

$$x^2 + a_1x + a_0 = (x - \lambda)^2 = x^2 - 2\lambda x + \lambda^2,$$

also $a_1 = -2\lambda$ und $a_0 = \lambda^2$. Nach dem oben Behaupteten sollte die Funktion $y(t) = te^{\lambda t}$ eine Lösung der DGL sein. Tatsächlich, mit der Produktregel berechnen wir

$$\begin{aligned} y'(t) &= t\lambda e^{\lambda t} + \lambda e^{\lambda t} = (1 + \lambda t)e^{\lambda t}, \\ y''(t) &= (1 + \lambda t)\lambda e^{\lambda t} + \lambda e^{\lambda t} = (\lambda^2 t + 2\lambda)e^{\lambda t}, \end{aligned}$$

woraus

$$\begin{aligned} y''(t) + a_1y'(t) + a_0y(t) &= \\ &= (\lambda^2 t + 2\lambda)e^{\lambda t} - 2\lambda(1 + \lambda t)e^{\lambda t} + \lambda^2 e^{\lambda t} = (\lambda^2 t + 2\lambda - 2\lambda - 2\lambda^2 t + \lambda^2)e^{\lambda t} = 0 \end{aligned}$$

folgt. Hat λ als Nullstelle von p noch größere Vielfachheit, so wird (wir verzichten auf den Beweis) auch $t \mapsto t^2 e^{\lambda t}$ zur Lösung der DGL etc. Analoges gilt auch für komplexe Nullstellen. Somit lautet der allgemeine Satz, der doppelte wie auch komplexe Nullstellen berücksichtigt:

Satz 3.2.3.2. *Das charakteristische Polynom p der homogenen linearen DGL n -ter Ordnung*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$$

habe folgende Zerlegung in Linearfaktoren:

$$p(x) = \sum_{i=0}^n a_n x^n = \prod_{j=1}^k (x - \lambda_j)^{e_j} \cdot \prod_{l=1}^m ((x - (\alpha_l + i\beta_l))(x - (\alpha_l - i\beta_l)))^{f_l}$$

mit $k, m \in \mathbb{N}$, $\sum_{j=1}^k e_j + \sum_{l=1}^m 2f_l = n$ und $a_n = 1$. Dabei seien die λ_j , $j = 1, \dots, k$, sämtliche paarweise verschiedenen reellen Nullstellen von p mit den Vielfachheiten $e_j > 0$, die $\alpha_l \pm i\beta_l$ mit $\beta_l \neq 0$ sämtliche paarweise verschiedenen komplexen Nullstellen von p mit den Vielfachheiten $f_l > 0$.

Dann bildet die Lösungsmenge L der gegebenen DGL einen n -dimensionalen Vektorraum von Funktionen. Eine Basis von L ist gegeben durch die sogenannten **Fundamentallösungen** der Gleichung. Das sind die Funktionen

$$t \mapsto t^\nu e^{\lambda_j t}, \quad j = 1, \dots, k, \quad \nu = 0, \dots, e_j - 1,$$

und die Funktionen

$$t \mapsto t^\nu e^{\alpha_l t} \cos(\beta_l t) \quad \text{und} \quad t \mapsto t^\nu e^{\alpha_l t} \sin(\beta_l t), \quad l = 1, \dots, m, \quad \nu = 0, \dots, f_l - 1.$$

Übungsaufgabe 195. (P) Zeigen Sie, dass $\cos \omega t$ und $\sin \omega t$ zwei linear unabhängige Lösungen von $y'' + \omega^2 y = 0$ sind und damit ein Fundamentalsystem darstellen. ($\omega \neq 0!$)

3 Differentialgleichungen (DGLen)

Übungsaufgabe 196. (P) Wie Aufgabe 195 mit den den Lösungen $e^{\omega t}$, $e^{-\omega t}$ der DGL $y'' - \omega^2 y = 0$.

Übungsaufgabe 197. (T) Lösen Sie

1. $y'' + 5y' + 6y = 0$

2. $y'' - 2y' + y = 0$

3. $y'' + 6y' + 10y = 0$

Übungsaufgabe 198. (T) Lösen Sie

1. $y''' - 3y'' + y' - 3y = 0$

2. $y''' - 3y'' + 3y' - y = 0$

3. $y''' + 3y'' - y' - 3y = 0$

4. $y''' - y = 0$.

Übungsaufgabe 199. (T) Lösen Sie

1. $y^{(4)} - 2y'' + y = 0$

2. $y^{(4)} + 2y'' + y = 0$

3. $y^{(4)} - y = 0$.

3.2.4 Inhomogene Gleichungen mit konstanten Koeffizienten

Inhalt in Kurzfassung: Kennt man eine (sogenannte partikuläre) Lösung y_p einer inhomogenen linearen DGL und die Lösungsmenge (allgemeine Lösung) L_0 der zugehörigen homogenen DGL, so hat die allgemeine Lösung L des inhomogenen Problems die Gestalt $L = y_p + L_0$. Hat man das homogene Problem z.B. mit den Mitteln aus 3.2.3 gelöst, so kommt es also nur noch darauf an, ein y_p zu finden. Ist der inhomogene Anteil (die Störfunktion) der inhomogenen DGL von spezieller Form, so gibt es auch ein y_p ganz ähnlicher Bauart. Beispiele dieser Art werden nun besprochen.

In diesem Unterabschnitt heißt die unabhängige Variable x (nicht wie zuletzt t).

Beginnen wir mit dem einfachsten Fall, nämlich der Ordnung 1: $y' + ay = b$ mit $a, b \in \mathbb{R}$. Wir greifen auf die Methode der Trennung der Variablen aus 3.3.2, schreiben $y' = \frac{dy}{dx}$, also $\frac{dy}{dx} + ay = b$ und formen um zu $\frac{dy}{b-ay} = dx$. Integration auf beiden Seiten liefert $-\frac{\ln(b-ay)}{a} = x$, also $b - ay = e^{-ax}$ und schließlich $y = y(x) = \frac{1}{a}(b - e^{-ax})$.

Diese Methode greift nicht mehr, wenn die Ordnung der DGL erhöht wird. Die Gleichungen, mit denen wir uns nun beschäftigen, haben folgende Bauart:

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \cdots + a_1y'(x) + a_0y(x) = f(x).$$

Man beachte, dass die Störfunktion f durchaus von der unabhängigen Variablen x abhängen darf, obwohl die Koeffizienten a_i konstant sind. Wir beschränken uns dabei auf Störfunktionen der Form

$$f(x) = (r_0 + r_1x + \cdots + r_mx^m) e^{\alpha x} \cos(\beta x) + (s_0 + s_1x + \cdots + s_mx^m) e^{\alpha x} \sin(\beta x), \quad (\text{i})$$

(Dabei soll wieder $r_m \neq 0$ oder $s_m \neq 0$ sein.) Darin sind auch die Spezialfälle

$$f(x) = (r_0 + r_1x + \cdots + r_mx^m) e^{\alpha x}, \quad \text{Spezialfall } \beta = 0 \quad (\text{ii})$$

und

$$f(x) = r_0 + r_1x + \cdots + r_mx^m. \quad \text{Spezialfall } \alpha = \beta = 0 \quad (\text{iii})$$

enthalten.

Die folgenden Aussagen über den Partikuläransatz sind im wesentlichen gleichlautend mit denen bei DGLs-Systemen (gekoppelten DGLen, siehe ab 3.2.6), bis auf einen Unterschied bei der Graderhöhung im Resonanzfall, den wir entsprechend hervorheben werden. Es genügt wieder, die Vorgangsweise im Fall (i) zu erläutern:

Ist $\alpha + i\beta$ keine Lösung der charakteristischen Gleichung, so setzt man an

$$y_p(x) = (b_0 + b_1x + \cdots + b_mx^m) e^{\alpha x} \cos(\beta x) + (c_0 + c_1x + \cdots + c_mx^m) e^{\alpha x} \sin(\beta x).$$

Fällt dagegen $\alpha + i\beta$ mit einer v -fachen Nullstelle λ zusammen („Resonanz“), so ist anzusetzen

$$y_p(x) = (b_0 + b_1x + \cdots + b_mx^m) x^v e^{\alpha x} \cos(\beta x) + (c_0 + c_1x + \cdots + c_mx^m) x^v e^{\alpha x} \sin(\beta x).$$

Der Polynomgrad wird wie bei DGL-Systemen um v erhöht, jedoch mit einem Unterschied: Hier kommen im Ansatz die niedrigen Potenzen x^0, \dots, x^{v-1} nicht vor. Sie vorzusehen, wäre nicht falsch, aber überflüssig!

Beispiel 3.2.4.1. Balken unter Last auf elastischer Unterlage.

Die Durchbiegung $y(x)$ eines schlanken Balkens mit Biegesteifigkeit 1, der auf einer elastischen Unterlage ruht und unter Last $f(x)$ steht, wird beschrieben durch die inhomogene DGL

$$y^{(4)}(x) + \omega^4 y(x) = f(x) \quad (\omega > 0).$$

Zuerst ist die zugehörige homogene Gleichung zu lösen. Deren charakteristische Gleichung $\lambda^4 + \omega^4 = 0$ hat die Lösungen

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \left\{ \frac{\omega}{\sqrt{2}} + i \frac{\omega}{\sqrt{2}}, \frac{\omega}{\sqrt{2}} - i \frac{\omega}{\sqrt{2}}, -\frac{\omega}{\sqrt{2}} + i \frac{\omega}{\sqrt{2}}, -\frac{\omega}{\sqrt{2}} - i \frac{\omega}{\sqrt{2}} \right\}.$$

Setzt man abkürzend $\mu = \frac{\omega}{\sqrt{2}}$, so erhält man für die reelle Homogenlösung

$$y_h(x) = c_1 e^{\mu x} \cos(\mu x) + c_2 e^{\mu x} \sin(\mu x) + c_3 e^{-\mu x} \cos(\mu x) + c_4 e^{-\mu x} \sin(\mu x).$$

3 Differentialgleichungen (DGLen)

Wir wenden uns jetzt der inhomogenen Gleichung zu. Als Lastfunktion sei zuerst $f(x) = \sin(\mu x)$ gewählt. Da $0 + i\mu$ mit keiner der charakteristischen Zahlen λ_1 bis λ_4 zusammenfällt, haben wir anzusetzen

$$y_p(x) = a \cos(\mu x) + b \sin(\mu x).$$

Ableiten und Einsetzen gibt

$$y_p^{(4)}(x) + \omega^4 y_p(x) = a\mu^4 \cos(\mu x) + b\mu^4 \sin(\mu x) + a\omega^4 \cos(\mu x) + b\omega^4 \sin(\mu x) = f(x) = \sin(\mu x).$$

Zusammenfassen der Sinus- und Cosinusglieder und Koeffizientenvergleich führt auf zwei Gleichungen für die zwei Unbekannten a und b :

$$\begin{aligned} \cos \mu x : \dots & \quad a(\mu^4 + \omega^4) = 0 \quad \Rightarrow \quad a = 0 \quad (\text{beachte : } \mu^4 + \omega^4 = 5\omega^4/4 \neq 0), \\ \sin \mu x : \dots & \quad b(\mu^4 + \omega^4) = 1 \quad \Rightarrow \quad b = 1/(\mu^4 + \omega^4) = 4/(5\omega^4). \end{aligned}$$

Beispiel 3.2.4.2.

Die Problemstellung ist dieselbe wie zuvor, wir illustrieren aber hier den Resonanzfall am Beispiel der Lastfunktion

$$f(x) = e^{\mu x} \cos(\mu x).$$

Da $\mu + i\mu$ mit der einfachen Wurzel $\lambda = \lambda_1$ zusammenfällt, muss der Partikuläranatz jetzt lauten

$$y_p(x) = a x e^{\mu x} \cos(\mu x) + b x e^{\mu x} \sin(\mu x).$$

Die Berechnung der reellen Konstanten a, b mit diesem Ansatz ist mühsam und wir führen sie nicht weiter aus.

Übungsaufgabe 200. (T) Lösen Sie allgemein

1. $\dot{x} + 5x = \cos(5t)$
2. $\dot{x} + 5x = e^{-5t}$.
3. $\dot{x} + 3tx = t$
4. $\dot{x} \cdot \frac{1}{\sin \varphi} + x = 1$
5. $\dot{x} + x \cot t = 5e^{\cos t}$

für die gesuchte Funktion $x = x(t)$.

Übungsaufgabe 201. (T) Bei den folgenden DG ist eine Partikularlösung $y(x)$ mittels Exponentialansatz zu bestimmen

1. $y'' - 2y' + 5y = x^2$
2. $y'' - 2y' + 5y = e^x$
3. $y'' - 2y' + 5y = 3x^2 - 2e^x$

4. $y'' + 2y' + 5y = \sin 2x$

5. $y'' + 2y' + 5y = e^x \sin 2x$

6. $y'' + 2y' + 5y = x^2 e^x \sin 2x$

Übungsaufgabe 202. (T^*) Wie lautet die Lösung von $my'' + m\omega^2 y = ma \sin \alpha t$ (ungedämpfte Schwingung mit harmonischer Erregung: $m > 0, \omega, a, \alpha \neq 0$). Beachten Sie die Fallunterscheidung $\omega \neq \alpha$ und $\omega = \alpha$.

Übungsaufgabe 203. (T^*) Wie lautet die Lösung von $my'' + 2mcy' + m\omega^2 y = ma \sin \alpha t$ ($= f(t)$) (gedämpfte Schwingung mit harmonischer Erregung: $c > 0, m > 0, \omega, a, \alpha \neq 0$). Für die homogene Lösung sind hier dieselben Fallunterscheidungen wie bei der gedämpften Schwingung ohne Erregung zu beachten, dagegen ist (im Gegensatz zum ungedämpften Fall) für den Partikuläransatz eine Fallunterscheidung entbehrlich. Diskutieren Sie folgende Fragen:

1. Für welchen Wert von α (Kreisfrequenz der Erregung) wird die Amplitude r maximal?
2. Wie groß ist die Phasenverschiebung $\Delta\varphi$ zwischen Lösung $y(t)$ und Erregung $f(t)$?

Übungsaufgabe 204. (T) Es seien $\{\lambda_1, \lambda_2, \dots, \lambda_6\} = \{2+i, 2+i, 2+i, 2-i, 2-i, 2-i\}$ die charakteristischen Werte einer linearen DGL 6. Ordnung $y^{(6)} + a_5 y^{(5)} + \dots + a_1 y' + a_0 y = 0$ mit konstanten Koeffizienten a_5, a_4, \dots, a_0 . Wie lauten die Koeffizienten? Wie lautet die Lösung der DGL mit Inhomogenität $f(x) = x^2 e^{2x}$?

Übungsaufgabe 205. (T) Wie Aufgabe 204 mit $\{\lambda_1, \lambda_2, \dots, \lambda_6\} = \{1, -1, i, i, -i, -i\}$ und $f(x) = (x^3 + x + 1)e^x \sin x$.

3.2.5 Einfache lineare DGLen mit nichtkonstanten Koeffizienten

Inhalt in Kurzfassung: Sind die Koeffizienten einer linearen DGL nicht mehr konstant, so versagen die bisher besprochenen Methoden. Immerhin gibt es für den homogenen Fall erster Ordnung ein einfaches Lösungsschema. Für die Ausdehnung auf den inhomogenen Fall, führt ein geeigneter sogenannter integrierender Faktor zum Ziel oder, alternativ, die Methode der Variation der Konstanten.

Bisher haben wir lineare DGLen fast ausschließlich mit konstanten Koeffizienten betrachtet. Will man sich von dieser Einschränkung lösen, hat man es naturgemäß mit weitaus komplizierteren Aufgaben zu tun. Unter diesen wiederum der einfachste Fall ist die homogene DGL erster Ordnung

$$y'(t) = a(t)y(t).$$

Sie lässt sich noch sehr leicht lösen: Die Umformungen der DGL zu

$$a(t) = \frac{y'(t)}{y(t)} = (\ln |y(t)|)'$$

3 Differentialgleichungen (DGLen)

und

$$\ln(|y(t)|) = \int a(t) dt + c$$

zeigen, dass

$$y(t) = c \cdot \exp\left(\int a(t) dt\right)$$

mit $c \in \mathbb{R}$ die allgemeine Lösung dieser DGL ist.

Eine erste Möglichkeit, auch der entsprechenden inhomogenen DGL

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

mit (stetigen) Funktionen $a(t)$ und $b(t)$ beizukommen, fußt auf dieser Methode. Wie wir in Abschnitt 3.3.3 über exakte DGLen sehen werden, können manche DGLen leichter gelöst werden, wenn man sie mit einem **integrierenden Faktor** $I(t)$ multipliziert. Es zeigt sich, dass im vorliegenden Fall ein $I(t)$ mit $I'(t) = a(t)I(t)$ zum Ziel führt, also eine Lösung einer DGL, wie wir sie im vorangegangenen Absatz behandelt haben. Von dort wissen wir, dass

$$I(t) = \exp\left(\int a(t) dt\right) = e^{A(t)}$$

mit einer Stammfunktion A von a die gewünschte Eigenschaft hat. Damit nimmt die ursprüngliche DGL die Form

$$y'(t)e^{A(t)} + a(t)y(t)e^{A(t)} = b(t)e^{A(t)}.$$

Nach der Produktregel ist die linke Seite dieser Gleichung die Ableitung von $y(t)e^{A(t)}$. Integration nach t liefert daher

$$y(t)e^{A(t)} = \int b(t)e^{A(t)} dt$$

und somit

$$y(t) = e^{-A(t)} \int b(t)e^{A(t)} dt.$$

Die Lösung der ursprünglichen DGL kann also zurückgeführt werden auf die Bestimmung von zwei Stammfunktionen, einer von $a(t)$ und daran anschließend einer von $b(t)e^{A(t)}$.

Als Methode für den Übergang von homogenen zu inhomogenen linearen DGLen erwähnen wir auch noch die sogenannte Methode der **Variation der Konstanten**, die wir nochmals in 3.2.8 aufgreifen werden. Um den entscheidenden Punkt deutlich zu erkennen, behandeln wir den einfachsten interessanten Fall wie oben. Wir suchen also nochmals nach einer Lösung von

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

und nehmen an, dass eine Lösung $y_h \neq 0$ des homogenen Problems $y'(t) + a(t)y(t) = 0$ bereits vorliegt. Dann ist cy_h , $c \in \mathbb{R}$, die allgemeine Lösung des homogenen Problems. Der

zielführende Trick besteht darin, die gesuchte partikuläre Lösung y_p des inhomogenen anzusetzen als $y_p(t) = c(t)y_h(t)$, wobei $c(t)$ jetzt eine von t abhängige, nun zu bestimmende Funktion und keine Konstante ist. Die linke Seite der inhomogenen Gleichung nimmt die Form

$$\begin{aligned} y_p' + a(t)y_p(t) &= c(t)y_h'(t) + c'(t)y_h(t) + a(t)c(t)y_h(t) = \\ &= c(t)(y_h'(t) + a(t)y_h(t)) + c'(t)y_h(t) = c'(t)y_h(t) \end{aligned}$$

an. Dabei wurde benutzt, dass y_h die homogene DGL löst. Es folgt $c'(t)y_h(t) = b(t)$, also $c(t) = \int \frac{b(t)}{y_h(t)} dt$. Somit ergibt sich eine partikuläre Lösung

$$y_p(t) = y_h(t) \int \frac{b(t)}{y_h(t)} dt.$$

Die Lösungsmenge des inhomogenen Problems besteht nach der allgemeinen Theorie linearer DGLen somit aus allen Funktionen y der Gestalt $y = y_p + c \cdot y_h$ mit $c \in \mathbb{R}$ und einer Lösung $y_h \neq 0$ des homogenen Problems. Ist eine Anfangsbedingung $y(t_0) = y_0$ vorgegeben, ergibt sich das entsprechende $c_0 \in \mathbb{R}$ aus der Gleichung $y_0 = y_p(t_0) + c_0 y_h(t_0)$ als

$$c_0 = y_h(t_0)^{-1}(y_0 - y_p(t_0)).$$

Die eindeutige Lösung y der ursprünglichen inhomogenen DGL ist also

$$y(t) = y_p(t) + c_0 y_h(t)$$

mit c_0, y_p und y_h wie oben.

Die Methode der Variation der Konstanten kann an allgemeinere Fälle linearer DGLen angepasst werden, siehe auch 3.2.8.

Übungsaufgabe 206. (E) Beschreiben und überprüfen Sie die Methode der Variation der Konstanten für inhomogene lineare Differentialgleichungen zweiter Ordnung.

Übungsaufgabe 207. (T) Zeigen Sie, dass $\left\{ \frac{1}{\cos x}, \frac{x}{\cos x} - \sin x \right\}$ ein Fundamentalsystem von $y'' - \frac{2}{\sin x \cos x} y' + y = 0$ ist.

Übungsaufgabe 208. (T) Zeigen Sie, dass $\{x^2, x^2 \ln x\}$ ein Fundamentalsystem von $y'' - \frac{3}{x} y' + \frac{4}{x^2} y = 0$ ist.

Übungsaufgabe 209. (T*) Die DGL $x^2 y'' - x y' + y = 0$ hat offenbar die Lösung $y_1(x) = x$. Bestimmen Sie eine zweite (linear unabhängige) Lösung durch den Reduktionsansatz $y_2(x) = z(x)y_1(x)$.

Übungsaufgabe 210. (T) Eine Lösung von $y'' + \frac{3x-1}{x^2-x} y' + \frac{1}{x^2-x} y = 0$ lautet $y_1(x) = \frac{1}{1-x}$. Bestimmen Sie durch Reduktion eine dazu linear unabhängige Lösung $y_2(x) = z(x)y_1(x)$.

Übungsaufgabe 211. (E) Betrachten Sie die Anfangswertaufgabe

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \dots + a_1y'(x) + a_0y(x) = 0$$

$$y^{(k)}(a) = 0 \quad \text{für } k = 0, \dots, n-2, \quad y^{(n-1)}(a) = 1.$$

Eine Lösung $y_a(x)$ dieser speziellen Anfangswertaufgabe führt zu folgender partikulären Lösung des inhomogenen Problems mit rechter Seite $f(x)$

$$y_p(x) := \int_a^x y_a(x-s+a)f(s) ds.$$

Bestimmen Sie mit dieser Methode eine partikuläre Lösung von

$$y''(x) - 2y'(x) + y(x) = \frac{1}{x^2}e^x.$$

Wählen Sie dazu $a = 1$.

3.2.6 Gekoppelte lineare DGLen erster Ordnung

Inhalt in Kurzfassung: Aus einer linearen DGL erster Ordnung der Gestalt $y'(t) = a(t)y(t)$ mit einer Koeffizientenfunktion a wird im gekoppelten Fall $\mathbf{y}'(t) = A(t)\mathbf{y}(t)$ mit einer quadratischen Matrix $A(t)$ von Funktionen $a_{ij}(t)$. Im Fall konstanter Koeffizienten steht eine Methode zur Verfügung, wo zum Beispiel die Eigenwertstruktur von A sowie Exponentialmatrizen (als Verallgemeinerung von Exponentialfunktionen) auftreten. Es folgt eine relativ ausführliche Darstellung dieser Methode.

Die Theorie linearer Differentialgleichungssysteme (1. Ordnung) ist nicht nur für Anwendungen wichtig, sondern umfasst auch, wie wir am Ende dieses Abschnitts sehen werden, die Theorie gewöhnlicher linearer Differentialgleichungen n -ter Ordnung. Ein allgemeines homogenes lineares DGL-System 1. Ordnung hat die Form:

$$\mathbf{y}'(t) = A(t)\mathbf{y}(t) \quad \text{bzw.} \quad \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}' = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

mit stetigen Koeffizientenfunktionen $a_{ik} : I \rightarrow \mathbb{R}$ (oder $G \subset \mathbb{C} \rightarrow \mathbb{C}$).

Von spezieller Bedeutung sind DGL-Systeme mit konstanten Koeffizienten, d.h. die Koeffizienten $a_{ik} \in \mathbb{R}$ oder \mathbb{C} sind konstant. Für solche Systeme gibt es eine einheitliche Lösungsformel bzw. ein algorithmisches Lösungsverfahren, die sogenannte *Eigenwert-Eigenvektormethode*, die wir in diesem Abschnitt ausführlich besprechen werden.

Es sei ab nun A stets eine konstante nicht notwendig reguläre $(n \times n)$ -Koeffizientenmatrix mit Koeffizienten $a_{ik} \in \mathbb{R}$ oder \mathbb{C} , $i, k = 1, \dots, n$. Es gilt dann die folgende sehr elegante und besonders für theoretische Zwecke nützliche Lösungsformel:

Satz 3.2.6.1. *Das homogene lineare DGS*

$$\mathbf{y}'(t) = A\mathbf{y}(t) \quad (3.1)$$

hat die allgemeine Lösung

$$\mathbf{y}(t) = e^{tA}\mathbf{c} \quad (3.2)$$

wobei $\mathbf{c} \in \mathbb{R}^n$ bzw. \mathbb{C}^n beliebig gewählt werden kann.

Die in Satz 3.2.6.1 auftretende Matrix e^{tA} ist die sogenannte Exponentialmatrix von tA . Diese ist wie folgt definiert:

Definition 3.2.6.2. *Ist $B = (b_{ij})$ eine reelle oder komplexe nicht notwendig reguläre $(n \times n)$ -Matrix, dann ist die **Exponentialmatrix** $\exp(B) = e^B$ von B definiert durch*

$$e^B = \sum_{m=0}^{\infty} \frac{1}{m!} B^m.$$

Um einzusehen, dass diese Exponentialreihe von Matrizen komponentenweise konvergiert, setzen wir $b = \max\{|b_{ij}| : i, j = 1, \dots, n\}$. Wie man durch Induktion sofort sieht, sind die Elemente $b_{ij}^{(k)}$ der Matrix B^k betragsmäßig durch $n^{k-1}b^k$ beschränkt. Für die Elemente $c_{ij}^{(k)}$ der Partialsumme $\sum_{m=0}^k \frac{1}{m!} B^m$ gilt daher

$$c_{ij}^{(k)} = \delta_{ij} + b_{ij} + \frac{1}{2!} b_{ij}^{(2)} + \dots + \frac{1}{k!} b_{ij}^{(k)}$$

($\delta_{ij} = 1$ für $i = j$ und 0 sonst). Diese Reihe muss aber absolut konvergieren, denn sie hat ja die konvergente Majorante

$$1 + b + \frac{1}{2!} nb^2 + \dots + \frac{1}{k!} n^{k-1} b^k + \dots \quad (< e^{nb}).$$

Für $n = 1$ degeneriert die Exponentialmatrix natürlich zur gewöhnlichen Exponentialfunktion.

Bevor wir uns an den Beweis von Satz 3.2.6.1 machen, benötigen wir noch zwei Hilfsaussagen über Exponentialmatrizen.

Lemma 3.2.6.3. *Es seien A, B zwei $(n \times n)$ -Matrizen ($n \geq 2$) mit $AB = BA$. Dann gilt*

$$\exp(A + B) = \exp(A) \exp(B).$$

Beweis. Der Beweis ist wörtlich identisch mit dem für den Fall $n = 1$ (vgl. dazu das Skriptum Mathematik 1). Man bildet genau wie dort das Cauchyprodukt der Potenzreihen $\exp(A)$ und $\exp(B)$ (die Aussagen über die Multiplikation von Potenzreihen lassen sich zwanglos auf Matrizenreihen übertragen). Die Vertauschbarkeit von A und B braucht man, damit der in den Beweis wesentlich eingehende binomische Lehrsatz seine Gültigkeit behält. (Ist $AB \neq BA$, dann gilt ja nicht einmal $(A + B)^2 = A^2 + 2AB + B^2$.)

■

3 Differentialgleichungen (DGLen)

Folgerung 3.2.6.4. *Es gelten die folgenden Aussagen:*

- (i) *Für jede beliebige $(n \times n)$ -Matrix A ist $\exp(A)$ regulär und hat $\exp(-A)$ als Inverse.*
- (ii) *Ist A reell, dann gilt sogar $\det(\exp(A)) > 0$.*

Beweis. ad (i): Zum Beweis wähle man $B = -A$ in Lemma 3.2.6.3. Da A und $-A$ kommutieren, erhält man $\exp(A)\exp(-A) = \exp(A - A) = \exp(0) = I$ (hier steht 0 für die Nullmatrix); damit ist $\exp(A)$ stets invertierbar und hat $\exp(-A)$ als Inverse.

ad (ii): Es sei nun A reell. Wir führen die Funktion $d_A(t) = \det \exp(tA)$ ein, die für alle $t \in \mathbb{R}$ definiert und offenbar stetig ist. Wie wir eben gesehen haben, ist $\exp(tA)$ niemals singular, das heißt, $d_A(t)$ kann an keiner Stelle $t \in \mathbb{R}$ verschwinden. Wegen der Stetigkeit kann die Funktion $d_A(t)$ also nirgends das Vorzeichen wechseln, sie ist somit entweder überall positiv oder überall negativ. Da aber $d_A(0) = \det \exp(0A) = \det \exp(O) = \det I = 1 > 0$ ist, muss $d_A(t)$ überall positiv sein. Speziell ist $d_A(1) = \det \exp(A) > 0$ für jede reelle Matrix A . ■

Beweis von Satz 3.2.6.1. Wir zeigen zunächst, dass $\mathbf{y}(t) = e^{tA}\mathbf{c}$ tatsächlich eine Lösung von $\mathbf{y}'(t) = A\mathbf{y}(t)$ ist. Dabei verwenden wir, dass alle Komponenten von e^{At} offenbar gleichmäßig konvergente Potenzreihen in t mit Konvergenzradius ∞ sind und daher gliedweises Differenzieren überall erlaubt ist:

$$\begin{aligned} \frac{d}{dt}\mathbf{y}(t) &= \frac{d}{dt} \left(\sum_{k=0}^{\infty} \frac{(At)^k}{k!} \mathbf{c} \right) = \left(\sum_{k=0}^{\infty} \frac{d}{dt} \frac{t^k}{k!} A^k \right) \mathbf{c} = \left(\sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k \right) \mathbf{c} = \\ &= A \left(\sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^{k-1} \right) \mathbf{c}. \end{aligned}$$

Setzen wir nun $j = k - 1$, so erhalten wir

$$\mathbf{y}'(t) = \frac{d}{dt}\mathbf{y}(t) = A \left(\sum_{j=0}^{\infty} \frac{t^j}{j!} A^j \right) \mathbf{c} = Ae^{At}\mathbf{c} = A\mathbf{y}(t)$$

was zu zeigen war.

Wir haben weiters zu zeigen, dass $\mathbf{y}'(t) = A\mathbf{y}(t)$ keine anderen Lösungen als solche von der Bauart $e^{tA}\mathbf{c}$ haben kann. Die Beweisidee ist im wesentlichen die gleiche wie im Fall der DG $y' = ay$. Bei der Übertragung der Argumente auf den n -dimensionalen Fall ist allerdings Vorsicht angebracht, da Matrizen bei Multiplikation nicht mehr generell kommutieren. Es sei also $\mathbf{z}(t)$ irgendein Vektor, der für alle reellen t aus einem gewissen Intervall $I = (a, b)$ (bzw. für alle komplexen t aus einem Gebiet G in \mathbb{C}) dem DGS $\mathbf{z}'(t) = A\mathbf{z}(t)$ genügt. Dann erhalten wir durch Anwendung der Produktregel und Vertauschung der Faktoren A und e^{-tA} :

$$\begin{aligned} \frac{d}{dt}(e^{-tA}\mathbf{z}(t)) &= -Ae^{-tA}\mathbf{z}(t) + e^{-tA}\mathbf{z}'(t) = -e^{-tA}A\mathbf{z}(t) + e^{-tA}\mathbf{z}'(t) = \\ &= -e^{-tA}A\mathbf{z}(t) + e^{-tA}A\mathbf{z}(t) = \mathbf{o}. \end{aligned}$$

Aus dem Eindeutigkeitsatz der Differentialrechnung, angewandt auf die Komponentenfunktionen von $e^{-tA}\mathbf{z}(t)$, folgt nun, dass $e^{-tA}\mathbf{z}(t)$ gleich einem konstanten Vektor $\mathbf{c} \in \mathbb{R}^n$ oder \mathbb{C}^n sein muss. Da nach Korollar 3.2.6.4 die Matrizen e^{-tA} und e^{tA} zueinander invers sind, ist daher $\mathbf{z}(t) = e^{tA}\mathbf{c}$, wie behauptet. (Da die Matrizenreihe e^{tA} für alle $t \in \mathbb{C}$ konvergiert, ist damit gleichzeitig klargeworden, dass jede Lösung von $\mathbf{y}'(t) = A\mathbf{y}(t)$ auf ganz \mathbb{C} , erst recht also auf ganz \mathbb{R} , definiert ist.) ■

Beispiel 3.2.6.5.

Zu lösen ist das (2×2) -DG-System

$$\begin{aligned} y_1'(t) &= y_1(t) + y_2(t) \\ y_2'(t) &= y_2(t) \end{aligned}$$

in Matrixschreibweise also

$$\mathbf{y}'(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}' = A\mathbf{y}(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

Die Potenzen von A lassen sich in diesem Fall leicht induktiv bestimmen. Es gilt für alle $k \geq 1$,

$$A^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.$$

Somit ist

$$I + tA + \frac{1}{2!}(tA)^2 + \cdots + \frac{1}{k!}(tA)^k = \begin{pmatrix} 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^k}{k!} & t + \frac{2t^2}{2!} + \cdots + \frac{kt^k}{k!} \\ 0 & 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^k}{k!} \end{pmatrix}.$$

Die Elemente dieser Matrix sind gerade die McLaurin-Polynome vom Grad k der Funktionen e^t bzw. te^t . Daher lautet die Grenzmatrix

$$e^{tA} = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix}.$$

Nach Satz 3.2.6.1 ist somit

$$\mathbf{y}(t) = e^{tA}\mathbf{c} = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

komponentenweise geschrieben also

$$\begin{aligned} y_1(t) &= c_1 e^t + c_2 te^t, \\ y_2(t) &= c_2 e^t, \end{aligned}$$

die allgemeine Lösung unseres Systems, was man durch Einsetzen auch sofort bestätigt.

Die Aussage (i) von Korollar 3.2.6.4 hat eine wichtige Konsequenz für die Struktur der Lösungsmenge von $\mathbf{y}'(t) = A\mathbf{y}(t)$: Wie wir gesehen haben, ist die Matrix $e^{tA} =$

3 Differentialgleichungen (DGLen)

$(\mathbf{z}_1(t), \dots, \mathbf{z}_n(t))$ stets regulär, hat also jedenfalls n linear unabhängige Spalten. Nach Satz 3.2.6.1 kann die allgemeine Lösung des Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$ in der Form

$$\mathbf{y}(t) = e^{tA}\mathbf{c} = (\mathbf{z}_1(t), \dots, \mathbf{z}_n(t))\mathbf{c} = \mathbf{z}_1(t)c_1 + \dots + \mathbf{z}_n(t)c_n$$

geschrieben werden. Damit erhalten wir direkt die folgende wichtige Aussage.

Folgerung 3.2.6.6. *Es sei A eine $(n \times n)$ -Matrix ($n \geq 2$). Dann bilden die Lösungen des DG-Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$ stets einen n -dimensionalen Vektorraum \mathcal{L} , in welchem die Spaltenvektoren $\mathbf{z}_1(t), \dots, \mathbf{z}_n(t)$ der Exponentialmatrix e^{tA} eine Basis, ein sogenanntes Fundamentalsystem, bilden.*

Bemerkungen:

- (i) Natürlich besitzt der Lösungsraum \mathcal{L} noch eine Fülle anderer Basen. Jede andere Basis $\mathbf{y}_1(t), \dots, \mathbf{y}_n(t)$ von \mathcal{L} erhält man allerdings aus der vorliegenden durch eine Transformation

$$(\mathbf{y}_1(t), \dots, \mathbf{y}_n(t)) = (\mathbf{z}_1(t), \dots, \mathbf{z}_n(t)) \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix}$$

mit einer regulären $n \times n$ -Matrix (w_{ij}) .

- (ii) Es wird bei der Beschreibung der Eigenwert-Eigenvektormethode klarwerden, dass sämtliche Komponentenfunktionen der Lösung von der Form $p_{ij}(t)e^{\lambda_j t}$ sind; die λ_j sind dabei die Eigenwerte von A und die p_{ij} Polynome in t . Genau dann, wenn A eine diagonalisierbare Matrix ist (wenn also n unabhängige Eigenvektoren $\mathbf{s}_1, \dots, \mathbf{s}_n$ existieren), sind alle p_{ij} vom Grad 0 (also Konstante).
- (iii) Auch bei einer nichtkonstanten $n \times n$ -Koeffizientenmatrix $A(t) = (a_{ik}(t))$ lässt sich zeigen, dass die Lösungen von $\mathbf{y}'(t) = A(t)\mathbf{y}(t)$ einen n -dimensionalen Vektorraum bilden. Es gibt dann aber keine universelle Lösungsformel wie im Fall konstanter Koeffizienten, und die Bestimmung eines Fundamentalsystems ist meist schwierig oder praktisch unmöglich.

Für die effektive Berechnung der Lösungen von $\mathbf{y}'(t) = A\mathbf{y}(t)$ ist Satz 3.2.6.1 leider im Allgemeinen

nicht geeignet. Ist jedoch die Koeffizientenmatrix A diagonalisierbar, so kann man e^{tA} leicht explizit ausrechnen und erhält, wie wir im folgenden ausführen, nach einem passenden Basiswechsel das Fundamentalsystem $\mathbf{y}_j(t) = \mathbf{s}_j e^{\lambda_j t}$ ($j = 1, \dots, n$).

Satz 3.2.6.7. *Es sei A eine diagonalisierbare $(n \times n)$ -Matrix ($n \geq 2$) mit den Eigenwerten $\lambda_1, \dots, \lambda_n$ und zugehörigen linear unabhängigen Eigenvektoren $\mathbf{s}_1, \dots, \mathbf{s}_n$. Dann bildet die Menge der Vektoren*

$$\{\mathbf{s}_1 e^{\lambda_1 t}, \mathbf{s}_2 e^{\lambda_2 t}, \dots, \mathbf{s}_n e^{\lambda_n t}\}$$

ein Fundamentalsystem für den Lösungsraum des DG-Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$.

Beweis. Da A diagonalisierbar ist, kann A dargestellt werden in der Form

$$A = SDS^{-1}.$$

Dabei ist die Diagonalmatrix $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ aus den Eigenwerten aufgebaut und die Matrix $S = (\mathbf{s}_1, \dots, \mathbf{s}_n)$ spaltenweise aus den Eigenvektoren. Für beliebige Potenzen A^k mit $k \in \mathbb{N}$ bedeutet dies

$$A^k = SD^kS^{-1}.$$

Wir setzen diese Ausdrücke für A^k in unsere Lösungsformel $\mathbf{y}(t) = e^{tA}\mathbf{c}$ ein, wobei wir immer gleich passend umklammern:

$$\mathbf{y}(t) = e^{tA}\mathbf{c} = \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \right) \mathbf{c} = \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} SD^kS^{-1} \right) \mathbf{c} = S \left(\sum_{k=0}^{\infty} \frac{t^k}{k!} D^k \right) (S^{-1}\mathbf{c}) = Se^{tD}(S^{-1}\mathbf{c}).$$

Beachten wir nun einerseits, dass mit \mathbf{c} auch $\mathbf{a} := S^{-1}\mathbf{c}$ den gesamten \mathbb{R}^n bzw. \mathbb{C}^n durchläuft und andererseits, dass $e^{tD} = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$, so folgt weiters

$$\mathbf{y}(t) = S \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n) \begin{pmatrix} e^{\lambda_1 t} a_1 \\ e^{\lambda_2 t} a_2 \\ \vdots \\ e^{\lambda_n t} a_n \end{pmatrix}.$$

Damit kann die allgemeine Lösung des DG-Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$ in der Form

$$\mathbf{y}(t) = \mathbf{s}_1 e^{\lambda_1 t} a_1 + \mathbf{s}_2 e^{\lambda_2 t} a_2 + \dots + \mathbf{s}_n e^{\lambda_n t} a_n$$

mit $a_1, \dots, a_n \in \mathbb{R}$ oder \mathbb{C} beliebig, dargestellt werden. ■

Beispiel 3.2.6.8.

Gesucht ist ein Fundamentalsystem für das DG-System $\mathbf{y}'(t) = A\mathbf{y}(t)$, wobei

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Die Eigenwerte der Matrix A berechnen sich dabei zu

$$\lambda_{1,2} = -2, \quad \lambda_3 = 1$$

und die zugehörigen Eigenvektoren sind

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

3 Differentialgleichungen (DGLen)

Zusammenfassend können wir das folgende Fundamentalsystem angeben:

$$\begin{aligned} \{ \mathbf{y}_1(t), \mathbf{y}_2(t), \mathbf{y}_3(t) \} &= \{ \mathbf{s}_1 e^{\lambda_1 t}, \mathbf{s}_2 e^{\lambda_2 t}, \mathbf{s}_3 e^{\lambda_3 t} \} = \\ &= \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{-2t}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} e^{-2t}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^t \right\}. \end{aligned}$$

Komponentenweise angeschrieben lautet die Lösung

$$\begin{aligned} y_1(t) &= c_1 e^{-2t} && + c_3 e^t \\ y_2(t) &= && c_2 e^{-2t} + c_3 e^t \\ y_3(t) &= -c_1 e^{-2t} - c_2 e^{-2t} + c_3 e^t \end{aligned}$$

mit Koeffizienten $c_1, c_2, c_3 \in \mathbb{R}$.

Für eine nicht diagonalisierbare (konstante) Koeffizientenmatrix A ist immer noch jeder Vektor der Form $\mathbf{y}(t) = \mathbf{s} e^{\lambda t}$ (\mathbf{s} Eigenvektor zu einem Eigenwert λ von A) eine Lösung des Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$, denn einfaches Einsetzen zeigt

$$\mathbf{y}'(t) = \mathbf{s}(\lambda e^{\lambda t}) = (\lambda \mathbf{s}) e^{\lambda t} = (A\mathbf{s}) e^{\lambda t} = A(\mathbf{s} e^{\lambda t}) = A\mathbf{y}(t).$$

Da es aber jetzt weniger als n unabhängige Eigenvektoren $\mathbf{s}_1, \dots, \mathbf{s}_m$ ($m < n$) gibt, bekommt man auf diese Art eben nur m unabhängige Lösungen, also keinesfalls das ganze Fundamentalsystem. Wir wollen im folgenden erläutern, wie man die restlichen $n - m$ Fundamentallösungen in einem solchen Fall bestimmt. Wir erinnern dazu zunächst an die (2×2) -Matrix

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

aus Beispiel 3.2.6.5, die nicht diagonalisierbar ist. Sie hat den Doppeleigenwert $\lambda_{1,2} = 1$, aber (bis auf Vielfache) nur einen einzigen Eigenvektor

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Dementsprechend bekommt man mit $\mathbf{y}_1(t) = \mathbf{s}_1 e^t$ nur eine Fundamentallösung, die aus dem Lösungsausdruck in Beispiel 3.2.6.5 auch direkt abzulesen ist. Wie man ebenfalls abliest, lautet die zweite Fundamentallösung

$$\mathbf{y}_2(t) = \begin{pmatrix} t \\ 1 \end{pmatrix} e^t,$$

sie ist also von polynomialem Typ. Das folgende Beispiel soll illustrieren, wie man allgemeiner solche polynomiellen Lösung auffinden kann.

Beispiel 3.2.6.9.

Gesucht ist ein Fundamentalsystem für das DG-System $\mathbf{y}'(t) = A\mathbf{y}(t)$, wobei

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

Die Eigenwerte der Matrix A berechnen sich dabei zu

$$\lambda_1 = -1, \quad \lambda_{2,3} = 1$$

und die zugehörigen Eigenvektoren sind

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Damit ergeben sich zwei der drei gesuchten Fundamentallösungen zu

$$\mathbf{y}_1(t) = \mathbf{s}_1 e^{-t} = \begin{pmatrix} e^{-t} \\ -e^{-t} \\ e^{-t} \end{pmatrix}, \quad \mathbf{y}_2(t) = \mathbf{s}_2 e^t = \begin{pmatrix} e^t \\ e^t \\ e^t \end{pmatrix}.$$

Zur Ermittlung der fehlenden dritten Fundamentallösung bestimmt man zunächst einen vom eben ermittelten Eigenvektor \mathbf{s}_2 linear unabhängigen Lösungsvektor \mathbf{u} des Gleichungssystems

$$(A - \lambda_{2,3}I)^2 \mathbf{u} = \mathbf{0},$$

also von

$$(A - I)^2 \mathbf{u} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \mathbf{u} = \begin{pmatrix} 1 & -2 & 1 \\ -1 & 2 & -1 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Offenbar ist jede Lösung des Gleichungssystems $(A - \lambda I)\mathbf{u} = \mathbf{0}$ auch eine Lösung des Systems

$(A - \lambda I)^2 \mathbf{u} = \mathbf{0}$. Das letztere System hat aber größeren Defekt und daher weitere von \mathbf{u} linear

unabhängige Lösungen! (Dabei versteht man unter dem *Defekt* einer $(n \times n)$ -Matrix A , die Zahl $n - \text{rang } A$ also die Dimension des Lösungsraumes des Gleichungssystems $A\mathbf{x} = \mathbf{a}$.) Wir wählen als zu \mathbf{s}_2 linear unabhängige Lösung des obigen Systems

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Mit Hilfe dieses Vektors ergibt sich eine dritte Fundamentallösung zu

$$\mathbf{y}_3(t) = [I + t(A - \lambda_{2,3}I)]\mathbf{u}e^{\lambda_{2,3}t} = \begin{bmatrix} 1-t & t & 0 \\ 0 & 1-t & t \\ -t & t & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^t = \begin{pmatrix} 1-t \\ -t \\ -1-t \end{pmatrix} e^t.$$

3 Differentialgleichungen (DGLen)

Dass $\mathbf{y}_3(t)$ wirklich Lösung unseres DG-Systems ist, sieht man so (wir argumentieren gleich allgemein): Nach Konstruktion von \mathbf{u} gilt $(A - \lambda I)(A - \lambda I)\mathbf{u} = \mathbf{o}$, also

$$A(A - \lambda I)\mathbf{u} = \lambda(A - \lambda I)\mathbf{u}. \quad (3.3)$$

Daraus erhalten wir

$$\mathbf{y}'_3(t) = \frac{d}{dt}[\mathbf{u} + t(A - \lambda I)\mathbf{u}]e^{\lambda t} = [\lambda\mathbf{u} + t\lambda(A - \lambda I)\mathbf{u} + (A - \lambda I)\mathbf{u}]e^{\lambda t}.$$

Einsetzen von (3.3) ergibt daher wie behauptet

$$\mathbf{y}'_3(t) = [tA(A - \lambda I)\mathbf{u} + A\mathbf{u}]e^{\lambda t} = A[t(A - \lambda I)\mathbf{u} + \mathbf{u}]e^{\lambda t} = A\mathbf{y}_3(t).$$

Dass \mathbf{y}_3 auch linear unabhängig von \mathbf{y}_2 ist, folgt einfach aus der Bauart dieser beiden Lösungen: Sie sind von der Form $\mathbf{y}_j(t) = \mathbf{p}_j e^{\lambda t}$, wobei \mathbf{p}_3 und \mathbf{p}_2 Vektorpolynome von verschiedenem Grad sind. Damit haben wir bei unserem Beispiel jetzt 3 unabhängige Lösungen $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ gefunden. Die allgemeine Homogenlösung erhält man daraus wie üblich durch LK:

$$\mathbf{y}(t) = a_1\mathbf{y}_1(t) + a_2\mathbf{y}_2(t) + a_3\mathbf{y}_3(t) \quad (a_1, a_2, a_3 \in \mathbb{R} \text{ oder } \mathbb{C} \text{ beliebig}).$$

Wir beschreiben jetzt das allgemeine Verfahren, die sogenannte *Eigenwert-Eigenvektor-Methode*, zur Bestimmung eines Fundamentalsystems des DG-Systems $\mathbf{y}'(t) = A\mathbf{y}(t)$ bei einer nicht diagonalisierbaren $n \times n$ -Matrix A . Dazu sei λ ein Eigenwert von A mit

algebraischer Vielfachheit v und geometrischer Vielfachheit $g < v$.

So etwas kann natürlich erst ab $v \geq 2$ eintreten, denn g ist sicher mindestens gleich 1. Wir brauchen also nur Doppel- oder Mehrfachwurzeln λ zu betrachten. Zu Einfachwurzeln λ gehören Fundamentallösungen $\mathbf{y}(t) = \mathbf{s}e^{\lambda t}$, \mathbf{s} Eigenvektor zu λ , und diese sind von anderen Fundamentallösungen linear unabhängig.

Eigenwert-Eigenvektor-Methode:

Alle zu einem Mehrfacheigenwert λ gehörigen Fundamentallösungen $\mathbf{y}_1, \dots, \mathbf{y}_v$ erhält man auf folgende Weise:

1. Schritt: Wähle g unabhängige Eigenvektoren $\mathbf{s}_1, \dots, \mathbf{s}_g$, d.h. g unabhängige Lösungen von

$$(A - \lambda I)\mathbf{s} = \mathbf{o}. \quad (3.4)$$

Damit ergeben sich g linear unabhängige Fundamentallösungen

$$\mathbf{y}_1(t) = \mathbf{s}_1 e^{\lambda t}, \quad \dots, \quad \mathbf{y}_g(t) = \mathbf{s}_g e^{\lambda t}.$$

Es fehlen uns nun wegen $g < v$ noch $v - g$ Fundamentallösungen zum Eigenwert λ . Zu ihrer Bestimmung machen wir den

2. Schritt: Da die Matrix $(A - \lambda I)^2$ größeren Defekt hat als $A - \lambda I$, hat das Gleichungssystem

$$(A - \lambda I)^2 \mathbf{u} = \mathbf{o} \quad (3.5)$$

sicher mehr Lösungen als (3.4), darunter offenbar jede Lösung von (3.4). Wir wählen so viele zusätzliche Lösungen \mathbf{u} wie möglich, die vom System $\mathbf{s}_1, \dots, \mathbf{s}_g$ linear unabhängig sind und bezeichnen sie mit

$$\mathbf{u}_{g+1}, \dots, \mathbf{u}_{g+h},$$

wobei $h = \text{Defekt}((A - \lambda I)^2) - \text{Defekt}(A - \lambda I)$. Diese Vektoren sind *keine* Eigenvektoren, sondern werden *Hauptvektoren* 1. Stufe genannt. Als zugehörige Fundamentallösungen ergeben sich:

$$\begin{aligned} \mathbf{y}_{g+1}(t) &= (I + t(A - \lambda I))\mathbf{u}_{g+1}e^{\lambda t}, \\ &\vdots \\ \mathbf{y}_{g+h}(t) &= (I + t(A - \lambda I))\mathbf{u}_{g+h}e^{\lambda t}. \end{aligned}$$

Ist nun $g + h = v$, dann sind wir fertig. Andernfalls fehlen immer noch Fundamentallösungen und wir führen den nächsten Schritt durch:

3. Schritt. Wiederum lässt sich zeigen, dass die Matrix $(A - \lambda I)^3$ größeren Defekt hat als die niedrigere Potenz $(A - \lambda I)^2$. Das bedeutet, dass das Gleichungssystem

$$(A - \lambda I)^3 \mathbf{v} = \mathbf{o}$$

wieder mehr Lösungen als (3.5) hat, darunter offenbar jede Lösung von (3.5) und (3.4). Wir wählen so viele zusätzliche Lösungen \mathbf{v} wie möglich, die vom System aller $\mathbf{s}_1, \dots, \mathbf{s}_g$ und $\mathbf{u}_{g+1}, \dots, \mathbf{u}_{g+h}$, linear unabhängig sind und bezeichnen sie mit

$$\mathbf{v}_{g+h+1}, \dots, \mathbf{v}_{g+h+j},$$

wobei $j = \text{Defekt}((A - \lambda I)^3) - \text{Defekt}(A - \lambda I)^2$. Als Fundamentallösungen kommen neu hinzu

$$\begin{aligned} \mathbf{y}_{g+h+1}(t) &= \left(I + t(A - \lambda I) + \frac{t^2}{2!}(A - \lambda I)^2 \right) \mathbf{v}_{g+h+1}e^{\lambda t}, \\ &\vdots \\ \mathbf{y}_{g+h+j}(t) &= \left(I + t(A - \lambda I) + \frac{t^2}{2!}(A - \lambda I)^2 \right) \mathbf{v}_{g+h+j}e^{\lambda t}. \end{aligned}$$

Nach endlich vielen Schritten der beschriebenen Art haben wir jedenfalls genügend viele linear unabhängige Fundamentallösungen (nämlich v viele) gefunden. Die Anzahl k der nötigen Schritte ist offenbar höchstens gleich $v - g$, denn begonnen haben wir mit g Fundamentallösungen, und in jedem Schritt kommt mindestens eine hinzu. Nach dem letzten Schritt, also dem Lösen von $(A - \lambda I)^k \mathbf{w} = \mathbf{o}$, erhöht sich der Defekt nicht mehr weiter, man hat dann

$$\text{Defekt}((A - \lambda I)^k) = \text{Defekt}((A - \lambda I)^{k+1}) = \dots = v.$$

Übungsaufgabe 212. (T) Berechnen Sie die Lösung des Systems $\mathbf{y}' = \mathbf{A}\mathbf{y}$ mit Anfangswert $\mathbf{y}(0) = (-2, 0, 3)^T$ und

$$\mathbf{A} = \begin{pmatrix} -3 & 7 & -3 \\ -4 & 7 & -2 \\ -3 & 3 & 1 \end{pmatrix}.$$

3.2.7 Homogene Systeme höherer Ordnung

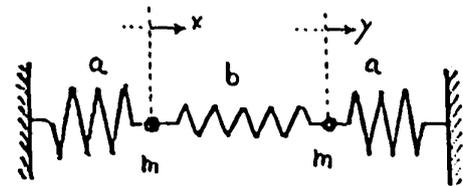
Inhalt in Kurzfassung: Der Übergang von gekoppelten Systemen erster zu solchen höherer Ordnung folgt den bereits aus 3.1.4 bekannten Überlegungen.

Jedes lineare DGS lässt sich auf eines von erster Ordnung zurückführen. Man führt dazu für alle auftretenden Ableitungen neue Funktionsbezeichnungen ein und ersetzt damit das gegebene DGS durch ein höherdimensionales DGS erster Ordnung. Die Vorgangsweise wird ausreichend illustriert durch das folgende Beispiel.

Beispiel 3.2.7.1.

Zwei gekoppelte Federschwinger mit gleichen Massen:

Die Anwendung des Newtonschen Grundgesetzes (Masse mal Beschleunigung ist gleich Summe der einwirkenden Kräfte, hier Rückstellkräfte der Federn) auf jede der beiden schwingenden (gleichen) Massen im skizzierten System führt auf die Bewegungsgleichungen



$$\begin{aligned} x'' &= -ax - b(x - y), \\ y'' &= -ay - b(y - x). \end{aligned}$$

Dabei sind $x(t)$ und $y(t)$ die von der Zeit t abhängigen Auslenkungen der Massen aus der jeweiligen Ruhelage. Sind beide Massen in Ruhelage ($x = y = 0$), dann sollen alle drei Federn entspannt sein. Der Einfachheit halber haben wir $m = 1$ gewählt. Ferner bezeichne $b (> 0)$ die Federkonstante der koppelnden Feder und $a (> 0)$ die Federkonstante jeder der beiden (gleich starken) äußeren Federn. Außer den Federkräften sollen keine weiteren Kräfte auf das System einwirken.

Setzt man

$$y_1 = x, \quad y_2 = y, \quad y_3 = x', \quad y_4 = y',$$

so wird aus dem gegebenen (2×2) -DG-System 2. Ordnung ein (4×4) -DG-System 1. Ordnung $\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t)$ mit

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a-b & b & 0 & 0 \\ b & -a-b & 0 & 0 \end{pmatrix}.$$

Dessen Eigenwerte lauten

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{i\sqrt{a}, -i\sqrt{a}; i\sqrt{a+2b}, -i\sqrt{a+2b}\}.$$

Wegen $a > 0$ und $b > 0$ sind alle Eigenwerte einfach, es ist also von vornherein klar, dass zu den paarweise konjugiert komplexen Eigenwerten vier unabhängige paarweise konjugiert komplexe Eigenvektoren $\mathbf{s}_1, \mathbf{s}_2 = \overline{\mathbf{s}_1}; \mathbf{s}_3, \mathbf{s}_4 = \overline{\mathbf{s}_3}$ existieren (d.h. A ist diagonalisierbar). Nach einfacher Rechnung erhält man, wobei abkürzend $\omega = \sqrt{a}, \eta = \sqrt{a+2b}$ gesetzt worden ist,

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ 1 \\ i\omega \\ i\omega \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 1 \\ 1 \\ -i\omega \\ -i\omega \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} 1 \\ -1 \\ i\eta \\ -i\eta \end{pmatrix}, \quad \mathbf{s}_4 = \begin{pmatrix} 1 \\ -1 \\ -i\eta \\ i\eta \end{pmatrix}.$$

Die Lösungsausdrücke für die ursprünglich gesuchten Auslenkungen x und y liest man nun aus der ersten und zweiten Zeile ab:

$$\begin{aligned} x(t) &= (c_1 e^{i\omega t} + c_2 e^{-i\omega t}) + (c_3 e^{i\eta t} + c_4 e^{-i\eta t}), \\ y(t) &= (c_1 e^{i\omega t} + c_2 e^{-i\omega t}) - (c_3 e^{i\eta t} + c_4 e^{-i\eta t}). \end{aligned}$$

Da man im Allgemeinen nur an reellen Lösung interessiert ist, geht man zu Real- und Imaginärteilen über und erhält so

$$\begin{aligned} x(t) &= (d_1 \cos \omega t + d_2 \sin \omega t) + (d_3 \cos \eta t + d_4 \sin \eta t), \\ y(t) &= (d_1 \cos \omega t + d_2 \sin \omega t) - (d_3 \cos \eta t + d_4 \sin \eta t). \end{aligned}$$

Den Lösungen sieht man noch deutlicher ihre Struktur an, wenn man sie in der Form

$$\begin{aligned} x(t) &= R \sin(\omega t - \alpha) + r \sin(\eta t - \beta), \\ y(t) &= R \sin(\omega t - \alpha) - r \sin(\eta t - \beta) \end{aligned}$$

schreibt. Man erkennt daraus, dass es zwei Grundsicherungen mit Kreisfrequenzen $\omega = \sqrt{a}$ und $\eta = \sqrt{a+2b}$ gibt, die sich (i.a. phasenverschoben und mit verschiedenen Amplituden) überlagern. Man kann diese aber im Experiment auch isoliert beobachten:

- (i) Wählt man $r = 0$, so hat man zu allen Zeitpunkten $x(t) = y(t)$, d.h. beide Massen schwingen synchron mit Kreisfrequenz ω (die koppelnde Feder ist dauernd entspannt, hat also gar keine Wirkung; die Massen schwingen so, als ob es die mittlere Feder gar nicht gäbe).
- (ii) Wählt man dagegen $R = 0$, so hat man $x(t) = -y(t)$, d.h. die Massen schwingen gegensynchron, die mittlere Feder wird jetzt periodisch komprimiert und dilatiert, das System ist also insgesamt steifer als bei (i) und schwingt dementsprechend schneller (die hier relevante Kreisfrequenz $\eta = \sqrt{a+2b}$ ist tatsächlich größer als die bei (i) auftretende $\omega = \sqrt{a}$).

3.2.8 Inhomogene Systeme

Inhalt in Kurzfassung: Für die Behandlung inhomogener Systeme gibt es zur Bestimmung einer partikulären Lösung sowohl Ansatzmethode als auch Variation der Konstanten Analoga für den gekoppelten Fall.

Inhomogen nennt man ein lineares DG-System, wenn es von der Gestalt

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{f}(t) \quad (3.6)$$

ist. Dabei ist $A = A(t) = (a_{kj}(t))$ eine vorerst nicht notwendig konstante $n \times n$ -Matrix mit stetigen Koeffizientenfunktionen a_{kj} auf einem offenen Intervall oder einem Gebiet in \mathbb{C} , und $\mathbf{f} = \mathbf{f}(t)$ der sogenannte *Störvektor*, dessen Komponenten $f_k(t)$ ebenfalls stetig von t abhängen sollen. Genau wie im Fall $n = 1$ zeigt man leicht, dass die allgemeine Lösung eines solchen Systems von folgender Form ist:

$$\mathbf{y}(t) = \mathbf{y}_h(t) + \mathbf{y}_p(t),$$

wobei unter \mathbf{y}_h die allgemeine Lösung des zugehörigen homogenen Systems $\mathbf{y}'_h(t) = A\mathbf{y}_h(t)$ (mit n freien Parametern) zu verstehen ist und für \mathbf{y}_p eine beliebige Lösung des inhomogenen Systems (3.6), eine sogenannte *Partikulärlösung*, genommen werden kann.

Die Variation der Konstanten zur Bestimmung einer Partikulärlösung.

Kennt man ein Fundamentalsystem

$$Y = Y(t) = (\mathbf{y}_1(t), \dots, \mathbf{y}_n(t)), \quad t \in I \text{ bzw. } t \in G \subset \mathbb{C}$$

des homogenen Systems (und damit die allgemeine Homogenlösung $Y\mathbf{c} = \mathbf{y}_1(t)c_1 + \dots + \mathbf{y}_n(t)c_n$), so kann man sich durch die Methode der *Variation der Konstanten* zumindest im Prinzip immer eine Partikulärlösung verschaffen. Man „variiert“ dazu die Homogenlösung, d.h. man wählt für die Partikulärlösung einen Ansatz der Form

$$\mathbf{y}_p(t) = \mathbf{y}_1(t)c_1(t) + \dots + \mathbf{y}_n(t)c_n(t) = Y(t)\mathbf{c}(t),$$

wobei die $c_j(t)$ jetzt *Funktionen* sind, die wir zu bestimmen haben. Ableiten nach der Produktregel und Einsetzen in (3.6) gibt

$$\mathbf{y}'_p(t) = Y(t)\mathbf{c}'(t) + Y'(t)\mathbf{c}(t) = AY(t)\mathbf{c}(t) + \mathbf{f}(t),$$

und da $Y'(t) = AY(t)$ ist (beachte: die Spaltenvektoren von Y sind die Fundamentallösungen \mathbf{y}_j , die ja nach Definition den Beziehungen $\mathbf{y}'_j(t) = A\mathbf{y}_j(t)$ genügen), vereinfacht sich das Ganze zu dem folgenden *linearen Gleichungssystem* mit n Gleichungen für die n Unbekannten $c'_j(t)$:

$$Y(t)\mathbf{c}'(t) = \mathbf{f}(t).$$

Dieses Gleichungssystem lässt sich aber für jedes $t \in I$ bzw. $t \in G$ eindeutig nach $\mathbf{c}'(t)$ auflösen, denn die Systemmatrix $Y(t) = (\mathbf{y}_1(t), \dots, \mathbf{y}_n(t))$ besteht ja aus n Fundamentallösungen, die nach Annahme auf I bzw. G linear unabhängig sind. $Y(t)$ ist also

dort jedenfalls regulär. Die gesuchten Funktionen $c_j(t)$ bekommt man am Ende durch Integration der offenbar stetigen Komponenten $c'_j(t)$ von $\mathbf{c}'(t)$.

Für die Matrix $Y = Y(t)$ hat sich die Bezeichnung *Wronskimatrix* eingebürgert, ihre Determinante heißt *Wronskideterminante*. Wie sich eben gezeigt hat, kommt ihr im Rahmen unseres Problems eine Schlüsselrolle zu. Die Wronskimatrix wird uns auch noch im nächsten Abschnitt bei der Lösung von Anfangswertproblemen wiederbegegnen.

Die Ansatzmethode zur Bestimmung einer Partikulärlösung.

Im Fall, dass A eine *konstante Koeffizientenmatrix* ist und der *Störvektor die spezielle Form*

$$\mathbf{f}(t) = \mathbf{p}(t)e^{(\alpha+i\beta)t}$$

hat mit einem Vektorpolynom \mathbf{p} beliebigen Grades, ist es meist vorteilhafter, mit einem geeigneten Ansatz für die Partikulärlösung zu arbeiten. Da man meist nur an reellen Lösungen interessiert ist, werden wir uns gleich nur mit Störvektoren von einer der folgenden Bauarten befassen:

$$\mathbf{f}(t) = (\mathbf{r}_0 + \mathbf{r}_1 t + \cdots + \mathbf{r}_m t^m)e^{\alpha t} \cos(\beta t) + (\mathbf{s}_0 + \mathbf{s}_1 t + \cdots + \mathbf{s}_m t^m)e^{\alpha t} \sin(\beta t). \quad (\text{i})$$

Wenigstens einer der führenden Koeffizienten \mathbf{r}_m oder \mathbf{s}_m soll dabei $\neq \mathbf{o}$ sein, d.h. wenigstens eines der beiden Vektorpolynome soll genau vom Grad m sein.

$$\mathbf{f}(t) = (\mathbf{r}_0 + \mathbf{r}_1 t + \cdots + \mathbf{r}_m t^m)e^{\alpha t}, \quad (\text{ii})$$

$$\mathbf{f}(t) = \mathbf{r}_0 + \mathbf{r}_1 t + \cdots + \mathbf{r}_m t^m. \quad (\text{iii})$$

Die obigen Fallunterscheidungen sind im Hinblick auf das praktische Rechnen sinnvoll, man käme jedoch mit (i) allein aus, wenn man alles im Reellen rechnen möchte (aus dieser Sicht stellen sich (ii) und (iii) als Spezialfälle von (i) dar mit $\beta = 0$ bzw. $\alpha = \beta = 0$). Lässt man dagegen auch Rechnung im Komplexen zu, dann käme man mit (ii) allein aus, man muss dann allerdings für α gegebenenfalls komplexe Werte zulassen (mit Hilfe der Eulerformeln kann ja (i) leicht in die Form (ii) gebracht werden).

Der Ansatz zum Auffinden einer Partikulärlösung \mathbf{y}_p wird in jedem Fall dem Störvektor nachgebaut und im Fall der sogenannten „Resonanz“ muss der Grad der Polynome im Ansatz erhöht werden. Wir erläutern die Vorgangsweise im Fall (i):

Zuerst ist zu prüfen, ob die komplexe Zahl $\alpha + i\beta$ in der Liste der Eigenwerte von A vorkommt. Ist das nicht der Fall, dann kopiert man die Bauart des Störvektors und setzt an

$$\mathbf{y}_p(t) = (\mathbf{a}_0 + \mathbf{a}_1 t + \cdots + \mathbf{a}_m t^m)e^{\alpha t} \cos(\beta t) + (\mathbf{b}_0 + \mathbf{b}_1 t + \cdots + \mathbf{b}_m t^m)e^{\alpha t} \sin(\beta t).$$

Ist dagegen $\alpha + i\beta$ mit einem der Eigenwerte λ identisch, dann spricht man von *Resonanz*. Hat λ algebraische Vielfachheit a und geometrische Vielfachheit g , dann ist der Grad der Polynome im Ansatz um $v = a - g + 1$ zu erhöhen:

$$\mathbf{y}_p(t) = (\mathbf{a}_0 + \mathbf{a}_1 t + \cdots + \mathbf{a}_{m+v} t^{m+v})e^{\alpha t} \cos(\beta t) + (\mathbf{b}_0 + \mathbf{b}_1 t + \cdots + \mathbf{b}_{m+v} t^{m+v})e^{\alpha t} \sin(\beta t).$$

3 Differentialgleichungen (DGLen)

Achtung: Auch wenn beim Cosinus- oder Sinusterm des Störvektors ein Polynom von niedrigerem Grad steht oder eventuell ganz fehlt, hat man den entsprechenden Term im Ansatz mit dem vollen Polynom vom Grad $m + v$ auszustatten!

Bemerkung: Kennt man die geometrische Vielfachheit von λ nicht, dann kann man den Grad einfach auch um a erhöhen. Für den Fall, dass $g = 1$ (also $a = v$) ist, wird sich dadurch nichts ändern, andernfalls aber ($g > 1$, d.h. $a > v$) wird die Graderhöhung größer ausfallen als eigentlich nötig wäre, d.h. die Rechnung wird für die führenden Koeffizienten $\mathbf{a}_{m+a} = \dots = \mathbf{a}_{m+a-g+2} = \mathbf{0}$ ergeben.

In den Spezialfällen (ii) und (iii) ist mit naheliegenden Modifikationen vorzugehen. In den folgenden Beispielen befassen wir uns jeweils nur mit der Partikulärlösung, die Homogenlösung ist schon aus Beispiel 3.2.6.5 bekannt.

Beispiel 3.2.8.1.

Gesucht ist eine Partikulärlösung für das inhomogene DG-System

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{r} + \mathbf{s}t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{y}(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} t.$$

Die Matrix A ist die aus Beispiel 3.2.6.5. Wir haben schon festgestellt, dass sie den Doppeleigenwert $\lambda = 1$ hat und nicht diagonalisierbar ist, wir haben also $a = 2$ und $g = 1$. Der Störvektor ist ein reines Vektorpolynom vom Grad 1, $\mathbf{f}(t) = (\mathbf{r} + \mathbf{s}t)e^{0t}$, also von der Form (iii). Resonanz liegt nicht vor, da $\alpha = 0$ kein Eigenwert ist, der Polynomgrad bleibt somit gleich 1, und der Ansatz lautet

$$\mathbf{y}_p(t) = \mathbf{a} + \mathbf{b}t.$$

Einsetzen in das gegebene inhomogene DG-System ergibt

$$\mathbf{y}'_p(t) = \mathbf{b} = A\mathbf{y}_p(t) + \mathbf{f}(t) = A(\mathbf{a} + \mathbf{b}t) + \mathbf{r} + \mathbf{s}t.$$

Koeffizientenvergleich (dabei ist es meist von Vorteil mit der höchsten Potenz zu beginnen!) liefert

$$\begin{aligned} t^1 : \dots & \quad A\mathbf{b} = -\mathbf{s}, \\ t^0 : \dots & \quad A\mathbf{a} = -\mathbf{r} + \mathbf{b}. \end{aligned}$$

Da die Matrix A regulär ist, sind das zwei eindeutig lösbar lineare Gleichungssysteme, aus denen sich sukzessive folgende Lösung ergibt:

$$\mathbf{b} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{a} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Beispiel 3.2.8.2.

Gesucht ist eine Partikulärlösung für das inhomogene DG-System

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{r}e^t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{y}(t) + \begin{pmatrix} 1 \\ 4 \end{pmatrix} e^t.$$

Die Matrix A ist wieder dieselbe wie oben. Der Störvektor $\mathbf{f}(t) = \mathbf{r} e^{1t}$ ist von der Form (ii), mit einem Vektorpolynom vom Grad 0. Diesmal ist $\alpha = 1 = \lambda$, es liegt also Resonanz vor, und der Polynomgrad ist beim Ansatz um $v = a - g + 1 = 2$ zu erhöhen:

$$\mathbf{y}_p(t) = (\mathbf{a} + \mathbf{b}t + \mathbf{c}t^2)e^t.$$

Einsetzen in das inhomogene DG-System ergibt

$$\mathbf{y}'_p(t) = (\mathbf{b} + 2\mathbf{c}t)e^t + (\mathbf{a} + \mathbf{b}t + \mathbf{c}t^2)e^t = A(\mathbf{a} + \mathbf{b}t + \mathbf{c}t^2)e^t + \mathbf{r}e^t.$$

Dividieren durch e^t und Koeffizientenvergleich liefert:

$$\begin{aligned} t^2 : \dots & (A - I)\mathbf{c} = \mathbf{o}, \\ t^1 : \dots & (A - I)\mathbf{b} = 2\mathbf{c}, \\ t^0 : \dots & (A - I)\mathbf{a} = \mathbf{b} - \mathbf{r}. \end{aligned}$$

In den drei Gleichungssystemen zur Bestimmung der unbekanntenen Vektoren \mathbf{c} , \mathbf{b} , \mathbf{a} tritt wieder ein und dieselbe (diesmal singuläre) Koeffizientenmatrix auf. Allgemein handelt es sich dabei um die Matrix $A - \mu I$ ($\mu = \alpha + i\beta$), welche genau im Resonanzfall singulär ist, wenn also μ gleich einem Eigenwert λ ist. In unserem Beispiel ist es die Matrix $A - I$, und die ist wegen $\mu = \lambda = 1$ singulär.

Aus dem ersten Gleichungssystem ist ersichtlich, dass \mathbf{c} ein Eigenvektor von A sein muss (das ist im Resonanzfall immer so). Wir haben somit

$$\mathbf{c} = \begin{pmatrix} c_1 \\ 0 \end{pmatrix}.$$

Es ist dabei zu beachten, dass der Eigenvektor \mathbf{c} allgemein anzusetzen ist. Damit gehen wir in das zweite System

$$(A - I)\mathbf{b} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = 2 \begin{pmatrix} c_1 \\ 0 \end{pmatrix},$$

welches als einzige Bedingung $b_2 = 2c_1$ liefert. Wir gehen daher mit

$$\mathbf{b} = \begin{pmatrix} b_1 \\ 2c_1 \end{pmatrix}$$

in das dritte System

$$(A - I)\mathbf{a} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ 2c_1 \end{pmatrix} - \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$

Der Nullzeile auf der linken Seite steht jetzt der nichttriviale Eintrag $2c_1 - 4$ auf der rechten Seite gegenüber. Damit ist das System also nur dann lösbar, wenn diese rechte Seite verschwindet, wir erhalten damit $c_1 = 2$ (hätte man \mathbf{c} nicht allgemein angesetzt, so könnte sich an dieser Stelle schon ein Widerspruch ergeben). Aus der ersten Zeile des

3 Differentialgleichungen (DGLen)

letzten Systems ergibt sich noch $a_2 = b_1 - 1$.

Zusammenfassend haben wir gefunden, dass die insgesamt sechs zu bestimmenden Komponenten $a_1, a_2, b_1, b_2, c_1, c_2$ nur den vier Bedingungen

$$c_2 = 0, \quad c_1 = 2, \quad b_2 = 4, \quad a_2 = b_1 - 1$$

unterworfen sind. Das Resultat ist somit keineswegs eindeutig, wir können hier a_1 und b_1 nach Belieben wählen, z.B. $a_1 = b_1 = 0$. Das ergibt eine Partikulärlösung

$$\mathbf{y}_p(t) = (\mathbf{a} + \mathbf{b}t + \mathbf{c}t^2) e^t = \left(\begin{pmatrix} 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \end{pmatrix} t + \begin{pmatrix} 2 \\ 0 \end{pmatrix} t^2 \right) e^t.$$

Es können sich hier je nach Wahl der freien Parameter verschiedene Partikulärlösungen ergeben, doch unterscheiden sich diese selbstverständlich nur um Homogenlösungen.

Beispiel 3.2.8.3.

Gesucht ist der Ansatz zur Bestimmung einer Partikulärlösung für das inhomogene DG-System

$$\mathbf{y}'(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{y} + (\mathbf{r} + \mathbf{s}t + \mathbf{u}t^2) e^{3t} \sin(2t).$$

Wegen $\alpha + i\beta = 3 + 2i \neq 1$ liegt keine Resonanz vor, der Partikuläranatz muss somit lauten

$$\mathbf{y}_p(t) = (\mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{a}_2 t^2) e^{3t} \cos(2t) + (\mathbf{b}_0 + \mathbf{b}_1 t + \mathbf{b}_2 t^2) e^{3t} \sin(2t).$$

Achtung: Im Störvektor fehlt hier der Cosinusterm gänzlich; trotzdem hat man ihn im Ansatz mit dem vollen Polynom vom Grad $m + v$ auszustatten.

Übungsaufgabe 213. (T) Für das folgende lineare Differentialgleichungssystem gebe man ein Fundamentalsystem an:

$$\mathbf{x}'(t) = \begin{pmatrix} 1 & 1 \\ 1 & \frac{t+1}{t-1} \end{pmatrix} \mathbf{x}(t).$$

Anleitung: Suchen Sie zuerst eine Fundamentallösung $\mathbf{x}_1(t)$, deren Komponenten jeweils ein Polynom sind. Machen Sie dann den sogenannten Reduktionsansatz

$$\mathbf{x}_2(t) = c(t)\mathbf{x}_1(t) + \mathbf{z}(t),$$

wobei $\mathbf{z}(t)$ ein Vektor ist, dessen letzte Komponente 0 ist.

Übungsaufgabe 214. (T) Finden Sie eine partikuläre Lösung des linearen Differentialgleichungssystem

$$\mathbf{x}'(t) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \mathbf{x}(t) - \begin{pmatrix} 2+t \\ 2t \end{pmatrix}, \quad \mathbf{x}(0) = \begin{pmatrix} 1 \\ 4 \end{pmatrix}.$$

3.2.9 Das Anfangswertproblem

Inhalt in Kurzfassung: Lösungen expliziter DGLen werden, wie wir bereits wissen, in der Regel erst durch Anfangsbedingungen eindeutig. Dies soll nun am Beispiel linearer DGLen illustriert werden.

Wir zeigen, dass die n Parameter in der Lösung eines inhomogenen linearen DG-Systems (mit nicht notwendig konstanter Koeffizientenmatrix)

$$\mathbf{y}'(t) = A(t)\mathbf{y}(t) + \mathbf{f}(t)$$

stets in eindeutiger Weise so gewählt werden können, dass die Lösung einer beliebig vorgeschriebenen „Anfangsbedingung“ (kurz: AB)

$$\mathbf{y}(t_0) = \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

genügt. Es sei I ein offenes Intervall in \mathbb{R} bzw. G ein Gebiet in \mathbb{C} , das t_0 enthält, und auf dem die allgemeine Lösung

$$\mathbf{y}(t) = Y(t)\mathbf{c} + \mathbf{y}_p(t) = (\mathbf{y}_1(t), \dots, \mathbf{y}_n(t)) \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + \begin{pmatrix} y_{p1}(t) \\ \vdots \\ y_{pn}(t) \end{pmatrix}$$

des inhomogenen Systems erklärt ist und die Spaltenvektoren $\mathbf{y}_1(t), \dots, \mathbf{y}_n(t)$ der Matrix $Y(t)$ ein System linear unabhängiger Homogenlösungen bilden. (Ist speziell A konstant und $\mathbf{f}(t)$ eine LK von Ausdrücken der früher betrachteten Form $\mathbf{p}(t)e^{(\alpha+i\beta)t}$, so kann man $I = \mathbb{R}$ bzw. $G = \mathbb{C}$ nehmen). Einsetzen der AB führt sofort auf das lineare Gleichungssystem

$$Y(t_0)\mathbf{c} = -\mathbf{y}_p(t_0) + \mathbf{a}.$$

Da die *Wronskimatrix* $Y(t_0)$ *regulär* ist, ist dieses Gleichungssystem für jede Vorgabe von \mathbf{a} *eindeutig* nach c_1, \dots, c_n *auflösbar*.

Beispiel 3.2.9.1.

Gesucht ist die Lösung des Anfangswertproblems

$$\mathbf{y}'(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{y}(t) + \begin{pmatrix} 1 \\ 4 \end{pmatrix} e^t, \quad \mathbf{y}(1) = \mathbf{a} = \begin{pmatrix} 2e \\ 4e \end{pmatrix}.$$

Die Resultate der Beispiele 3.2.6.5 und 3.2.8.2 ergeben zusammen die allgemeine Lösung

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{y}_h(t) + \mathbf{y}_p(t) = Y(t)\mathbf{c} + \mathbf{y}_p(t) = \\ &= \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 4 \end{pmatrix} t + \begin{pmatrix} 2 \\ 0 \end{pmatrix} t^2 \right\} e^t. \end{aligned}$$

3 Differentialgleichungen (DGLen)

Einsetzen der AB führt auf das Gleichungssystem $Y(1) \mathbf{c} = -\mathbf{y}_p(1) + \mathbf{a}$, also

$$\begin{pmatrix} e & e \\ 0 & e \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = - \begin{pmatrix} 2e \\ 3e \end{pmatrix} + \begin{pmatrix} 2e \\ 4e \end{pmatrix}.$$

Dieses hat die eindeutige Lösung

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Somit lautet die gesuchte Lösung

$$\begin{aligned} y_1(t) &= (-1 + t + 2t^2)e^t, \\ y_2(t) &= 4te^t. \end{aligned}$$

3.2.10 Lösungsmethode bei konstanten Koeffizienten

Inhalt in Kurzfassung: Nochmals wird die Rückführung (einfacher) linearer DGLen höherer Ordnung auf (gekoppelte) lineare DGLen erster Ordnung im Sinne von 3.1.4 durchgespielt.

Die homogene Gleichung:

Gewöhnliche *homogene* lineare Differentialgleichungen n -ter Ordnung lassen sich auf einfache Weise in lineare DG-Systeme erster Ordnung „übersetzen“. Wir demonstrieren das Prinzip dieser Übersetzung am Beispiel der allgemeinsten homogenen linearen Differentialgleichung vierter Ordnung

$$y^{(4)}(t) + a(t)y'''(t) + b(t)y''(t) + c(t)y'(t) + d(t)y(t) = 0 \quad (3.7)$$

mit stetigen Koeffizientenfunktionen $a(t), b(t), c(t), d(t)$, an dem man schon alles Wesentliche sieht. Um diese DG zu einem DGS erster Ordnung umzuschreiben, setzen wir

$$y_1 := y, \quad y_2 := y', \quad y_3 := y'', \quad y_4 := y''' \quad (3.8)$$

und erhalten damit aus (3.7) und (3.8)

$$\begin{array}{rcccc} y_1' & = & y' & = & y_2 \\ y_2' & = & y'' & = & y_3 \\ y_3' & = & y''' & = & y_4 \\ y_4' & = & y^{(4)} & = & -d(t)y_1 - c(t)y_2 - b(t)y_3 - a(t)y_4 \end{array}$$

Wir sind also, wie angekündigt, auf ein lineares DGS mit 4 Differentialgleichungen 1. Ordnung für 4 unbekannte Funktionen geführt worden (dabei interessiert uns hier allerdings nur $y_1 = y$). In Matrixschreibweise $\mathbf{y}'(t) = A(t)\mathbf{y}(t)$ sieht das Ganze so aus:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}' = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ -d(t) & -c(t) & -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}. \quad (3.9)$$

Die sogenannte *Begleitmatrix* einer linearen Differentialgleichung ist also von spezieller Bauart.

Wir wollen nun das System (3.9) für den Fall konstanter Koeffizienten $a, b, c, d \in \mathbb{R}$ mit Hilfe der Eigenwert-Eigenvektor-Methode lösen. Die Eigenwerte $\lambda_1, \dots, \lambda_4$ von A ergeben sich als Nullstellen des charakteristischen Polynoms

$$p(\lambda) = \det \begin{pmatrix} -\lambda & 1 & & \\ & -\lambda & 1 & \\ & & -\lambda & 1 \\ -d & -c & -b & -a - \lambda \end{pmatrix} = \lambda^4 + a\lambda^3 + b\lambda^2 + c\lambda + d.$$

Man nennt die Eigenwerte von A auch die *charakteristischen Zahlen* der Differentialgleichung (3.7), von der wir ausgegangen sind. Ist λ ein Eigenwert, so erhält man jeden zugehörigen Eigenvektor aus dem Gleichungssystem

$$(A - \lambda I)\mathbf{s} = \begin{pmatrix} -\lambda & 1 & & \\ & -\lambda & 1 & \\ & & -\lambda & 1 \\ -d & -c & -b & -a - \lambda \end{pmatrix} \begin{pmatrix} r \\ s \\ u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Der Rang von $A - \lambda I$ ist genau 3 (bei einer $(n \times n)$ -Begleitmatrix genau $n - 1$), denn einerseits muss er kleiner als 4 sein, da für einen Eigenwert (und nur für einen solchen) die Matrix $A - \lambda I$ jedenfalls singular ist, andererseits ist er mindestens gleich 3, da in unserem speziellen Beispiel die ersten 3 Zeilen dieser Matrix ersichtlich unabhängig sind. Die volle Information unseres Gleichungssystems ist daher schon in den ersten drei Zeilen enthalten. Man kann also die letzte Zeile ignorieren und erhält

$$\begin{aligned} -\lambda r + s &= 0 & \Rightarrow & s = \lambda r, \\ -\lambda s + u &= 0 & \Rightarrow & u = \lambda s = \lambda^2 r, \\ -\lambda u + v &= 0 & \Rightarrow & v = \lambda u = \lambda^3 r, \end{aligned}$$

also die einparametrische Lösungsschar

$$\begin{pmatrix} r \\ s \\ u \\ v \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{pmatrix} t, \quad t \in \mathbb{R} \text{ beliebig.}$$

Sind die Eigenwerte $\lambda_1, \dots, \lambda_4$ alle einfach, also paarweise verschieden, so hat A die vier linear unabhängigen Eigenvektoren

$$\mathbf{s}_1 = \begin{pmatrix} 1 \\ \lambda_1 \\ \lambda_1^2 \\ \lambda_1^3 \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} 1 \\ \lambda_2 \\ \lambda_2^2 \\ \lambda_2^3 \end{pmatrix}, \quad \mathbf{s}_3 = \begin{pmatrix} 1 \\ \lambda_3 \\ \lambda_3^2 \\ \lambda_3^3 \end{pmatrix}, \quad \mathbf{s}_4 = \begin{pmatrix} 1 \\ \lambda_4 \\ \lambda_4^2 \\ \lambda_4^3 \end{pmatrix}.$$

3 Differentialgleichungen (DGLen)

Baut man nun aus dem Fundamentalsystem

$$\mathbf{y}_1(t) = \mathbf{s}_1 e^{\lambda_1 t}, \quad \mathbf{y}_2(t) = \mathbf{s}_2 e^{\lambda_2 t}, \quad \mathbf{y}_3(t) = \mathbf{s}_3 e^{\lambda_3 t}, \quad \mathbf{y}_4(t) = \mathbf{s}_4 e^{\lambda_4 t}$$

die allgemeine Lösung

$$\mathbf{y} = a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2 + a_3 \mathbf{y}_3 + a_4 \mathbf{y}_4$$

auf und liest daraus den Ausdruck für die erste Komponente ab, so ergibt sich

$$y(t) = a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t} + a_3 e^{\lambda_3 t} + a_4 e^{\lambda_4 t}.$$

Diese Ergebnisse lassen sich auf den allgemeinen Fall einer linearen Differentialgleichung n -ter Ordnung mit konstanten Koeffizienten wie folgt übertragen: Jeder Eigenwert λ der Begleitmatrix einer solchen DG hat *geometrische Vielfachheit* 1, es gibt nämlich bis auf Vielfache nur einen einzigen Eigenvektor mit Komponenten $1, \lambda, \lambda^2, \dots, \lambda^{n-1}$, gleichgültig welche *algebraische Vielfachheit* der Eigenwert λ hat. Hat die betreffende DG also auch nur einen einzigen mehrfachen Eigenwert, so gibt es insgesamt offenbar weniger als n unabhängige Eigenvektoren, d.h. die

Begleitmatrix ist dann nicht diagonalisierbar. Sind die Eigenwerte paarweise verschieden, dann lautet die aus den Eigenvektoren zusammengesetzte Matrix

$$S = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \dots & \lambda_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix}.$$

Das ist eine sogenannte *Vandermonde*-Matrix. Sie ist genau dann regulär, wenn die λ_j paarweise verschieden sind, was aus der Bauart der „Vandermonde-Determinante“ unmittelbar zu erkennen ist. Diese hat nämlich, wie man induktiv zeigt, die Faktorisierung

$$\det S = \prod_{1 \leq j < k \leq n} (\lambda_k - \lambda_j) = \prod_{k=2}^n \prod_{j=1}^{k-1} (\lambda_k - \lambda_j).$$

Zur Lösung gewöhnlicher linearer Differentialgleichungen genügt es also ein lineares DG-System 1. Ordnung mit der entsprechenden Begleitmatrix zu lösen. Diese Methode ist jedoch für die praktische Lösung von Differentialgleichungen höherer Ordnung in der Regel ungeeignet, daher geben wir hier eine gesonderte zusammenfassende Darstellung. Eine Begründung aller folgenden Aussagen ließe sich durch eine konsequente Anwendung der Eigenwert-Eigenvektormethode auf das entsprechende DG-System geben. Wir werden diese jedoch nicht ausführen.

Hat man eine lineare homogene Differentialgleichung mit konstanten Koeffizienten der Form

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \dots + a_1y'(x) + a_0y(x) = 0 \quad (3.10)$$

zu lösen, dann kann man den *Ansatz*

$$y(x) = e^{\lambda x}$$

in die DG einsetzen, wobei λ eine noch zu bestimmende Konstante ist. Diese Funktion und ihre Ableitungen $y'(x) = \lambda e^{\lambda x}$, $y''(x) = \lambda^2 e^{\lambda x}$, \dots , $y^{(n)}(x) = \lambda^n e^{\lambda x}$ setzt man ein und erhält

$$(\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0) e^{\lambda x} = 0.$$

Da $e^{\lambda x}$ niemals verschwinden kann, muss λ eine Lösung der *charakteristischen Gleichung*

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0 \quad (3.11)$$

sein. Umgekehrt liefert jede Lösung von (3.11) eine Lösung $e^{\lambda x}$ von (3.10). Die Lösungen von (3.11) werden *charakteristische Zahlen* genannt und sind nichts anderes als die Eigenwerte der Begleitmatrix der DG (3.10). Wenn alle n Lösungen der charakteristischen Gleichung einfach sind, so erhält man auf diese Weise n unabhängige Lösungen, also ein Fundamentalsystem. Wie wir gesehen haben, bildet die Lösungsmenge einer homogenen linearen DG n -ter Ordnung stets einen n -dimensionalen Vektorraum. Ist hingegen auch nur eine der Nullstellen mehrfach, dann fehlen noch Fundamentallösungen, und es muss beschrieben werden, wie man sich diese beschafft.

Es sei also λ eine Lösung von (3.11) mit Vielfachheit $v \geq 1$. Dann bilden die insgesamt v Funktionen

$$e^{\lambda x}, \quad x e^{\lambda x}, \quad \dots, \quad x^{v-1} e^{\lambda x}$$

ein System linear unabhängiger Lösungen von (3.10) (bei $v = 1$ hat man lediglich $e^{\lambda x}$). Auf diese Weise arbeitet man alle Nullstellen ab. Zusammengesetzt ergeben die Teilsysteme gerade die richtige Anzahl unabhängiger Lösungen und damit ein Fundamentalsystem $\{y_1, y_2, \dots, y_n\}$, aus dem man die allgemeine Homogenlösung zusammenbaut:

$$y = c_1 y_1 + c_2 y_2 + \dots + c_n y_n, \quad c_1, \dots, c_n \in \mathbb{R} \text{ beliebig.}$$

Beispiel 3.2.10.1.

Gesucht ist die *allgemeine* Lösung der Differentialgleichung

$$y''' - 2y'' - y' + 2y = 0.$$

Die charakteristische Gleichung der DG ist gegeben durch $\lambda^3 - 2\lambda^2 - \lambda + 2 = 0$, womit sich die charakteristischen Zahlen ergeben zu

$$\{\lambda_1, \lambda_2, \lambda_3\} = \{-1, 1, 2\}$$

und damit ein Fundamentalsystem gegeben ist durch

$$\{e^{-x}, e^x, e^{2x}\}.$$

Beispiel 3.2.10.2.

3 Differentialgleichungen (DGLen)

Gesucht ist ein Fundamentalsystem für die Differentialgleichung

$$y^{(4)} - 3y''' + 3y'' - y' = 0.$$

Die charakteristische Gleichung der DG ist gegeben durch $\lambda^4 - 3\lambda^3 + 3\lambda^2 - \lambda = 0$, womit sich die charakteristischen Zahlen ergeben zu

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{0, 1, 1, 1\}$$

und damit ein Fundamentalsystem gegeben ist durch

$$\{1, e^x, xe^x, x^2e^x\}.$$

Beispiel 3.2.10.3.

Gesucht ist ein Fundamentalsystem für die Differentialgleichung

$$y^{(4)} - 8y''' + 26y'' - 40y' + 25y = 0.$$

Die charakteristische Gleichung der DG ist gegeben durch

$$\lambda^4 - 8\lambda^3 + 26\lambda^2 - 40\lambda + 25 = (\lambda^2 - 4\lambda + 5)^2 = 0,$$

womit sich die charakteristischen Zahlen ergeben zu

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{2 + i, 2 + i, 2 - i, 2 - i\}$$

und damit ein komplexes Fundamentalsystem gegeben ist durch

$$\{e^{(2+i)x}, xe^{(2+i)x}, e^{(2-i)x}, xe^{(2-i)x}\}.$$

Durch Ablesen der Real- und Imaginärteile erhält man daraus ein reelles Fundamentalsystem

$$\{e^{2x} \cos x, xe^{2x} \cos x, e^{2x} \sin x, xe^{2x} \sin x\}.$$

3.2.11 Das Anfangswertproblem für lineare DG

Inhalt in Kurzfassung: Die Untersuchungen aus 3.2.10 werden in Hinblick auf das Anfangswertproblem fortgesetzt.

Wir zeigen, dass die n Parameter in der Lösung einer inhomogenen linearen Differentialgleichung n -ter Ordnung (mit nicht notwendig konstanten Koeffizienten)

$$y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \cdots + a_1(x)y'(x) + a_0(x)y(x) = f(x)$$

stets in eindeutiger Weise so gewählt werden können, dass die Lösung den Anfangsbedingungen

$$y(x_0) = \eta_0, \quad y'(x_0) = \eta_1, \quad \cdots, \quad y^{(n-1)}(x_0) = \eta_{n-1}$$

mit beliebig vorgegebenen Werten $\eta_k \in \mathbb{C}$ genügt. Zum Beweis könnte man auf das zugeordnete DG-System zurückgreifen, doch ist die „Rückübersetzung“ recht undurchsichtig. Daher ist eine gesonderte Darstellung vorzuziehen. Es sei I ein offenes Intervall in \mathbb{R} bzw. G ein Gebiet in \mathbb{C} , das x_0 enthält, und auf dem die allgemeine Lösung

$$y(x) = y_1(x) c_1 + y_2(x) c_2 + \cdots + y_n(x) c_n + y_p(x)$$

des inhomogenen Systems erklärt ist und die Funktionen $y_1(x), \dots, y_n(x)$ ein System linear unabhängiger Homogenlösungen bilden (sind speziell die Koeffizienten a_k konstant und ist f eine LK von Ausdrücken der Form $p(x) e^{(\alpha+i\beta)x}$, so kann man $I = \mathbb{R}$ bzw. $G = \mathbb{C}$ nehmen). Einsetzen der AB führt unmittelbar auf das lineare Gleichungssystem

$$Y(x) \mathbf{c} = \begin{pmatrix} y_1(x_0) & y_2(x_0) & \cdots & y_n(x_0) \\ y_1'(x_0) & y_2'(x_0) & \cdots & y_n'(x_0) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x_0) & y_2^{(n-1)}(x_0) & \cdots & y_n^{(n-1)}(x_0) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} -y_p(x_0) & + \eta_0 \\ -y_p'(x_0) & + \eta_1 \\ \vdots & \vdots \\ -y_p^{(n-1)}(x_0) & + \eta_{n-1} \end{pmatrix}$$

Die Matrix Y ist nichts anderes als die *Wronskimatrix*, die uns schon beim Anfangswertproblem für DG-Systeme begegnet ist. Sie ist wegen der Unabhängigkeit der Homogenlösungen y_1, \dots, y_n auf I bzw. G an jeder Stelle $x_0 \in I$ (bzw. $x_0 \in G$) *regulär*. Daher ist das Anfangswertproblem für jedes x_0 und jeden vorgegebenen Vektor $(\eta_0, \dots, \eta_{n-1})$ eindeutig lösbar.

Dass $Y(x)$ auf G tatsächlich genau dann regulär ist, wenn die erste Zeile aus einem System von (auf G) unabhängigen Funktionen y_1, y_2, \dots, y_n aufgebaut ist, sieht man so: Unabhängigkeit ist gleichbedeutend damit, dass sich die Funktionen y_1, y_2, \dots, y_n auf G nur trivial zur Nullfunktion linear kombinieren lassen, anders ausgedrückt:

$$\begin{aligned} \text{Das Bestehen der Identität} \quad c_1 y_1(x) + c_2 y_2(x) + \cdots + c_n y_n(x) = 0 \quad (\forall x \in G) & \quad (*) \\ \text{impliziert} \quad c_1 = c_2 = \cdots = c_n = 0. & \end{aligned}$$

Leitet man die Identität (*) nach x wiederholt ab, so gewinnt man aus ihr die n ebenfalls auf G bestehenden Identitäten

$$\left. \begin{aligned} c_1 y_1(x) &+ c_2 y_2(x) &+ \cdots &+ c_n y_n(x) &= 0 \\ c_1 y_1'(x) &+ c_2 y_2'(x) &+ \cdots &+ c_n y_n'(x) &= 0 \\ &\vdots &&& \\ c_1 y_1^{(n-1)}(x) &+ c_2 y_2^{(n-1)}(x) &+ \cdots &+ c_n y_n^{(n-1)}(x) &= 0 \end{aligned} \right\} Y(x) \mathbf{c} = \mathbf{o}. \quad (**)$$

Bestehen umgekehrt alle Identitäten (**) gleichzeitig, so trivialerweise auch (*). Das bedeutet: (*) ist äquivalent damit, dass das homogene Gleichungssystem (**) für alle $x \in G$ erfüllt ist. Soll dieses aber für jedes $x \in G$ nur die triviale Lösung $\mathbf{c} = \mathbf{o}$ haben, so muss $Y(x)$ auf ganz G regulär sein, und ist umgekehrt $Y(x)$ auf G regulär, so hat (*) nur die triviale Lösung, das System $\{y_1, y_2, \dots, y_n\}$ ist dann also auf G unabhängig.

3.3 Lösungsverfahren für ausgewählte nichtlineare DGLen

In diesem Abschnitt besprechen wir einige weitere Methoden, mit denen manche DGLen gelöst werden können. Die Schlagworte dazu lauten Potenzreihenansatz (3.3.1), Trennung der Variablen (3.3.2), exakte DGLen (3.3.3), integrierende Faktoren (3.3.4), Variablensubstitution (3.3.5) und numerische Methoden (3.3.6).

3.3.1 Potenzreihenansatz

Inhalt in Kurzfassung: In Mathematik 1 sind wir auf die Exponentialreihe gestoßen, indem wir eine Funktion f mit $f(0) = 1$ und $f' = f$ gesucht haben und dafür allgemein eine Potenzreihe $f(x) = \sum_{n=0}^{\infty} a_n x^n$ um $x_0 = 0$ mit zunächst unbekanntem Koeffizienten a_n angesetzt haben. Aus der Forderung $f' = f$ lässt sich daraus $(n+1)a_{n+1} = a_n$ für alle n ableiten, was wegen $a_0 = f(0) = 1$ zur eindeutigen Lösung $a_n = \frac{1}{n!}$ für alle $n \in \mathbb{N}$ führt. Diese Methode des sogenannten Potenzreihenansatzes ist sehr allgemein anwendbar und wird nun kurz besprochen.

Wir kennen bereits aus Mathematik 1 ein einfaches Beispiel eines Potenzreihenansatzes, als wir uns die Aufgabe stellten, eine Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

mit positivem Konvergenzradius zu finden, so dass $f' = f$ und $f(0) = 1$ gilt, die also Lösung der Anfangswertaufgabe $y' = y$, $y(0) = 1$ ist. Wir sind dabei notgedrungen auf die Potenzreihe mit den (rekursiv sich ergebenden) Gliedern $a_n = \frac{1}{n!}$ gestoßen. Von der konnten wir zeigen, dass sie auf ganz \mathbb{R} konvergiert und tatsächlich die gegebene Anfangswertaufgabe löst. (Außerdem konnten wir zeigen, dass die durch diese Potenzreihe dargestellte reelle Funktion f die Funktionalgleichung $f(x+y) = f(x) \cdot f(y)$ erfüllt und deshalb eine Exponentialfunktion ist, deren Basis wir e nannten. Das ist aber für das Folgende nicht der entscheidende Punkt.)

Diese Vorgangsweise, dass man gesuchte Funktionen als Potenzreihen mit zunächst unbekanntem Koeffizienten a_n anschreibt und versucht, diese a_n aus den Vorgaben zu ermitteln, heißt **Potenzreihenansatz**. Sie ist deshalb bei vielen DGLen zielführend, weil Potenzreihen gliedweise differenziert werden, die Bildung ihrer Ableitungen also kein Problem darstellt. Kommen auch andere Operationen vor, so sind natürlich mannigfaltige Komplikationen denkbar. Treten beispielsweise Multiplikationen auf, so müssen die entsprechenden Cauchyprodukte gebildet werden. Auch das ist bei Potenzreihen noch routinemäßig möglich. Bei Divisionen, Verkettungen etc. wird das Verfahren aber immer komplizierter.

Potenzreihenansätze führen typischerweise zu Rekursionsgleichungen für die Koeffizienten a_n . (Im Fall der Exponentialreihe war es die Rekursion $a_{n+1} = \frac{a_n}{(n+1)!}$ mit Startwert $a_0 = 1$.) Es ist aber keineswegs garantiert, dass sich diese Rekursionsgleichungen leichter explizit lösen lassen als die ursprüngliche DGL. Und selbst wenn man zu einer expliziten

3.3 Lösungsverfahren für ausgewählte nichtlineare DGLen

Formel für die a_n gelangt, heißt das noch nicht, dass man die Funktion f dadurch hinreichend gut versteht. Ist der Konvergenzradius der erhaltenen Reihe positiv, so erhält man aber wenigstens in einer Umgebung der Entwicklungsstelle x_0 (im Beispiel war $x_0 = 0$) Taylorpolynome als lokale Approximationen von f , über deren Güte der Satz von Taylor (siehe Mathematik 1, für mehrere Variablen: 2.4.3) Auskunft gibt. Hier begnügen wir uns mit diesen Bemerkungen und einigen Übungsaufgaben.

Übungsaufgabe 215. (T) Bestimmen Sie für das Anfangswertproblem

$$f'(x) = -xf(x)(1 + f(x)), \quad f(0) = 1$$

durch Potenzreihenansatz $f(x) = c_0 + c_1x + c_2x^2 + \dots$ die Koeffizienten c_0, \dots, c_4 .

Übungsaufgabe 216. (T) Bestimmen Sie für das Anfangswertproblem $y' = e^x(y - y^2) - 1$, $y(0) = 1$, die ersten Glieder der Reihenentwicklung $y = c_0 + c_1x + c_2x^2 + \dots$

Übungsaufgabe 217. (E) Verwenden Sie den Ansatz $y_1(x) = \sqrt{x} \cdot \sum_{n=0}^{\infty} c_n x^n$ um eine Lösung folgender Gleichung zu bestimmen:

$$x^2 y'' + xy' + \left(-\frac{1}{4} + x^2\right) y = 0$$

$$\text{(Lösung: } y_1(x) = c_0 \frac{\sin x}{\sqrt{x}})$$

Übungsaufgabe 218. (E) Verwenden Sie den Reduktionsansatz $y_2(x) = c(x)y_1(x)$ um eine zweite l.u. Lösung der Differentialgleichung zu gewinnen:

$$x^2 y'' + xy' + \left(-\frac{1}{4} + x^2\right) y = 0$$

$$\text{(Lösung: } y_2(x) = c_0 \frac{\cos x}{\sqrt{x}})$$

3.3.2 Implizite Lösungen vermittelt Trennung der Variablen

Inhalt in Kurzfassung: Explizite DGLen der speziellen Form $y' = f(t, y) = g(t)h(y)$ lassen sich durch sogenannte Trennung der Variablen behandeln.

Ein Spezialfall expliziter, aber nicht autonomer und schon gar nicht linearer DGLen

$$y' = f(t, y)$$

sind jene, wo sich f als Produkt zweier Funktionen schreiben lässt, die jeweils nur von einer Variablen abhängen, also $f(t, y) = g(t)h(y)$. Für solche Gleichungen hat sich der auf den ersten Blick begrifflich zwar etwas fragwürdige, jedoch sehr praktische Formalismus mittels Differentialschreibweise eingebürgert, der sich überdies streng rechtfertigen lässt.

Man verwendet die Symbole dt und dy für die sogenannten, an dieser Stelle nicht exakt definierten **Differentiale** für die Variablen t und y . Man stellt sich vor, dass y als

3 Differentialgleichungen (DGLen)

Funktion $y(t)$ in Abhängigkeit von t die Änderungsrate $dy = y' dt$ hat, die sich also nach der Regel „Änderungsrate von t mal Ableitung von y nach t “ ergibt. Es gilt also $y' = \frac{dy}{dt}$. Damit lässt sich die Differentialgleichung $y' = f(t, y) = g(t)h(y)$ nach Multiplikation mit dem Differential dt umschreiben zu

$$\frac{dy}{h(y)} = g(t)dt.$$

(Die Variable y kommt nur auf der linken, die Variable t nur auf der rechten Seite vor. Man spricht deshalb auch von Methode der **Trennung der Variablen**.)

Diese Gleichung wird nun integriert zu

$$\int \frac{dy}{h(y)} = \int g(t)dt.$$

Wenn es gelingt eine Stammfunktion $H(y)$ von $\frac{1}{h(y)}$ (als Funktion in y) und eine Stammfunktion $G(t)$ von $g(t)$ (als Funktion in t) zu finden, so liefert (unter geeigneten, typischerweise erfüllten Voraussetzungen) für jedes $c \in \mathbb{R}$ der Zusammenhang

$$H(y) = G(t) + c$$

eine implizite Darstellung von y als Funktion von t . (Bezüglich des Themenkreises *implizite Funktionen* sei auf Satz 2.3.4.1 verwiesen.) Die Konstante c kann gegebenenfalls an eine Anfangsbedingung angepasst werden.

Diese Vorgangsweise mag Unbehagen erzeugen, weil der Begriff der Differentiale dt und dy nicht geklärt ist und deshalb auch der Vorgang des Integrierens auf einer rein symbolischen Manipulation zu beruhen scheint, der keine mathematischen Realitäten entsprechen. Bevor wir dieses Unbehagen ausräumen, sei die Funktionsweise dieser Trennung der Variablen aber an einem einfachen Beispiel illustriert.

Gegeben sei die (nichtlineare aber autonome) DGL $y' = y^2$. Mit obiger Notation ist also $h(y) = y^2$ und $g(t) = 1$, was zu

$$H(y) = \int \frac{dy}{y^2} = -\frac{1}{y} + c_1 \quad \text{und} \quad G(t) = \int dt = t + c_2$$

führt. Aus der resultierenden impliziten Gleichung ergibt sich für jedes $c = c_1 - c_2$ eine entsprechende Lösung $y_c(t) = -\frac{1}{t+c}$. Geben wir eine Anfangsbedingung $y(0) = y_0$ vor, so liefert Einsetzen der Stelle $t = 0$ die Beziehung $y_0 = -\frac{1}{0+c} = -\frac{1}{c}$, also ist $c = -y_0^{-1}$ zu setzen. Somit ist

$$y(t) = -\frac{1}{t - y_0^{-1}} = \frac{1}{y_0^{-1} - t}$$

die Lösung der DGL $y' = y^2$ mit Anfangswert $y(0) = y_0$.

Die Rechtfertigung dieser Methode erfolgt, indem man von einer Funktion $y(t)$ in der Variablen t ausgeht, so dass für alle t aus einem geeigneten Definitionsbereich die Gleichung

$$H(y(t)) = G(t) + c$$

3.3 Lösungsverfahren für ausgewählte nichtlineare DGLen

gilt, wobei $H'(y) = \frac{1}{h(y)}$ und $G'(t) = g(t)$ gelte. Ableiten nach t liefert für die linke Seite nach der Kettenregel $y'(t)H'(y(t)) = \frac{y'(t)}{h(y)}$, für die rechte $g(t)$, woraus die Gleichung $\frac{y'(t)}{h(y)} = g(t)$, also $y'(t) = g(t)h(y)$ folgt. Die Funktion $y(t)$ ist also tatsächlich eine Lösung der DGL. Dass c in der beschriebenen Weise an einen gegebenen Anfangswert angepasst werden kann, ergibt sich durch Einsetzen.

Übungsaufgabe 219. (T) Bestimmen Sie mittels Trennung der Variablen alle Lösungen der Differentialgleichung

$$y'(t) = 3y^{2/3}(t) \cos t.$$

Übungsaufgabe 220. (T) Bestimmen Sie die Lösung der folgenden Differentialgleichungen jeweils mit einer geeigneten Methode:

1. $y' = x^2y$, mit Anfangswert $y(0) = 1$.
2. $y' + y = x$, mit Anfangswert $y(0) = 0$. (Hinweis: Potenzreihenansatz)
3. $y' = \frac{x}{y}$, mit Anfangswert $y(1) = 2$.

Übungsaufgabe 221. (T) Lösen Sie möglichst allgemein

1. $y' + \frac{2y}{x} = 4x$.
2. $y' + \frac{1-2x}{x^2}y = 1$.
3. $y' + y = \cos x$.

3.3.3 Exakte DGLen

Inhalt in Kurzfassung: Sei F eine (zweimal stetig differenzierbare) Funktion in zwei Variablen, so dass die Gleichung $F(x, y) = 0$ implizit eine Funktion $y(x)$ definiert. Differentiation nach x gemäß der mehrdimensionalen Kettenregel liefert $F'(x, y(x))' = F_x + F_y y' = 0$. Ist umgekehrt eine DGL der Gestalt $f(x, y) + g(x, y)y' = 0$, so kann man sich fragen, ob es ein F gibt mit $F_x = f$ und $F_y = g$. Nach dem Satz von Schwarz muss in diesem Fall die sogenannte Integrabilitätsbedingung $f_y = F_{xy} = F_{yx} = g_x$ erfüllt sein. Man spricht von einer exakten DGL. Liegt Exaktheit vor, so hat man sehr gute Chancen, durch entsprechende Integration der Komponentenfunktionen f und g die Funktion F und somit eine implizite Darstellung einer Lösung der DGL zu rekonstruieren.

Exakte Differentialgleichungen sind spezielle explizite Differentialgleichungen, für die sich die rechte Seite in einer ganz bestimmten Weise als Quotient schreiben lässt. Um die Idee der Lösungsmethode für solche Gleichungen besser zu verstehen, empfiehlt sich eine etwas andere Notation.

3 Differentialgleichungen (DGLen)

Definition 3.3.3.1. Eine Differentialgleichung der Form

$$(1) \quad f(x, y) + g(x, y)y' = 0 \text{ mit } f, g : G \rightarrow \mathbb{R} \text{ stetig}$$

heißt **exakt**, wenn es eine Funktion $F : G \rightarrow \mathbb{R}$ mit stetigen partiellen Ableitungen gibt mit

$$(2) \quad F_x = f, \quad F_y = g.$$

Man nennt F eine **Stammfunktion** von f und g .

Beispiel. Die Differentialgleichung

$$(2x + y) + (x + 2y)y' = 0$$

ist exakt. Eine Stammfunktion ist

$$F(x, y) = x^2 + xy + y^2.$$

Statt (1) schreibt man auch gerne

$$(1') \quad f(x, y)dx + g(x, y)dy = 0.$$

(Ist die Differentialgleichung exakt und F eine Stammfunktion, dann ist (1') offenbar das sogenannte (totale) Differential von F .)

Lösung einer exakten Differentialgleichung

Wir betreiben wieder **Heuristik**: Ist $y = y(x)$ eine Lösung von (1), so gilt wegen (2)

$$f(x, y(x)) + g(x, y(x))y'(x) = F_x(x, y(x)) + F_y(x, y(x))y'(x) = 0.$$

Nach der Kettenregel für Funktionen in mehreren Veränderlichen gilt also

$$\frac{d}{dx}F(x, y(x)) = 0$$

oder

$$F(x, y(x)) = \text{const.}$$

Wir präzisieren den Satz über exakte Differentialgleichungen:

Satz 3.3.3.2. Es seien $f, g : G \rightarrow \mathbb{R}$ stetig, so dass

$$(1) \quad f(x, y) + g(x, y)y' = 0$$

eine exakte Differentialgleichung ist. F sei eine Stammfunktion von f, g . Ist $(a, b) \in G$ und gilt $g(a, b) \neq 0$, dann gibt es eine Umgebung $U \subset \mathbb{R}$ von a und eine eindeutig bestimmte Lösung $y : U \rightarrow \mathbb{R}$ von (1) mit

$$y(a) = b.$$

Diese Lösung ergibt sich durch Auflösung nach y aus der impliziten Darstellung

$$F(x, y) = F(a, b).$$

Exakte Differentialgleichungen haben angenehme Eigenschaften, die sie gegenüber vielen anderen Differentialgleichungen auszeichnen, z.B. lassen sie sich leicht lösen — wie wir noch sehen werden. Es erhebt sich daher die Frage,

Wie erkennt man eine exakte Differentialgleichung?

Für viele Fälle reicht das folgende Kriterium aus.

Satz (Exaktheitskriterium). Eine Differentialgleichung der Form

$$f(x, y) + g(x, y)y' = 0 \quad (f, g : G \rightarrow \mathbb{R}, \text{ stetig})$$

ist jedenfalls dann exakt, wenn sie folgende Bedingungen erfüllt:

- (i) G ist einfach zusammenhängend (d.h. G hat keine Löcher).
- (ii) f, g haben stetige partielle Ableitungen.
- (iii) **Integrabilitätsbedingung:** $f_y = g_x$ auf G .

Die Integrabilitätsbedingung ist ganz naheliegend: Aus

$$f = F_x, \quad g = F_y$$

folgt wegen der stetigen Differenzierbarkeit von F nach dem Satz von Schwarz

$$f_y = F_{xy} = F_{yx} = g_x.$$

Das Exaktheitskriterium folgt aus einem Satz über Vektorfelder.

Beispiel.

$$2xy^3 + (3x^2y^2 - 2)y' = 0$$

ist exakt, denn

$$G = \mathbb{R}^2, \quad \frac{\partial}{\partial y}(2xy^3) = 6xy^2 = \frac{\partial}{\partial x}(3x^2y^2 - 2).$$

Wir haben oben gesehen, dass man die Lösung einer exakten Differentialgleichung gleich hinschreiben kann — jedenfalls in impliziter Form — wenn man eine Stammfunktion kennt. Es läuft also die Lösung auf die Beantwortung folgender Frage hinaus:

Wie bestimmt man eine Stammfunktion?

Wir geben zuerst ein allgemeines Resultat an (**Bestimmung von Stammfunktionen**) und überlegen uns dann, wie man üblicherweise vorgeht.

Satz 3.3.3.3. Es sei

$$f(x, y) + g(x, y)y' = 0 \quad (f, g : G \rightarrow \mathbb{R}, \text{ stetig})$$

eine exakte Differentialgleichung. Dann ist eine Stammfunktion F gegeben durch

$$F(x, y) = \int_{(a,b)}^{(x,y)} f(x, y)dx + g(x, y)dy \quad \text{für } (x, y) \in G.$$

Dabei ist $(a, b) \in G$ fest und das Integral ist das Kurvenintegral über einen beliebigen, in G gelegenen Integrationsweg von (a, b) nach (x, y) . Die weiteren Stammfunktionen sind genau die Funktionen, die sich von F um eine additive Konstante unterscheiden.

3 Differentialgleichungen (DGLen)

Ist $(a, b) \in G$ so, dass für jeden Punkt $(x, y) \in G$ der Streckenzug mit den Ecken

$$(a, b), (x, b), (x, y)$$

in G liegt, dann geht das Kurvenintegral in gewöhnliche Integrale über und es folgt

$$F(x, y) = \int_a^x f(s, b) ds + \int_b^y g(x, t) dt \text{ für } (x, y) \in G.$$

Beispiel. Wir führen unser obiges Beispiel fort.

$$2xy^3 + (3x^2y^2 - 2)y' = 0, \quad y(0) = 1.$$

Stammfunktion:

$$F(x, y) = \int_0^x (2s1^3) ds + \int_1^y (3x^2t^2 - 2) dt = 3x^2 \frac{y^3}{3} - 2y + 2 = x^2y^3 - 2y + 2.$$

Implizite Darstellung der Lösung:

$$x^2y^3 - 2y + 2 = 0.$$

Übungsaufgabe 222. (T) Entscheiden Sie ob eine exakte Differentialgleichung vorliegt und lösen Sie das Anfangswertproblem

1. $\cos x \cos y - (\sin x \sin y + y^2) y' = 0$ und $y(0) = 1$.
2. $\frac{2x}{y^3} + \frac{y^2 - 3x^2}{y^4} y' = 0$ und $y(1) = 1$.

Übungsaufgabe 223. (T) Bestimmen Sie die allgemeine Lösung der folgenden Differentialgleichungen:

1. $(\sin x \cos y + \tan^2 x) + (\cos x \sin y) y' = 0$
2. $(6xy^3 - y) + (12x^2y^2 - 2x + y) y' = 0$
3. $(2axy - y^2) dx + (2ax^2 - 3xy + 1) y' = 0$

3.3.4 Integrierende Faktoren

Inhalt in Kurzfassung: Von einer beliebig gegebenen DGL kann man nicht erwarten, dass sie exakt ist und somit auf die in 3.3.3 beschriebene Weise gelöst werden kann. Eventuell lassen sich aber sogenannte integrierende Faktoren finden, mit denen multipliziert eine nicht exakte DGL zu einer exakten wird. Allerdings führt die Suche nach integrierenden Faktoren in der Regel selbst zu neuen DGLen, die nur in gewissen Fällen leicht zu lösen sind.

Kann man eine Differentialgleichung exakt machen?

Sehr viele Differentialgleichungen der Form

$$(1) \quad f(x, y) + g(x, y)y' = 0$$

lassen sich mit einem passenden Faktor

$$m = m(x, y) : G \rightarrow \mathbb{R}, \quad m(x, y) \neq 0 \text{ für } (x, y) \in G$$

multiplizieren, so dass die lösungsgleiche Differentialgleichung

$$(2) \quad m(x, y)f(x, y) + m(x, y)g(x, y)y' = 0$$

exakt ist. m nennt man gerne einen **integrierenden Faktor** oder **Eulerschen Multiplikator** von (1). Sind die Funktionen m, f, g stetig differenzierbar, so müssen, falls (2) exakt ist, die Ausdrücke

$$\frac{\partial}{\partial y}(mf) = m_y f + m f_y \quad \text{und} \quad \frac{\partial}{\partial x}(mg) = m_x g + m g_x$$

gleich sein, d.h. es muss gelten

$$(3) \quad m_x g - m_y f = m(f_y - g_x).$$

Die Behandlung von (3) kann schwieriger sein, als die von (1). In manchen Sonderfällen kommt man aber doch leicht zum Ziel. Einer davon ist in folgendem **Satz vom integrierenden Faktor** enthalten.

Satz 3.3.4.1. *Es sei G einfach zusammenhängend, $f, g : G \rightarrow \mathbb{R}$ haben stetige partielle Ableitungen und es sei $g(x, y) \neq 0$ für $(x, y) \in G$.*

$$\frac{f_y - g_x}{g} \text{ hänge nur von } x \text{ ab.}$$

I sei die Projektion von G auf die x -Achse. Dann ist jede Lösung $m : I \rightarrow \mathbb{R}$ der Differentialgleichung

$$m' = \frac{f_y - g_x}{g} m \text{ mit } m \neq 0$$

ein integrierender Faktor der Differentialgleichung

$$f(x, y) + g(x, y)y' = 0.$$

Bemerkung. Hängt

$$\frac{f_y - g_x}{f} \text{ nur von } y \text{ ab,}$$

so betrachtet man die gewöhnliche Differentialgleichung

$$m' = -\frac{f_y - g_x}{f} m$$

3 Differentialgleichungen (DGLen)

und geht ähnlich vor.

Beispiel. Man löse die folgende Differentialgleichung

$$y - x y' = 0.$$

Es ist

$$\frac{\partial}{\partial y}(y) = 1 \quad \text{und} \quad \frac{\partial}{\partial x}(-x) = -1,$$

womit

$$\frac{1 - (-1)}{-x} = \frac{2}{-x}$$

nur von x abhängt. Integrierender Faktor:

$$m' = \frac{-2}{x} m,$$

$$m = e^{-\int \frac{2dx}{x}} = e^{-2 \ln x} = \frac{1}{x^2}.$$

Exakte Differentialgleichung:

$$\frac{y}{x^2} - \frac{1}{x} y' = 0.$$

Stammfunktion:

$$F(x, y) = \int_a^x \frac{b}{x^2} dx - \int_b^y \frac{dy}{x} = -\frac{b}{x} \Big|_a^x - \frac{y}{x} \Big|_b^y = -\frac{y}{x} + \frac{b}{a}.$$

Lösung durch (a, b) :

$$-\frac{y}{x} + \frac{b}{a} = 0 \quad \text{oder} \quad y = \frac{b}{a} x.$$

Beispiel (lineare Differentialgleichungen 1. Ordnung). Für stetige $f, g : I \rightarrow \mathbb{R}$ ist

$$y' = f(x)y + g(x) \quad \text{oder} \quad (-f(x)y - g(x)) + 1y' = 0$$

nicht exakt. Wegen

$$\frac{\partial}{\partial y}(-fy - g) = -f \quad \text{und} \quad \frac{\partial}{\partial x}(1) = 0$$

hängt

$$\frac{-f(x) - 0}{1}$$

nur von x ab und man erhält als integrierenden Faktor

$$m(x) = e^{-\int_a^x f(s) ds}.$$

Multipliziert man die Differentialgleichung damit und löst sie, so erhält man dasselbe Ergebnis wie im Abschnitt über lineare DGLen beschrieben.

3.3.5 Variablensubstitution

Inhalt in Kurzfassung: So wie bei der Suche nach Stammfunktionen führen auch bei manchen DGLen Variablensubstitutionen zum Ziel. Es folgen zwei illustrative Beispiele hierzu.

So wie bei der Integralrechnung führen auch bei der Lösung von DGLen gelegentlich Variablensubstitutionen zum Ziel, indem komplizierte auf einfachere Probleme zurückgeführt werden. In der Zeit vor dem Computer gehörte das zu den wenigen vorhandenen Optionen. Heutzutage spielen jedoch numerische Verfahren, wie wir sie in 3.3.6 noch etwas ausführlicher besprechen werden, eine viel größere Rollen. Sie machen sich die enorme Leistungsfähigkeit moderner Rechenmaschinen zunutze. Im Vergleich dazu wird Trickreichtum bei geschickten Substitutionen zur recht brotlosen Kunst, die nur mehr in recht speziellen Einzelfällen Erfolg verspricht. Immerhin zwei Beispiele dieser Art sollen hier trotzdem kurz dargestellt werden.

Rückführung homogener DGLen auf solche mit Trennung der Variablen:

(Achtung: Hier bedeutet „homogen“ etwas ganz anderes als bei linearen DGLen!) Lässt sich in einer expliziten DGL $y' = f(t, y)$ die Funktion f auch als Funktion g anschreiben, die nur von einer Variablen $u = \frac{y}{t}$ abhängt, also $f(t, y) = g(u) = g(\frac{y}{t})$, so nennt man die DGL **homogen**. In diesem Fall gilt $y = ut$ und daher mittels Produktregel (wir verwenden die Differential Schreibweise aus 3.3.2)

$$g(u) = f(t, y) = \frac{dy}{dt} = \frac{d(ut)}{dt} = u \frac{dt}{dt} + \frac{du}{dt} t = u + t \frac{du}{dt} = u + tu'.$$

Sammeln wir links die Ausdrücke in t , rechts die in u , so erhalten wir

$$tu'(t) = g(u) - u,$$

also eine DGL in der unbekanntem Funktion u , die mit der Methode der Trennung der Variablen gelöst werden kann (siehe 3.3.2).

Bernoullische DGL: Darunter versteht man sehr spezielle explizite, aber im Allgemeinen weder lineare noch autonome DGLen der Bauart

$$y' = p(t)y + q(t)y^n,$$

mit irgendwelchen Koeffizientenfunktionen p und q und $n \in \mathbb{Z}$. Für $n = 0, 1$ läge eine lineare DGL vor, die mit den entsprechenden Methoden behandelt werden kann. Es zeigt sich, dass man aber auch den Fall $n \geq 2$ auf den linearen zurückführen kann, und zwar mit Hilfe der Variablensubstitution $u = y^{1-n}$. Denn dann ergibt sich mit der Kettenregel

$$u' = (1-n)y^{-n}y' = (1-n)y^{-n}(p(t)y + q(t)y^n) = (1-n)p(t)y^{1-n} + q(t) = (1-n)p(t)u + q(t),$$

tatsächlich eine (inhomogene) in u lineare DGL mit variablen Koeffizienten.

3 Differentialgleichungen (DGLen)

Übungsaufgabe 224. (T) Lösen Sie die Differentialgleichungen indem Sie zuerst die angegebene Substitution durchführen. Vergessen Sie nicht, die Lösung am Ende wieder zurück zu transformieren.

1. $(x + y)y' = (x + y)^2 + 1$ mit Substitution $z(x) = x + y$.

2. $y \cdot \left(\cos\left(\frac{y}{x}\right) + \frac{y}{x} \sin\left(\frac{y}{x}\right) \right) = \left(y \sin\left(\frac{y}{x}\right) - x \cos\left(\frac{y}{x}\right) \right) y'$ mit Substitution $z(x) = \frac{y}{x}$

3. $2x^3 + 3xy^2 + y^3y' = 0$ mit Substitution $z(x) = \frac{y}{x}$

4. $y' + xy + \frac{x^3}{2}y^3 = 0$ mit Substitution $y = z^{-1/2}$.

Übungsaufgabe 225. (T) Lösen Sie das Anfangswertproblem mit Hilfe der Substitution $u(x) = x^\alpha y(x)$. Der Parameter α ist dabei jeweils passend zu bestimmen.

1. $3x^2y^2 - 2y - xy' = 0, y(1) = 1$ und

2. $xyy' = 2x^2 + 3xy + 2y^2, y(1) = 1$.

Übungsaufgabe 226. (T) Lösen Sie die Differentialgleichung $y' = \frac{x + 2y + 3}{2x - y - 4}$, indem Sie $X(x) = x - 1$ und $Y(y) = y + 2$ substituieren.

3.3.6 Numerische Methoden

Inhalt in Kurzfassung: Numerische Methoden führen zwar nicht zu expliziten Formeln für die Lösungen von DGLen. Sie sind aber – vor allem Dank der explodierenden Leistungsfähigkeit moderner Computer – sehr universell einsetzbar und deshalb zum wichtigsten Werkzeug für die anwendungsorientierte Lösung von DGLen geworden.

Die bisherigen Lösungsmethoden und erst recht die für lineare DGLen, wie wir sie in ?? und ?? kennen gelernt haben, sind auf ganz bestimmte Typen von DGLen ausgerichtet, bieten nur qualitative Ansätze oder sind aus einem anderen Grund für viele praktische Anwendungen nicht befriedigend. Bis weit ins 20. Jahrhundert hinein beschränkten diese Schwierigkeiten die technischen Möglichkeiten in vielen Bereichen beträchtlich. Während der letzten Jahrzehnte jedoch eröffnete die Entwicklung des Computers völlig neue Möglichkeiten. Sowohl die Kapazität der Speicher als auch die Geschwindigkeit der Prozessoren haben sich in astronomischen Größenordnungen weiterentwickelt. Kleine Computer, wie sie heutzutage schon fast jeder Schüler besitzt, können innerhalb weniger Sekunden Rechnungen durchführen, die zur Zeit der ersten Mondlandung im Jahr 1969 vermutlich überhaupt nicht möglich gewesen wären, selbst wenn man sämtliche Rechenmaschinen der NASA zusammengespannt hätte. Und nur 30 Jahre früher gab es noch überhaupt keine Computer.

Das näherungsweise (numerische) Lösen von DGLen ist heute einer der wichtigsten Anwendungsbereiche von automatischen Rechenmaschinen. Das liegt daran, dass man

3.3 Lösungsverfahren für ausgewählte nichtlineare DGLen

nicht mehr darauf angewiesen ist, durch geschickte symbolische Manipulation explizite Formeln zu finden. In den meisten etwas komplizierter gelagerten Fällen wäre man damit auch auf verlorenem Posten. Wie schon in Mathematik 1 bemerkt, lässt sich beweisen, dass es keine explizite Darstellung von Stammfunktion der relativ einfach gebauten Funktionen e^{-x^2} und $\frac{\sin x}{x}$ gibt. Damit ist klar, dass man auch bei den allermeisten etwas komplizierteren DGLen keine einfach gebauten Lösungsformeln erwarten darf. Mindestens ebenso wie bei der Integration spielen daher in praktischen Anwendungen auch bei der Lösung von DGLen sogenannte **numerische Methoden** eine immer größere Rolle. Dabei werden ausgereifte und teilweise auch hoch spezialisierte Programmpakete verwendet, die dem Anwender die Rechenarbeit abnehmen. Nicht abgenommen wird ihm hingegen die Entscheidung, zu welchen der unzähligen verfügbaren Hilfsmitteln er greift. Die wichtigste Entscheidungsgrundlage ist ein grundsätzliches Verständnis für die zugrundeliegenden Methoden und eine realistische Einschätzung ihrer Stärken und Schwächen. Da es sich in aller Regel nicht um exakte, sondern um Näherungsverfahren handelt, müssen bei ihrem Einsatz verschiedene Arten von Qualität und Kosten (z.B. für Rechenzeit, Speicheraufwand, Genauigkeit) gegeneinander abgewogen werden.

An dieser Stelle begnügen wir uns mit einer elementaren Idee zur numerischen Lösung eines Anfangswertproblems $y' = f(t, y)$, $y(t_0) = y_0$. Die gesuchte Lösung y hat an der Stelle t_0 die Ableitung $y'(t_0) = f(t_0, y_0)$. In einem hinreichend kleinen Intervall um t_0 gilt daher

$$y(t) \approx y_0 + (t - t_0)y'(t_0) = y_0 + (t - t_0)f(t_0, y_0).$$

Liegt t_1 hinreichend nahe bei t_0 , wird also $y_1 := y_0 + (t_1 - t_0)f(t_0, y_0)$ eine akzeptable Approximation für $y(t_1)$ sein. Die entsprechende Überlegung führt zu einer Approximation $y_2 := y_1 + (t_2 - t_1)f(t_1, y_1)$ für $y(t_2)$. Dieses Verfahren kann fortgeführt werden, so dass schließlich endlich viele Approximationen y_1, y_2, \dots, y_n für die Funktionswerte $y(t_1), y(t_2), \dots, y(t_n)$ der gesuchten Funktion an den Stellen t_1, t_2, \dots, t_n vorliegen. Es ist zu erwarten, dass diese Funktionswerte genauer werden, wenn die Abstände zwischen den t_i klein sind, was allerdings die Anzahl n der Schritte im Verfahren in die Höhe treibt. Auch stellen mehr Werte mehr Information dar. Allerdings kosten die zusätzlichen Schritte kostbare Rechenzeit. Überdies ist noch ein weiterer Effekt zu bedenken, die Rundungsfehler. Der Computer kann reelle Zahlen, die in der Regel ja erst durch unendlich viele Nachkommastellen exakt bestimmt sind, nur näherungsweise abspeichern. Unterhalb einer gewissen Größenordnung (etwa der Abstände $t_{i+1} - t_i$) werden die Rundungsfehler deshalb so groß, dass eine weitere Verkleinerung der Schrittlänge sich auf die Genauigkeit des Ergebnisses kontraproduktiv auswirken würde. Eine kluge Entscheidung für sinnvolle Parameter bei der Berechnung setzt also feine mathematische Überlegungen voraus, die das tägliche Brot numerischer Mathematiker sind.

Die angedeutete Methode ist für explizite (gewöhnliche) Differentialgleichungen maßgeschneidert. Für einen tatsächlichen Einsatz in der Praxis, sind unterschiedlichste Verfeinerungen im Gebrauch. Natürlich wurden Methoden zur numerischen Berechnung von Lösungen auch partieller Differentialgleichungen entwickelt. Dort müssen Näherungen nicht nur entlang der Zeitachse errechnet werden, sondern in mehrere Dimensionen, was alles entsprechend komplizierter werden lässt. Lediglich als Schlagwort sei die sogenannte

3 Differentialgleichungen (DGLen)

Methode der **finiten Elemente** erwähnt.

4 Integralrechnung in mehreren Variablen

Im Vergleich mit dem eindimensionalen Riemann- oder auch Lebesgueintegral ist die Situation im Höherdimensionalen wesentlich vielfältiger. Am unmittelbarsten als direkte Verallgemeinerung lassen sich Bereichsintegrale verstehen, wo nämlich lediglich nach dem n -dimensionalen Lebesguemaß λ_n statt nach λ integriert wird. Begrifflicher Hintergrund von Bereichsintegralen sowie Berechnungsmethoden wie der Satz von Fubini und die Substitutionsregel sind die zentralen Inhalte von Abschnitt 4.1. Kompliziertere Varianten höherdimensionaler Integrale sind Kurven- und Oberflächenintegrale, denen die Abschnitte 4.2 bzw. 4.3 gewidmet sind. Von Kurven- und Oberflächenintegralen gibt es jeweils zwei Typen: Bei jenen erster Art wird über Skalarfelder (reellwertige Funktionen) integriert, bei jenen zweiter Art über Vektorfelder. Weitreichende Verbindungen zwischen all diesen Konzepten werden durch Vektoranalysis und Integralsätze hergestellt, die als Abschluss des Kapitels in Abschnitt 4.4 behandelt werden.

4.1 Bereichsintegrale

Der nächstliegende Schritt der Verallgemeinerung eindimensionaler Integrale, die zur Berechnung von (zweidimensionalen) Flächen dienen, ist der zu Bereichsintegralen zur Berechnung von beispielsweise dreidimensionalen Volumina. Dieser Schritt ist Hauptgegenstand dieses Abschnitts. Ein angemessener Rahmen, der sich sehr gut für diverse Verallgemeinerungen eignet, fußt auf der Integration nach einem Maß, wie sie in ihren Grundzügen auch schon in Mathematik 1 behandelt wurde. Das wird in 4.1.1 nochmals rekapituliert. Mit dem Satz von Fubini (4.1.2) kann die Berechnung von höherdimensionalen Integralen auf die iterierte Berechnung eindimensionaler Integrale zurückgeführt werden. Angenehme Vereinfachungen sind in vielen anwendungsrelevanten Situationen mit Hilfe der höherdimensionalen Substitutionsregel möglich (4.1.4). Als wesentlicher Bestandteil darin tritt die Funktionaldeterminante, deren anschauliche Bedeutung noch davor in 4.1.3 behandelt wird. Darüber hinaus gehört sonst nur noch die Berechnung eines für die Stochastik wichtigen Integrals in 4.1.7 zum Prüfungsstoff. Die anderen Unterabschnitte dieses Abschnitts (4.1.5, 4.1.6, 4.1.8 und 4.1.9) sind vor allem Wiederholungen und Illustrationen des Stoffes. Sie werden zur wenigstens einmaligen Lektüre empfohlen, ergeben sich aber mehr oder weniger unmittelbar aus dem Bisherigen und sollten deshalb nicht als neuer Stoff empfunden werden.

4.1.1 Ein gemeinsamer Rahmen

Inhalt in Kurzfassung: Wir rekapitulieren Grundideen zur Integralrechnung aus Mathematik 1, insbesondere die Integration nach einem Maß. Damit wird ein gemeinsamer

4 Integralrechnung in mehreren Variablen

begrifflicher Rahmen für mannigfaltige Aufgabenstellungen verfügbar, der zum Beispiel auch in der Stochastik wertvolle Dienste leisten wird.

In Mathematik 1 haben wir den Einstieg zur Integralrechnung über das (eindimensionale) Riemannintegral gewählt. Die Aufgabe besteht darin, den Flächeninhalt rechteck-sähnlicher Flächen zu berechnen, bei denen allerdings eine der vier Seiten (wir stellen sie uns als die „obere“ vor) durch eine krummlinige Begrenzungslinie ersetzt wird. Diese Begrenzungslinie wird ihrerseits durch eine Funktion f beschrieben. In einem zweidimensionalen Koordinatensystem stellen wir uns die untere Begrenzungslinie als den Abschnitt der x -Achse zwischen zwei Punkten $a < b$ vor, die rechte und linke Begrenzungslinien als Abschnitte der senkrechten Geraden $x = a$ (links) und $x = b$ (rechts) zwischen der x -Achse (mit der Gleichung $y = 0$) und den Werten $y = f(a)$ bzw. $y = f(b)$. Ist f nichtnegativ und stetig (auch allgemeinere Voraussetzungen reichen hin), so beschreibt das Riemannintegral

$$\int_a^b f(x) dx$$

die Fläche zwischen diesen vier Linien: $y = 0$, $x = a$, $y = f(x)$ und $x = b$. Begrifflich präzise wird diese Aufgabe mit Hilfe des Riemannintegrals gelöst, die mit Hilfe von Ober- und Untersummen oder äquivalent mittels Riemannsummen erfolgt.

Ein alternativer Zugang, der ebenfalls schon in Mathematik 1 angedeutet wurde, eröffnet sich unter Verwendung des Lebesgueschen oder auch irgendeines anderen Maßes μ . Zur Erinnerung: Ähnlich wie bei der Definition von Untersummen approximiert man eine gegebene Funktion f durch Treppenfunktionen $t \leq f$, die nur endlich viele verschiedene Werte c_i annehmen, jeden auf einer *messbaren* Teilmenge A_i , $i = 1, \dots, n$, des Integrationsbereichs D . Für so ein t definiert man das Integral

$$\int_D t d\mu := \sum_{i=1}^n c_i \mu(A_i)$$

und damit für allgemeines $f \geq 0$

$$\int_D f d\mu := \sup_{t \leq f} \int_D t d\mu.$$

Für reellwertiges f mit eventuell auch negativen Werten verwendet man die Zerlegung $f = f^+ - f^-$ in Positiv- und Negativteil ($f^+ := \max\{f, 0\}$, $f^- := (-f)^+ = \max\{-f, 0\}$), für komplexwertiges f die Darstellung $f = f_1 + if_2$ mit Realteil f_1 und Imaginärteil f_2 , um das Integral in linearer Weise zu definieren:

$$\int_D f d\mu := \int_D f^+ d\mu - \int_D f^- d\mu \quad \text{bzw.} \quad \int_D f_1 + if_2 d\mu := \int_D f_1 d\mu + i \int_D f_2 d\mu.$$

Damit lässt sich eine Integrationstheorie entwickeln, die dem Riemannschen Integralbegriff in mehrfacher Hinsicht überlegen ist.¹

¹Vorausgesetzt wird dabei, dass die auftretenden Funktionen f messbar bezüglich μ sind. Und zwar

Ein erster Eindruck von dieser Überlegenheit entsteht, wenn man sich die Allgemeinheit der Begriffe vor Augen führt. Angenommen, wir wollen nicht wie beim Riemannintegral die zweidimensionale Fläche unter dem Graphen einer Funktion $f : [a, b] \rightarrow \mathbb{R}$, sondern zum Beispiel das (dreidimensionale) Volumen V eines räumlichen Bereichs B definieren, der über einer zweidimensionalen Grundfläche D liegt und nach oben durch eine Funktion $f : D \rightarrow \mathbb{R}$ mit $f \geq 0$ begrenzt ist. Dann gelingt das, weil die Beziehung

$$V = \lambda_3(B) = \int_D f \, d\lambda_2$$

zwischen dreidimensionalem Maß λ_3 und dem Integral über das zweidimensionale Maß λ_2 gilt.

Dabei ist $B = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, 0 \leq z \leq f(x, y)\}$. Wollte man statt dessen im Geiste des Riemannintegrals vorgehen, könnte man Zerlegungen von zwei- oder höherdimensionalen Bereichen samt Ober- und Untersummen betrachten. Das ist zwar möglich, wirkt im Vergleich zu der obigen konzisen Notation eher schwerfällig. So wie im Eindimensionalen stimmt der Wert des Integrals im Riemannschem Sinn (sofern definiert) aber mit dem Lebesgueschen überein, und man kann risikolos auch

$$V = \int_D f(x, y) \, dx \, dy$$

schreiben, was in vielen ingenieurwissenschaftlichen Anwendungen bevorzugt wird.

Der entscheidende mathematische Satz, mit dem sich die Lebesguesche Integralschreibweise rechtfertigen lässt, ist jener von der Existenz des Lebesguemaßes λ_n in jeder Dimension n , hier verwendet für $n = 2$ (Flächenmaß) und $n = 3$ (Raummaß). Wann immer es nützlich erscheint, werden wir davon ebenso Gebrauch machen wie von der Möglichkeit, Integrale bezüglich eines beliebigen Maßes zu bilden.

Damit ist begriffliche Klarheit darüber geschaffen, wie beispielsweise dreidimensionale Volumina als Integrale dargestellt werden können. Wie man sie berechnen kann, ist damit allerdings noch nicht geklärt. So wie bei den meisten in der Praxis auftretenden eindimensionalen Integralen, sind die Berechnungsmethoden für Riemann- und Lebesgueintegrale dieselben. Beide Integralbegriffe liefern nämlich zum Beispiel für (stückweise) stetige Funktionen auf hinreichend einfach gebauten Integrationsbereichen (zum Beispiel mit endlich vielen Rechtecken, Kreisscheiben u.ä. als Bestandteilen) immer denselben Wert.

Darüber hinaus wird sich der Zugang zur Integralrechnung bezüglich eines Maßes besonders in der Stochastik (siehe Kapitel 5) bewähren, für die wir dann die Maß- und Integrationstheorie einfach übernehmen können und nicht nochmals extra entwickeln müssen.

heißt eine reellwertige Funktion f messbar, wenn die Urbilder $f^{(-1)}(I)$ von allen Intervallen I unter f bezüglich μ messbar sind. Eine komplexwertige Funktion f heißt messbar, wenn sowohl Real- als auch Imaginärteil messbar sind. Außerdem verlangt man für reellwertiges f meist $\int f^+ \, d\mu < \infty$ oder $\int f^- \, d\mu < \infty$ oder beides, in welchem Fall f integrierbar heißt. Ein komplexwertiges f heißt integrierbar, wenn sowohl Real- als auch Imaginärteil integrierbar sind. Dann gilt auch $\int |f| \, d\mu < \infty$.

Es ist äußerst nützlich, sich klarzumachen, dass in der Integralrechnung **Nullmengen** bezüglich des zugrunde liegenden Maßes μ , das heißt Mengen N mit $\mu(N) = 0$, weitgehend vernachlässigt werden können, weil Integrale über Nullmengen immer den Wert 0 haben sie folglich keinen Einfluss auf den Wert eines Integrals haben:

$$\int_{A \cup N} f \, d\mu = \int_A f \, d\mu$$

Praktisch ist das oft, wenn ein Integrand f auf dem Rand ∂D des Definitionsbereichs $D \subseteq \mathbb{R}^n$ anders definiert ist als sonst. Bezüglich des n -dimensionalen Lebesguemaßes λ_n sind Ränder oft Nullmengen. In solchen Fällen gilt für alle Integranden f :

$$\int_D f \, d\lambda_n = \int_{D \setminus \partial D} f \, d\lambda_n$$

4.1.2 Der Satz von Fubini

Inhalt in Kurzfassung: Der Satz von Fubini ermöglicht die Rückführung höherdimensionaler Integrale auf die Iteration eindimensionaler. Die Grundidee kann man sich so vorstellen, dass man das Volumen eines dreidimensionalen Bereiches erhält, wenn man entlang einer Achse über die Querschnittsflächen integriert.

Die wichtigste Methode, mit der die aus dem Eindimensionalen vertrauten Integrationsmethoden (meist auf Basis des Hauptsatzes der Differential- und Integralrechnung) auf höhere Dimensionen hochgezogen werden kann, fußt auf dem **Satz von Fubini**.

Die Idee dazu lässt sich am einfachsten anhand eines dreidimensionalen Bereiches $B = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, 0 \leq z \leq f(x, y)\}$ mit einer Funktion (eines Funktionsgebietes) $f \geq 0$ über einem Definitionsbereich $D \subseteq \mathbb{R}^2$ illustrieren. Der Einfachheit halber setzen wir den Integrationsbereich als Rechteck $D = [a, b] \times [c, d]$ und f als stetig voraus. Das Volumen von B ist die Summe der Volumina dünner Scheiben, die sich ergeben, wenn man B durch Schnitte parallel zur y - z -Ebene (normal zur x -Achse) zerschneidet. Sei $Z = \{a = x_0 < x_1 < \dots < x_n = b\}$ eine Zerlegung von $[a, b]$ und S_i die i -te Scheibe, d.h. $S_i = \{(x, y, z) \in B : x_{i-1} \leq x \leq x_i\}$ für $i = 1, \dots, n$. Dass dabei die Schnittflächen $F_{x'} = \{(x, y, z) \in B : x = x'\}$ mit $x' = x_1, \dots, x_{n-1}$ doppelt gezählt werden, stört nicht, weil sie ja dreidimensionales Volumen 0 haben. Mit f ist auch die Funktion $x \mapsto \lambda_2(F_x)$ stetig. (Für einen strengen Beweis verwendet man ähnlich wie bei der Riemannintegrierbarkeit stetiger Funktionen im Eindimensionalen, dass stetige Funktionen auf kompakten Bereichen auch im Höherdimensionalen sogar gleichmäßig stetig sind.) Deshalb gilt bei hinreichend feiner Zerlegung Z von $[a, b]$, d.h. bei hinreichend kleinen Abständen $x_i - x_{i-1}$ eine Approximation $\lambda_3(S_i) \approx (x_i - x_{i-1})\lambda_2(F_i)$ mit einer beliebig vorgegebenen Güte. Bezeichnen wir für ein beliebiges mit F_x den zweidimensionalen Bereich $\{(y, z) \in \mathbb{R}^2 : (x, y, z) \in B\}$, so lässt sich genauer sagen: Die Riemannsumme

$$\sum_{i=1}^n (x_i - x_{i-1})\lambda_2(F_{x_i})$$

konvergiert, wenn die Feinheit der Zerlegung Z gegen 0 geht. Wegen $\lambda_2(F_x) = \int_c^d f(x, y) dy$ ist der Grenzwert das entsprechende Riemannintegral

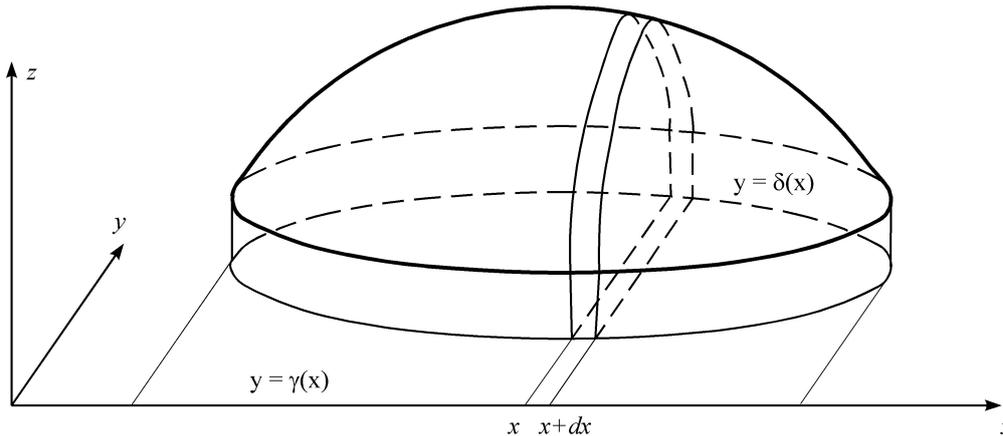
$$\lambda_3(B) = \int_a^b \lambda_2(F_x) dx = \int_a^b \int_c^d f(x, y) dy dx,$$

dessen Integrand selbst ein Riemannintegral ist, allerdings bezüglich der zweiten Integrationsvariablen y . Man kann natürlich auch symmetrisch vorgehen und zuerst das Intervall $[c, d]$ auf der y -Achse zerlegen und in Scheiben normal dazu zerlegen. Es kommt also nicht auf die Reihenfolge der Integration an. Zusammenfassend können wir schreiben:

$$\int_{[a,b] \times [c,d]} f d\lambda_2 = \int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dy dx$$

Das erste Integral ist eines im Lebesgueschen Sinn, das zweite und dritte sind Riemannsche Doppelintegrale. Die Aussage, dass höherdimensionale Integrale (unter sehr allgemeinen Voraussetzungen, die jedenfalls bei stetigem Integranden erfüllt sind) durch Iteration von eindimensionalen berechnet werden können, ist auch als **Satz von Fubini** bekannt.

Eine Veranschaulichung der oben angestellten Überlegungen mit etwas anderer Notation:



Der Satz von Fubini lässt sich auch für die Berechnung des Volumens einer Kugel

$$K_r = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}$$

mit Radius r anwenden. Die Querschnittsfläche bei einem bestimmten x -Wert besteht aus allen Punkten $(x, y, z) \in \mathbb{R}^3$ mit $x^2 + y^2 + z^2 \leq r^2$, also $y^2 + z^2 \leq r^2 - x^2$. Das ist eine Kreisscheibe mit Radius $\sqrt{r^2 - x^2}$. Diese hat nach der wohlbekannten Formel (siehe auch Mathematik 1) die Fläche $(r^2 - x^2)\pi$. Dabei variiert x zwischen $-r$ und r . Eine Stammfunktion von $f(x) := (r^2 - x^2)\pi$ ist die (ungerade) Funktion $F(x) := (r^2x - \frac{x^3}{3})\pi$. Somit ist das Volumen der Kugel gegeben durch

$$\lambda_3(K_r) = \int_{-r}^r (r^2 - x^2)\pi dx = F(r) - F(-r) = 2F(r) = 2(r^3 - \frac{r^3}{3})\pi = \frac{4}{3}r^3\pi.$$

4 Integralrechnung in mehreren Variablen

Übungsaufgabe 227. (T) Sei $f(x, y) = 1$ und D das von den drei Geraden $x = 0$, $2y + x = 2$ und $x + y = 2$ begrenzte Dreieck. Skizzieren Sie D und berechnen Sie sowohl $\iint_D f(x, y) dx dy$ als auch $\iint_D f(x, y) dy dx$ (d.h. berechnen Sie zweimal das Integral mit unterschiedlicher Integrationsreihenfolge).

Übungsaufgabe 228. (T) Vertauschen Sie die angegebene Integrationsreihenfolge und berechnen Sie

$$\int_0^4 \int_{\sqrt{y}}^2 \sin(\pi x^3) dx dy.$$

Übungsaufgabe 229. (T) Vertauschen Sie die angegebene Integrationsreihenfolge und berechnen Sie

$$\int_1^2 \int_1^{x^2} \frac{x}{y} dy dx.$$

Übungsaufgabe 230. (T) Es sind die beiden folgenden Integrale allgemein mit vertauschter Integrationsreihenfolge darzustellen:

$$1. \int_0^1 \int_y^{\sqrt{y}} [f(x, y) + g(x)] dx dy$$

$$2. \int_1^2 \int_x^{2x} [f(x, y) + g(y)] dy dx$$

Übungsaufgabe 231. (T) Vertauschen Sie die Integrationsreihenfolge und berechnen Sie

$$1. \int_0^4 \int_{\sqrt{y}}^2 \sin(\pi x^3) dx dy$$

$$2. \int_1^2 \int_1^{x^2} \frac{x}{y} dy dx$$

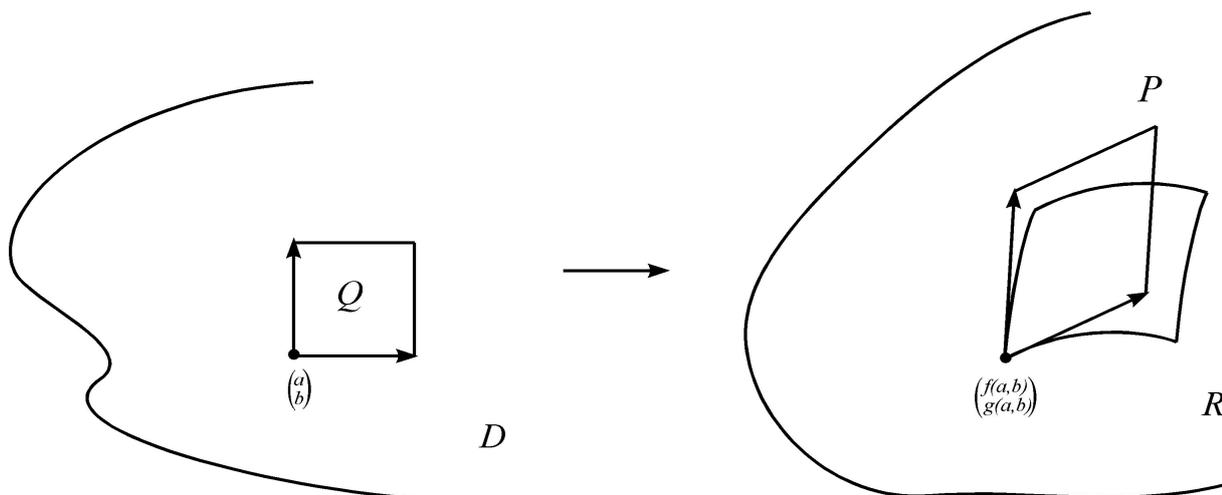
Übungsaufgabe 232. (E) Bestätigen Sie durch Nachrechnen, dass gilt

$$-\frac{\pi}{4} = \int_{y=0}^1 \int_{x=0}^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx dy \neq \int_{x=0}^1 \int_{y=0}^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy dx = \frac{\pi}{4}.$$

Verwenden Sie dabei die Identitäten

$$\frac{x^2 - y^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x} \left(\frac{-x}{x^2 + y^2} \right) = \frac{\partial}{\partial y} \left(\frac{y}{x^2 + y^2} \right).$$

Widerspricht dieses Ergebnis dem Satz von Fubini?



4.1.3 Funktionaldeterminante

Inhalt in Kurzfassung: Zur Einstimmung auf die Substitutionsregel folgt eine Veranschaulichung der Funktionaldeterminante im Zweidimensionalen.

Funktionaldeterminante

Es sei $D \subset \mathbb{R}^2$ offen und $f, g : D \rightarrow \mathbb{R}$ seien stetig differenzierbar. Wir fassen die beiden Funktionen zusammen zu einer Abbildung

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} : D \rightarrow \mathbb{R}^2.$$

Ist Q in D ein kleines Quadrat mit den Ecken

$$\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a+h \\ b \end{pmatrix}, \begin{pmatrix} a \\ b+h \end{pmatrix}, \begin{pmatrix} a+h \\ b+h \end{pmatrix},$$

dann ist der Bildbereich R von Q ein krummlinig berandeter, parallelogrammähnlicher Bereich mit den Ecken

$$\begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix}, \begin{pmatrix} f(a+h, b) \\ g(a+h, b) \end{pmatrix}, \begin{pmatrix} f(a, b+h) \\ g(a, b+h) \end{pmatrix}, \begin{pmatrix} f(a+h, b+h) \\ g(a+h, b+h) \end{pmatrix}.$$

Da f, g als Funktionen in zwei Variablen differenzierbar sind, sind die Ecken annähernd

$$\begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix}, \begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix} + h \cdot \begin{pmatrix} f_x(a, b) \\ g_x(a, b) \end{pmatrix}, \begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix} + h \cdot \begin{pmatrix} f_y(a, b) \\ g_y(a, b) \end{pmatrix}, \\ \begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix} + h \begin{pmatrix} f_x(a, b) \\ g_x(a, b) \end{pmatrix} + h \begin{pmatrix} f_y(a, b) \\ g_y(a, b) \end{pmatrix}.$$

4 Integralrechnung in mehreren Variablen

Es sei P das Parallelogramm mit diesen Ecken.
Dann gilt für die Flächeninhalte

$$I(Q) = h^2, \\ I(R) \sim I(P) = \left| \det \left(h \begin{pmatrix} f_x(a, b) \\ g_x(a, b) \end{pmatrix}, h \begin{pmatrix} f_y(a, b) \\ g_y(a, b) \end{pmatrix} \right) \right| = \left| h^2 \begin{vmatrix} f_x(a, b) & f_y(a, b) \\ g_x(a, b) & g_y(a, b) \end{vmatrix} \right|.$$

Die hier auftretende Determinante nennt man die **Funktionaldeterminante** von f, g an der Stelle (a, b) , Symbole

$$\frac{\partial(f, g)}{\partial(x, y)}(a, b), \quad \frac{\partial(f, g)}{\partial(x, y)}.$$

Es gilt also annähernd

$$\frac{I(R)}{I(Q)} \sim \left| \frac{\partial(f, g)}{\partial(x, y)}(a, b) \right| \text{ für } |h| \text{ klein.}$$

Das drückt man gerne auch so aus:

Geometrische Interpretation der Funktionaldeterminante

Die Funktionaldeterminante

$$\frac{\partial(f, g)}{\partial(x, y)}$$

gibt im zweidimensionalen Fall die lokale Flächenvergrößerung (oder -verkleinerung) bei folgender Abbildung an:

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} : D \rightarrow \mathbb{R}^2.$$

Analog beschreibt die Funktionaldeterminante im dreidimensionalen Fall den Faktor der Volumsvergrößerung.

4.1.4 Die Substitutionsregel

Inhalt in Kurzfassung: Transformiert man einen Bereich B (in welcher Dimension n auch immer) mittels einer Transformation T zu einem Bereich $T(B)$, so verändert sich im Allgemeinen das Volumen. Handelt es sich bei T um eine lineare Transformation, so multipliziert sich das Volumen mit einem Faktor, den wir in der Linearen Algebra (Kapitel 1) als Determinante $|\det T|$ bestimmt haben. Bei einer allgemeineren Transformation ist dieser Faktor variabel. Wenn T stetig differenzierbar, kann man aber in jedem Punkt \mathbf{x} die Funktionaldeterminante $|\det T'_{\mathbf{x}}|$ nehmen. Die resultierende Formel nennt man Substitutionsregel. Für $n = 1$ handelt es sich im Wesentlichen um die bereits aus Mathematik 1 bekannte Regel.

Nach dem Satz von Fubini lohnt es, auch noch eine alternative Methode zu besprechen, weil sie zur wichtigsten Formel für Integrale führt, der Substitutionsregel für höherdimensionale Integrale. Wir erinnern uns an Kugelkoordinaten, mit denen dreidimensionale Objekte beschrieben werden können. Und zwar werden die Koordinaten des Punktes

(x, y, z) dargestellt als

$$\begin{aligned}x &= r \cos(\alpha) \cos(\beta) \\y &= r \sin(\alpha) \cos(\beta) \\z &= r \sin(\beta)\end{aligned}$$

mit $r \geq 0$, $0 \leq \alpha < 2\pi$ und $-\frac{\pi}{2} \leq \beta \leq \frac{\pi}{2}$. Unter diesen Einschränkungen sind r, α, β eindeutig bestimmt, sofern nur $(x, y, z) \neq \mathbf{o}$. Die Kugel K_R mit Radius R besteht genau aus jenen Punkten mit Koordinate $r \leq R$. Wir betrachten die Transformation

$$T : [0, R] \times [0, 2\pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} r \\ \alpha \\ \beta \end{pmatrix} \mapsto \begin{pmatrix} x := r \cos(\alpha) \cos(\beta) \\ y := r \sin(\alpha) \cos(\beta) \\ z := r \sin(\beta) \end{pmatrix}.$$

Sie ist auf dem Quader $Q_R := [0, R] \times [0, 2\pi] \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ definiert und ordnet dem Tripel von Kugelkoordinaten (r, α, β) die zugehörigen kartesischen Koordinaten (x, y, z) zu. Das Bild $T(Q_R)$ ist gerade die Kugel K_R .

Eine Approximation des Volumens von K_R lässt sich folgendermaßen erhalten. Wir zerlegen den Definitionsbereich Q_R von T in kleine Teilquader Q_i . T ist differenzierbar und kann deshalb auf hinreichend kleinen Teilquadern um einen Punkt \mathbf{v}_i beliebig gut durch die lineare Funktion $T'_{\mathbf{v}_i}$ approximiert werden. Als lineare Funktion verzerrt $T'_{\mathbf{v}_i}$ das Volumen um den Faktor $|\det(T'_{\mathbf{v}_i})|$, dem Betrag der aus ?? bekannten Funktionaldeterminante von T im Punkt \mathbf{v}_i . Somit ergibt sich für das Volumen des Bildbereichs

$$\lambda_3(T(Q)) \approx \sum_i |\det(T'_{\mathbf{v}_i})| \lambda_3(Q_i).$$

Durch Verfeinerung der Zerlegung von Q wird aus \approx im Grenzwert $=$ und aus der Summe das Lebesgueintegral

$$\lambda_3(T(Q)) = \int_Q |\det(T')| d\lambda_3.$$

Dieselben Überlegungen gelten, wenn zusätzlich noch über eine auf $T(Q)$ definierte reell- oder auch komplexwertige Funktion f integriert wird. Auf diese Weise erhält man die allgemeine **Substitutionsregel** in mehreren Variablen:

$$\int_{T(Q)} f(\mathbf{y}) d\lambda_n(\mathbf{y}) = \int_Q f(T(\mathbf{x})) |\det(T'_{\mathbf{x}})| d\lambda_n(\mathbf{x})$$

Voraussetzungen an T sind dabei: $T : Q \rightarrow T(Q)$ ist fast injektiv (d.h. injektiv auf einer Menge $Q \setminus N$ mit einer λ_n -Nullmenge N) und stetig differenzierbar. Außerdem ist für Q , $T(Q)$ und f Messbarkeit vorauszusetzen, was in der Praxis aber so gut wie immer erfüllt ist.

Bemerkung zum Fall $n = 1$: Für $n = 1$ stimmt die Determinante der Ableitung mit dem Wert der Ableitung selbst überein. Deshalb geht die Substitutionsregel in diesem Fall fast in jene für das gewöhnliche Riemannintegral aus Mathematik 1 über. Und zwar deshalb nur *fast*, weil dort kein Betrag der Ableitung auf für $n = 1$ aufgetreten ist. Das hat mit der Orientierung des Riemannintegrals (Stammfunktion ausgewertet am

4 Integralrechnung in mehreren Variablen

Endpunkt minus ausgewertet am Anfangspunkt, weshalb sich bei Vertauschung der Integrationszeichen das Vorzeichen umdreht) zusammen, die im höherdimensionalen nicht in so einfacher Weise sinnvoll ist. In diesem Punkt unterscheiden sich das vorzeichenbehaftete Riemannintegral von Lebesgueintegral, wo positive Funktionen nie negative Integrale haben können.

Für unser konkretes (und sehr wichtiges) Beispiel mit den Kugelkoordinaten wollen wir die Funktionaldeterminante berechnen. Sie ergibt sich aus

$$\begin{aligned} \det T'_{(r,\alpha,\beta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial \beta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial \beta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \alpha} & \frac{\partial z}{\partial \beta} \end{vmatrix} = \begin{vmatrix} \frac{\partial r \cos(\alpha) \cos(\beta)}{\partial r} & \frac{\partial r \cos(\alpha) \cos(\beta)}{\partial \alpha} & \frac{\partial r \cos(\alpha) \cos(\beta)}{\partial \beta} \\ \frac{\partial r \sin(\alpha) \cos(\beta)}{\partial r} & \frac{\partial r \sin(\alpha) \cos(\beta)}{\partial \alpha} & \frac{\partial r \sin(\alpha) \cos(\beta)}{\partial \beta} \\ \frac{\partial r \sin(\beta)}{\partial r} & \frac{\partial r \sin(\beta)}{\partial \alpha} & \frac{\partial r \sin(\beta)}{\partial \beta} \end{vmatrix} = \\ &= \begin{vmatrix} \cos(\alpha) \cos(\beta) & -r \sin(\alpha) \cos(\beta) & -r \cos(\alpha) \sin(\beta) \\ \sin(\alpha) \cos(\beta) & r \cos(\alpha) \cos(\beta) & -r \sin(\alpha) \sin(\beta) \\ \sin(\beta) & 0 & r \cos(\beta) \end{vmatrix} \end{aligned}$$

Nach der Regel von Sarrus und unter mehrmaliger Verwendung von $\sin^2 + \cos^2 = 1$ folgt daraus weiter

$$\begin{aligned} \det T'_{(r,\alpha,\beta)} &= r^2 \cos^2(\alpha) \cos^3(\beta) + r^2 \sin^2(\alpha) \sin^2(\beta) \cos(\beta) + \\ &\quad + r^2 \cos^2(\alpha) \sin^2(\beta) \cos(\beta) + r^2 \sin^2(\alpha) \cos^3(\beta) = \\ &= r^2 \cos^3(\beta) + r^2 \sin^2(\beta) \cos(\beta) = r^2 \cos(\beta). \end{aligned}$$

Wenn wir als Integranden die konstante Funktion $f = 1$ nehmen, erhalten wir damit für das Kugelvolumen:

$$\lambda_3(K) = \lambda_3(T(Q)) = \int_Q |r^2 \cos(\beta)| d\lambda_3(r, \alpha, \beta) = \int_0^R r^2 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(\beta) d\beta d\alpha dr.$$

Das innerste Integral hat den Wert

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(\beta) d\beta = 2 \sin\left(\frac{\pi}{2}\right) = 2,$$

also

$$\lambda_3(K) = \int_0^R r^2 \int_0^{2\pi} 2 d\alpha dr = \int_0^R r^2 4\pi dr = \frac{4}{3} R^3 \pi,$$

in Übereinstimmung mit unserer ersten Berechnung.

Übungsaufgabe 233. (T) Berechnen Sie das Doppelintegral

$$\iint_D x^3 y dx dy$$

wobei D jener Bereich der Ebene ist, der von den Geraden $y = x$ und $y = 4x$ und von den Hyperbeln $xy = 1$ und $xy = 9$ begrenzt wird.

Anleitung: Substituieren Sie $x(u, v) = \frac{u}{v}$, $y(u, v) = uv$. Überlegen Sie, wo diese Substitution wohldefiniert ist (nämlich nicht auf ganz \mathbb{R}^2 !) und skizzieren Sie den Integrationsbereich D sowohl in (x, y) -Koordinaten als auch in (u, v) -Koordinaten.

Übungsaufgabe 234. (T) Berechnen Sie die von den gegebenen Kurven eingeschlossenen Flächeninhalte mit Hilfe der angegebenen Substitutionen

1. $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ ($x = ra \cos t$, $y = rb \sin t$) mit $a, b > 0$ und $t \in [0, 2\pi]$.
2. $(x^2 + y^2)^2 = xy$ ($x = r \cos \phi$, $y = r \sin \phi$).

Übungsaufgabe 235. (T) Berechnen Sie

$$\iint_D x^2 + y^2 \, dx dy,$$

wobei D von den Hyperbeln $x^2 - y^2 = 1$, $x^2 - y^2 = 3$, $xy = 2$ und $xy = 4$ begrenzt wird. Substituieren Sie analog zu Aufgabe 233 $u(x, y) = x^2 - y^2$, $v(x, y) = 2xy$.

Übungsaufgabe 236. (T) Berechnen Sie das Volumen jenes Körpers im ersten Oktanten, der von den Koordinatenebenen und der Fläche

$$f(x, y) = (1 + x + 3y)^{-3}$$

begrenzt wird. Achtung: Hier tritt ein uneigentliches Doppelintegral auf.

4.1.5 Definition und Grundeigenschaften der Doppelintegrale

Inhalt in Kurzfassung: Es geht nun um einen alternativen Zugang zu Doppelintegralen. Im Gegensatz zu der bisher dargestellten Integration nach einem Maß ähnelt er stärker dem eindimensionalen Riemann-Integral. Da in der Vorlesung dieser Weg nicht besprochen wird, handelt es sich auch um keinen Prüfungstoff. Zur Verbesserung der Anschauung und zur Verbreiterung der Verständnisbasis wird dennoch empfohlen, diesen Unterabschnitt wenigstens einmal durchzulesen.

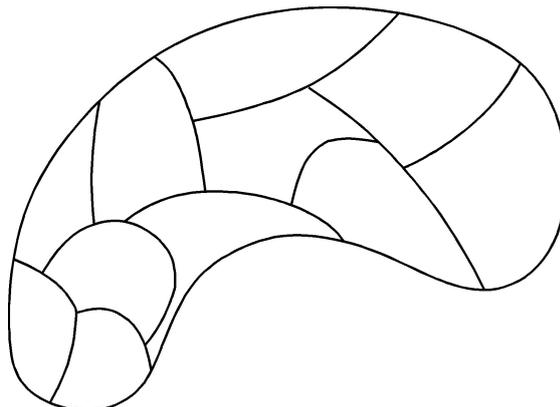
Integrationsbereiche

Unter einem **elementaren Bereich** im \mathbb{R}^2 verstehen wir eine beschränkte Teilmenge des \mathbb{R}^2 , deren Rand sich aus endlich vielen, aneinander anschließenden Kurvenstücken zusammensetzt, die sich in der Form

$$y = f(x), x \in I \text{ oder } x = g(y), y \in J$$

darstellen lassen. Dabei verlangen wir, dass diese Funktionen f, g stetig und konvex oder konkav sind. Wir merken an, dass man einem elementaren Bereich $D \subset \mathbb{R}^2$ in eindeutiger Weise einen **Inhalt** $I(D) = |D|$ zuordnen kann. Zur Inhaltsbestimmung kann man vielfach die Leibnizsche Sektorformel heranziehen.

Unter dem **Durchmesser** einer Teilmenge des \mathbb{R}^2 versteht man das Supremum der Abstände von je zwei ihrer Punkte. Man sagt, die elementaren Bereiche D_1, \dots, D_n bilden eine **Zerlegung** oder **Einteilung** des elementaren Bereichs D , wenn gilt $D = D_1 \cup \dots \cup D_n$ und je zwei verschiedene der Bereiche D_1, \dots, D_n höchstens Randpunkte gemeinsam haben. Die **Feinheit** der Zerlegung D_1, \dots, D_n von D ist das Maximum der Durchmesser der einzelnen Bereiche D_i .



Von nun an sei D ein elementarer Bereich im \mathbb{R}^2 , ebenso D_1, \dots, D_n .

Definition des Doppelintegrals

Es sei $f : D \rightarrow \mathbb{R}$. Gibt es eine Zahl $I(f)$ mit folgender Eigenschaft: zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für jede Zerlegung D_1, \dots, D_n von D mit Feinheit kleiner δ für beliebige Wahl von Punkten $(\xi_1, \eta_1) \in D_1, \dots, (\xi_n, \eta_n) \in D_n$ gilt

$$\left| I(f) - \sum_{k=1}^n f(\xi_k, \eta_k) |D_k| \right| < \varepsilon,$$

dann heißt die Funktion f **(Riemann-) integrierbar** auf D und $I(f)$ ihr **(Riemann-) Integral**. Symbole für $I(f)$:

$$\iint_D f \, dx dy, \dots, \int_D f \, dx dy, \int_D f(x, y) \, dx dy$$

f heißt **Integrand**, D **Integrationsbereich** und x, y die **Integrationsvariablen**.

Eindeutigkeit und Existenz

Ohne besondere Schwierigkeiten sieht man folgenden

Satz (Eindeutigkeit des Integrals). *Existiert das Riemannintegral einer Funktion $f : D \rightarrow \mathbb{R}$, dann ist es eindeutig bestimmt.*

Sind Integrand f und Integrationsbereich hinreichend regulär, so existiert das Riemannintegral auch tatsächlich. Wir verzichten hier allerdings auf Details.

Rechenregeln

Aus der Definition des Doppelintegrals leitet man einfach die folgenden Eigenschaften (i) – (iii) her. Die Cauchy-Schwarzsche Ungleichung zeigt man genauso wie bei Vektoren. Daraus folgt dann die Dreiecksungleichung.

Satz. *Es seien $f, g : D \rightarrow \mathbb{R}$ integrierbar und $a \in \mathbb{R}$. Dann gilt*

(i) af ist integrierbar und $\iint_D af \, dx dy = a \iint_D f \, dx dy$.

(ii) $f + g$ ist integrierbar und $\iint_D (f + g) dx dy = \iint_D f dx dy + \iint_D g dx dy$.

(iii) $|f|$ ist integrierbar und $\left| \iint_D f dx dy \right| \leq \iint_D |f| dx dy$.

(iv) **Cauchy-Schwarzsche Ungleichung:** fg, f^2, g^2 sind integrierbar und

$$\langle f, g \rangle = \iint_D fg dx dy \leq \left(\iint_D f^2 dx dy \right)^{1/2} \left(\iint_D g^2 dx dy \right)^{1/2} \quad (= \|f\|_2 \|g\|_2).$$

(v) **Dreiecksungleichung:** $(f + g)^2, f^2, g^2$ sind integrierbar und

$$\|f + g\|_2 = \left(\iint_D (f + g)^2 dx dy \right)^{1/2} \leq \left(\iint_D f^2 dx dy \right)^{1/2} + \left(\iint_D g^2 dx dy \right)^{1/2} \quad (= \|f\|_2 + \|g\|_2).$$

Ist $f : D \rightarrow \mathbb{R}$ integrierbar und $f \geq 0$, dann betrachten wir den Bereich (Gebirgsstock)

$$B = \{(x, y, z) : (x, y) \in D, 0 \leq z \leq f(x, y)\}.$$

Sein **Volumen** $V(B) = |B|$ kann man dann definieren durch

$$V(B) = \iint_D f(x, y) dx dy.$$

Eine einfache geometrische Überlegung (Gebirgsstock bzw. Hochebene mit demselben Grundriss) macht das folgende Ergebnis plausibel:

Satz (Mittelwertsatz für Doppelintegrale). *Es sei D zusammenhängend und $f : D \rightarrow \mathbb{R}$ stetig. Dann gibt es eine Stelle $\mathbf{a} \in D$ mit*

$$\iint_D f(x, y) dx dy = |D| f(\mathbf{a}).$$

Ebenso leicht einleuchtend wie der vorhergehende Satz ist das folgende Resultat.

Satz. *Es sei D_1, D_2 eine Zerlegung von D . Dann ist eine Funktion $f : D \rightarrow \mathbb{R}$ genau dann integrierbar, wenn die Einschränkungen von f auf D_1 und D_2 integrierbar sind. In diesem Falle gilt*

$$\iint_D f(x, y) dx dy = \iint_{D_1} f(x, y) dx dy + \iint_{D_2} f(x, y) dx dy.$$

4.1.6 Methoden zur Berechnung und weitere Eigenschaften

Inhalt in Kurzfassung: Es folgen Wiederholungen am Beispiel von Doppelintegralen sowie Ergänzungen, die nicht Prüfungsstoff sind.

Rückführung auf Einfachintegrale

Da man gewöhnliche Riemannintegrale verhältnismäßig gut beherrscht und auch berechnen kann, ist das folgende Resultat ein effektives Hilfsmittel zur Berechnung von Doppelintegralen.

Satz. Der elementare Bereich D habe die Form

$$D = \{(x, y) : a \leq x \leq b, \gamma(x) \leq y \leq \delta(x)\} \quad (\gamma, \delta : [a, b] \rightarrow \mathbb{R} \text{ stetig}).$$

Dann gilt für jede stetige Funktion $f : D \rightarrow \mathbb{R}$

$$\iint_D f(x, y) dx dy = \int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x, y) dy \right) dx.$$

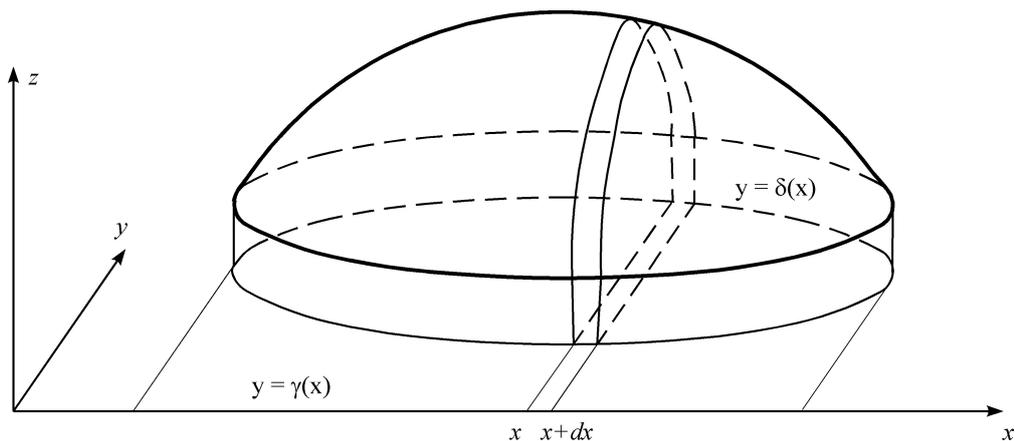
Ganz analog ergibt sich für

$$D = \{(x, y) : \alpha(y) \leq x \leq \beta(y), a \leq y \leq b\} \quad (\alpha, \beta : [a, b] \rightarrow \mathbb{R} \text{ stetig})$$

die Integrationsformel

$$\iint_D f(x, y) dx dy = \int_a^b \left(\int_{\alpha(y)}^{\beta(y)} f(x, y) dx \right) dy.$$

Diese Ergebnisse sind sehr einleuchtend: Das Volumen eines Brotlaibs ergibt sich, indem man ihn in Scheiben schneidet, den Flächeninhalt jeder Scheibe mit ihrer Dicke multipliziert und über alle Scheiben summiert.



Ein wichtiger Sonderfall unseres Satzes tritt dann auf, wenn $D = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$ ist. Dann nimmt (1) folgende Gestalt an

$$\iint_{\substack{a \leq x \leq b \\ c \leq y \leq d}} f(x, y) dx dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

Gilt darüber hinaus $f(x, y) = g(x)h(y)$, dann erhält man folgende Formel:

$$(2) \quad \iint_{\substack{a \leq x \leq b \\ c \leq y \leq d}} g(x)h(y) dx dy = \left(\int_a^b g(x) dx \right) \cdot \left(\int_c^d h(y) dy \right).$$

Beispiel

$$\begin{aligned} \iint_{\substack{-2 \leq x \leq 2 \\ 0 \leq y \leq \sqrt{4-x^2}}} x^2 y dx dy &= \int_{-2}^2 \left(\int_0^{\sqrt{4-x^2}} x^2 y dy \right) dx = \int_{-2}^2 \left(\frac{x^2 y^2}{2} \Big|_0^{\sqrt{4-x^2}} \right) dx = \\ &= \int_{-2}^2 \frac{x^2(4-x^2)}{2} dx = \frac{64}{15}. \end{aligned}$$

Transformation von Doppelintegralen

Manchmal kann man ein Doppelintegral durch „Einführung von neuen Variablen“ vereinfachen. Wie das gemacht wird ist Inhalt des folgenden Resultats.

Satz. Es seien C, D elementare Bereiche im \mathbb{R}^2 . Durch

$$(u, v) \rightarrow (x(u, v), y(u, v))$$

sei eine umkehrbar eindeutige Abbildung von C auf D gegeben, wobei x und y stetige partielle Ableitungen erster Ordnung besitzen und für die Funktionaldeterminante gilt

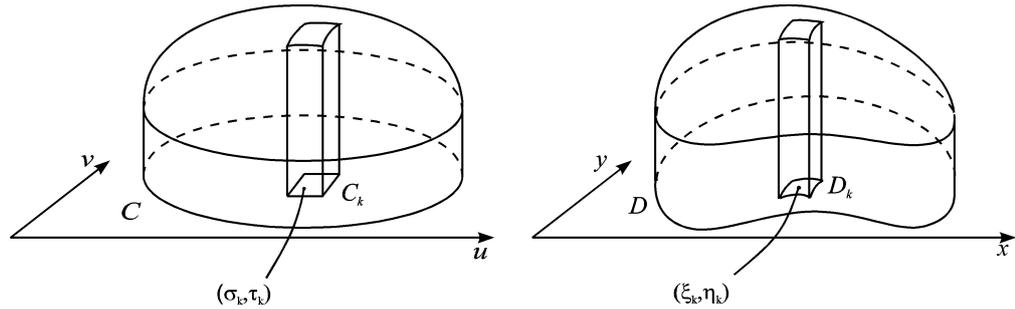
$$(3) \quad \left| \frac{\partial(x, y)}{\partial(u, v)}(u, v) \right| = \begin{vmatrix} x_u(u, v) & x_v(u, v) \\ y_u(u, v) & y_v(u, v) \end{vmatrix} \neq 0 \text{ für alle } (u, v) \in C.$$

Ist $f : D \rightarrow \mathbb{R}$ integrierbar, dann ist auch $f(x(u, v), y(u, v)) : C \rightarrow \mathbb{R}$ integrierbar und es gilt

$$(4) \quad \iint_C f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)}(u, v) \right| dudv = \iint_D f(x, y) dx dy.$$

Wir versuchen, dieses Ergebnis plausibel zu machen. Es sei D_1, \dots, D_n eine Zerlegung von D und C_1, \dots, C_n die entsprechende Zerlegung von C

4 Integralrechnung in mehreren Variablen



Nach einem Satz über Funktionaldeterminanten ist

$$\left| \frac{\partial(x, y)}{\partial(u, v)}(\sigma_k, \tau_k) \right| |C_k| \sim |D_k|$$

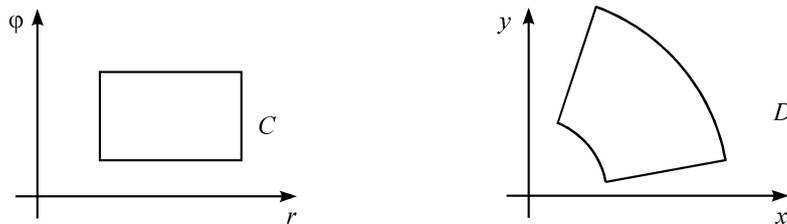
und daher

$$f(\underbrace{x(\sigma_k, \tau_k)}_{\xi_k}, \underbrace{y(\sigma_k, \tau_k)}_{\eta_k}) \left| \frac{\partial(x, y)}{\partial(u, v)}(\sigma_k, \tau_k) \right| |C_k| \sim f(\xi_k, \eta_k) |D_k|.$$

Summiert man von 1 bis n , so sieht man, dass die Zwischensummen der beiden Integrale in (4) beinahe übereinstimmen, wobei die Übereinstimmung umso besser ist, je feiner die Einteilung ist. Damit gilt (4).

Bemerkung. Wir halten fest, dass die Transformationsformel auch dann noch gilt, wenn die umkehrbare Eindeutigkeit der Abbildung oder Bedingung (3), dass die Funktionaldeterminante nicht verschwinden darf, in einzelnen Punkten oder auf Kurven verletzt ist.

Beispiel (Transformation auf Polarkoordinaten).



$$(5) \quad \iint_C f(r \cos \varphi, r \sin \varphi) r dr d\varphi = \iint_D f(x, y) dx dy.$$

Parameterintegrale

Für manche Anwendungen ist das folgende Ergebnis wichtig.

Satz (Differentiation von Parameterintegralen). *Es sei D beschränkt und abgeschlossen. $f : D \times I \rightarrow \mathbb{R}$ sei stetig und habe eine stetige partielle Ableitung nach der letzten Variablen t . Dann ist die Funktion $F : I \rightarrow \mathbb{R}$, definiert durch*

$$F(t) = \iint_D f(x, y, t) dx dy \text{ für } t \in I,$$

stetig differenzierbar und es gilt

$$F'(t) = \frac{d}{dt} \iint_D f(x, y, t) dx dy = \iint_D \frac{\partial f}{\partial t}(x, y, t) dx dy \text{ für } t \in I.$$

4.1.7 Gaußsches Fehlerintegral

Inhalt in Kurzfassung: Als originelle Anwendung bisheriger Integrationsmethoden berechnen wir das Integral $\int_{\mathbb{R}} e^{-x^2} dx$, das auch in der Stochastik noch eine wichtige Rolle spielen wird.

Für $a > 0$ sei

$$I(a) = \iint_{\substack{-a \leq x \leq a \\ -a \leq y \leq a}} e^{-(x^2+y^2)} dx dy \quad \text{und} \quad J(a) = \int_{-a}^a e^{-x^2} dx.$$

Nach (2) ist

$$(6) \quad I(a) = \int_{-a}^a e^{-x^2} dx \int_{-a}^a e^{-y^2} dy = J(a)^2.$$

Wir schätzen $I(a)$ ab. Offenbar gilt

$$(7) \quad \iint_{x^2+y^2 \leq a^2} e^{-(x^2+y^2)} dx dy \leq I(a) \leq \iint_{x^2+y^2 \leq 2a^2} e^{-(x^2+y^2)} dx dy.$$

Nun ist nach (5) und (2)

$$\begin{aligned} \iint_{x^2+y^2 \leq b^2} e^{-(x^2+y^2)} dx dy &= \iint_{\substack{0 \leq r \leq b \\ 0 \leq \varphi \leq 2\pi}} e^{-r^2} r dr d\varphi = \left(-\frac{1}{2} \int_0^b e^{-r^2} (-2r) dr \right) \int_0^{2\pi} d\varphi \\ &= \left(-\frac{1}{2} e^{-r^2} \Big|_0^b \right) 2\pi = \pi (1 - e^{-b^2}). \end{aligned}$$

Daraus und aus (6) und (7) folgt

$$\sqrt{\pi} (1 - e^{-a^2})^{1/2} \leq J(a) \leq \sqrt{\pi} (1 - e^{-2a^2})^{1/2},$$

4 Integralrechnung in mehreren Variablen

also

$$J(a) \rightarrow \sqrt{\pi} \text{ bei } a \rightarrow +\infty,$$

womit

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}.$$

4.1.8 Anwendungen

Inhalt in Kurzfassung: Die folgenden Anwendungen der mehrdimensionalen Integralrechnung sind zwar nicht Prüfungsstoff, können aber zur Ausweitung der Vorstellungsvermögens und somit zur Verbesserung des Verständnisses dienen. Entsprechend wird empfohlen, sie sich wenigstens einmal kurz zu Gemüte zu führen.

Masse:

Die Masse einer Platte D variabler (Flächen-)Dichte $\rho : D \rightarrow \mathbb{R}$ ist

$$m = \iint_D \rho(x, y) dx dy.$$

Statische Massenmomente und Schwerpunkt:

Die statischen Massenmomente oder Massenmomente erster Ordnung einer Platte D variabler Dichte ρ bezüglich der y - und der x -Achse sind

$$m_y = \iint_D x \rho(x, y) dx dy \quad \text{und} \quad m_x = \iint_D y \rho(x, y) dx dy.$$

Daraus bestimmt man sich die Schwerpunktskoordinaten

$$x_s = \frac{m_y}{m} \quad \text{und} \quad y_s = \frac{m_x}{m}.$$

Massenträgheitsmomente:

Die Massenträgheitsmomente oder Massenmomente zweiter Ordnung einer Platte D mit variabler Dichte ρ sind

$$I_{xx} = \iint_D y^2 \rho(x, y) dx dy, \quad I_{xy} = I_{yx} = \iint_D xy \rho(x, y) dx dy, \quad I_{yy} = \iint_D x^2 \rho(x, y) dx dy.$$

Erwartungswert: Diese wichtige Anwendung wird in Kapitel 5 eine zentrale Rolle spielen.

4.1.9 Dreifachintegrale

Inhalt in Kurzfassung: Zur Illustration folgen noch einige Wiederholungen und Illustrationen zu Dreifachintegralen. Die Lektüre kann zur Absicherung des Verständnisses dienen.

Wir geben nur einige Formeln ohne Details an und dann einige Beispiele.

Rückführung auf Doppel- und Einfachintegrale

Gilt

$$D = \{(x, y, z) : (x, y) \in D', \alpha(x, y) \leq z \leq \beta(x, y)\}$$

oder

$$D = \{(x, y, z) : z \in [a, b], (x, y) \in D(z)\},$$

dann gilt

$$\iiint_D f(x, y, z) dx dy dz = \iint_{D'} \left(\int_{\alpha(x, y)}^{\beta(x, y)} f(x, y, z) dz \right) dx dy \quad (\text{Säulenintegration})$$

$$\iiint_D f(x, y, z) dx dy dz = \int_a^b \left(\iint_{D(z)} f(x, y, z) dx dy \right) dz \quad (\text{Schichtenintegration})$$

Sonderfall:

$$\iiint_{\substack{a \leq x \leq b \\ c \leq y \leq d \\ e \leq z \leq k}} f(x)g(y)h(z) dx dy dz = \int_a^b f(x) dx \int_c^d g(y) dy \int_e^k h(z) dz.$$

Transformation von Dreifachintegralen

Es sei

$$\Phi : (u, v, w) \rightarrow (x(u, v, w), y(u, v, w), z(u, v, w))$$

eine umkehrbare eindeutige Abbildung von C auf D , wobei x, y, z stetige partielle Ableitungen erster Ordnung haben und

$$\det \Phi := \left| \frac{\partial(x, y, z)}{\partial(u, v, w)}(u, v, w) \right| \neq 0 \text{ für alle } (u, v, w) \in C.$$

Dann gilt

$$\iiint_C f(x(u, v, w), y(u, v, w), z(u, v, w)) \cdot \det \Phi \, du \, dv \, dw = \iiint_D f(x, y, z) dx dy dz.$$

Beispiel. Transformiert man auf Zylinder- bzw. auf Kugelkoordinaten, so sind

$$r \text{ bzw. } r^2 \sin \vartheta$$

4 Integralrechnung in mehreren Variablen

die Funktionaldeterminanten.

Beispiel. (Trägheitsmoment T einer Vollkugel K vom Radius R , dem Mittelpunkt \mathbf{o} und der Dichte ϱ bezüglich einer Achse durch \mathbf{o}). Aus Symmetriegründen ist T von der speziellen Wahl der Achse unabhängig. Daher gilt

$$\begin{aligned} T &= \varrho \iiint_K (x^2 + y^2) dx dy dz = \varrho \iiint_K (y^2 + z^2) dx dy dz \\ &= \varrho \iiint_K (z^2 + x^2) dx dy dz. \end{aligned}$$

Addition und Einführung von Kugelkoordinaten gibt

$$\begin{aligned} T &= \frac{2\varrho}{3} \iiint_K (x^2 + y^2 + z^2) dx dy dz = \frac{2\varrho}{3} \iiint_{\substack{0 \leq r \leq R \\ 0 \leq \vartheta \leq \pi \\ 0 \leq \varphi \leq 2\pi}} r^2 r^2 \sin \vartheta dr d\vartheta d\varphi \\ &= \frac{2\varrho}{3} \cdot \frac{R^5}{5} \cdot 2 \cdot 2\pi = \frac{8\varrho\pi R^5}{15}. \end{aligned}$$

4.2 Kurvenintegrale

Bei Kurvenintegralen ist der Integrand in Teilbereichen des n -dimensionalen Raum definiert, integriert wird aber nur entlang einer Kurve, die durch eine eindimensionale Parametrisierung gegeben ist. Mit deren Hilfe lassen sich Kurvenintegrale als gewöhnliche eindimensionale (Riemann-)Integrale darstellen. Die Unterscheidung zwischen Kurvenintegralen erster und zweiter Art (4.2.1 bzw. 4.2.2) entspricht den beiden Möglichkeiten für den Integranden: Skalar- oder Vektorfeld. Im zweiten Fall zeichnen sich Potentialfelder durch Wegunabhängigkeit der Kurvenintegrale aus, was ihnen einen besonderen Status verleiht, siehe 4.2.3.

4.2.1 Kurvenintegrale erster Art (über Funktionen)

Inhalt in Kurzfassung: Kurvenintegralen erster Art orientieren sich am gewöhnlichen eindimensionalen Integral. Allerdings wird nicht entlang der (linearen) x -Achse integriert (aufsummiert), sondern entlang einer Kurve im Raum. Es geht also darum, die (lokale) Kurvenlänge ins Spiel zu bringen, was über die Ableitung der Parametrisierung der Kurve gelingt.

Bereichsintegrale $\int_D f d\lambda_n$, wie wir sie in Abschnitt 4.1 behandelt haben, lassen sich folgendermaßen deuten: In einem n -dimensionalen Bereich D ist eine Substanz mit unterschiedlicher Dichte verteilt. Diese Dichte wird durch die Funktion beschrieben. Das Integral gibt die Gesamtmenge der Substanz in D an.

Eine ähnliche Idee liegt Kurvenintegralen erster Art zugrunde. Der Unterschied besteht darin, dass die Substanz (es kann sich auch um elektrische Ladung o.ä. handeln) nicht in einem Teilbereich D des \mathbb{R}^n von voller Dimension n verteilt ist, sondern auf einer (idealisiert) eindimensionalen Teilmenge, einer Kurve. Auf dieser Teilmenge gibt (analog zu Bereichsintegralen) die Funktion f an, welche Dichte $f(\mathbf{x})$ die Substanz an einem Punkt \mathbf{x} auf der Kurve hat. Im Gegensatz zu Bereichsintegralen, wo die Dichte zu verstehen ist als Grenzwert „Masse pro Raumeinheit“ für klein werdende Raumanteile, haben wir es nun mit „Masse pro Längeneinheit“ zu tun. Schnell wird klar, wie wir weiter vorzugehen haben, wenn wir uns an den Begriff der Kurvenlänge aus Mathematik 1 erinnern. Der Unterschied zum gewöhnlichen eindimensionalen Integral besteht lediglich darin, dass nun nicht entlang der x -Achse integriert wird, sondern entlang einer in der Regel gekrümmten Kurve im Raum, deren Länge berücksichtigt werden muss.

Wir betrachten also Kurven, die mit Hilfe eines reellen Intervalls $[a, b]$ parametrisiert sind:

$$\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n, \quad t \mapsto \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

Ist $Z = \{a = t_0 \leq t_1 \leq \dots \leq t_{n-1} \leq t_n = b\}$ eine hinreichend feine Zerlegung des Parameterintervalls, so wird die Länge L der Kurve (unter Voraussetzungen, die wir gleich rekapitulieren werden) approximiert durch die Summe der Abstände $\|\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})\|$, $i = 1, \dots, n$, zwischen zwei aufeinanderfolgenden Unterteilungspunkten der Kurve.

Wäre die Parameterdarstellung \mathbf{x} auf jedem der Teilintervalle linear, so wäre

$$\|\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})\| = (t_i - t_{i-1})\|\mathbf{x}'(\tau_i)\|$$

für alle τ_i mit $t_{i-1} < \tau_i < t_i$, und wir hätten

$$L = \sum_{i=1}^n (t_i - t_{i-1})\|\mathbf{x}'(\tau_i)\|.$$

Das ist gerade die Riemann-Summe für das Integral

$$\int_a^b \|\mathbf{x}'(t)\| dt$$

bezüglich der Belegung $\{\tau_1, \dots, \tau_n\}$ für die Belegung Z . Ist \mathbf{x}' stetig auf $[a, b]$ (die Parametrisierung \mathbf{x} der Kurve also stetig differenzierbar), so konvergieren (Mathematik 1) die Riemannsummen gegen den Wert dieses Integrals, sofern die Feinheit von Z gegen 0 geht. Das Integral ist deshalb auch gleich dem Supremum der Summen

$$\sum_{i=1}^n \|\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})\|$$

4 Integralrechnung in mehreren Variablen

über sämtliche Zerlegungen Z von $[a, b]$, das man üblicherweise zur Definition der Länge einer Kurve heranzieht. Für diese Definition genügt es, von der Parametrisierung nur Stetigkeit vorauszusetzen.

Doch zurück zum Kurvenintegral erster Art: Ist die Dichte f der Substanz auf der Kurve konstant $= 1$, so ist die Gesamtmenge schlicht die Länge der Kurve. Doch hindert uns nichts diesen konstanten Faktor 1 durch den Integranden f zu ersetzen:

Definition 4.2.1.1. Sei $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve im \mathbb{R}^n mit Bildbereich $C = \mathbf{x}([a, b])$ und $f : C \rightarrow \mathbb{R}$ eine (als Masseverteilung auf C interpretiert) stetige reellwertige Funktion.² Dann heißt

$$\int_C f(\mathbf{x}) \, d\mathbf{x} := \int_a^b f(\mathbf{x}(t)) \|\mathbf{x}'(t)\| \, dt$$

das Kurvenintegral erster Art von f über die Kurve \mathbf{x} (bzw. über C bezüglich der Parametrisierung \mathbf{x}).

In der ursprünglichen Aufgabenstellung hing das gesuchte Ergebnis nur von der geometrischen Lage der Kurve und der darauf durch f gegebenen Masseverteilung ab, in Definition 4.2.1.1 jedoch von der speziellen Parametrisierung. Tatsächlich: Handelt es sich bei der Kurve zum Beispiel um einen Kreis, so könnte dieser laut Definition 4.2.1.1 z.B. auch zweimal durchlaufen werden. Das würde natürlich auch den Wert des Kurvenintegrals verdoppeln. Um sich zu vergewissern, dass das Kurvenintegral tatsächlich die jeweils intendierte Bedeutung hat, ist also beispielsweise darauf zu achten, dass die Kurve nicht mehrmals durchlaufen wird (eventuell mit der Ausnahme einzelner Anfangs-, End- oder Überschneidungspunkte vom Maß 0). Ansonsten liefern verschiedene Parametrisierungen aber dasselbe Ergebnis. Der allgemeine Satz dazu lautet:

Satz 4.2.1.2. Seien $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$ und $\mathbf{y} : [c, d] \rightarrow \mathbb{R}^n$ zwei stetig differenzierbare Parametrisierungen derselben Kurve in folgendem Sinn: Es gibt eine bijektive und stetig differenzierbare Funktion (Parametertransformation) $\varphi : [a, b] \rightarrow [c, d]$ mit $\mathbf{y} \circ \varphi = \mathbf{x}$, also $\mathbf{x}(s) = \mathbf{y}(\varphi(s))$ für alle $s \in [a, b]$. Dann gilt

$$\int_a^b f(\mathbf{x}(s)) \|\mathbf{x}'(s)\| \, ds = \int_c^d f(\mathbf{x}(t)) \|\mathbf{y}'(t)\| \, dt$$

Übungsaufgabe 237. (E) Man beweise Satz 4.2.1.2. Hinweis: Kettenregel fürs Differenzieren und Substitutionsregel fürs Integrieren.

4.2.2 Kurvenintegrale zweiter Art (über Vektorfelder)

Inhalt in Kurzfassung: Im Vergleich zu Kurvenintegralen erster Art wird der dort skalarwertige Integrand f durch ein Vektorfeld \mathbf{v} ersetzt und entsprechend das Produkt

²Stetigkeit von f ist hier nicht einmal nötig. Es genügt z.B., dass $t \mapsto f(t)\|\mathbf{x}'(t)\|$ auf $[a, b]$ integrierbar ist.

$f(\mathbf{x}(t)) \cdot \|\mathbf{x}'(t)\|$ von Zahlen durch das Skalarprodukt $\mathbf{v}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$ von Vektoren. Die übliche physikalische Interpretation: Das Vektorfeld \mathbf{v} ist ein Kraftfeld, $\mathbf{x}'(t)$ die Momentangeschwindigkeit entlang der Kurve und das Integral über das Skalarprodukt die gesamte Arbeit, die zur Überwindung des Feldes \mathbf{v} entlang der Kurve geleistet werden muss.

Man stelle sich ein Kraftfeld, zum Beispiel ein Gravitationsfeld vor. Wir interessieren uns für die physikalische Arbeit A (Energie), die notwendig ist, um eine Einheitsmasse entlang eines bestimmten Weges durch dieses Gravitationsfeld zu bewegen. Wäre das Kraftfeld (repräsentiert durch den Vektor \mathbf{v}) konstant und der Weg geradlinig (mit Richtungsvektor \mathbf{x}), so wäre die einfache Formel *Arbeit = Kraft mal Weg* anzuwenden. Die pro Streckeneinheit zu überwindende Kraft ergäbe sich als die Projektion des Kraftvektors \mathbf{v} auf \mathbf{x} , also als Skalarprodukt $\mathbf{v}\mathbf{x}$. Ist das Kraftfeld hingegen nicht konstant, so hängt \mathbf{v} vom Ort ab, also auch vom Weg, den wir durchlaufen. Wir fassen \mathbf{x} als zeitabhängige Größe $\mathbf{x}(t) = (x(t), y(t), z(t))$, $t \in [0, T]$ auf. In einem kleinen Zeitintervall zwischen den Zeitpunkten t_0 und t_1 bewegen wir uns von $\mathbf{x}(t_0)$ bis $\mathbf{x}(t_1)$. Es ist also der Vektor $\mathbf{x}(t_1) - \mathbf{x}(t_0)$ zu überwinden. So wie bei den Kurvenintegralen erster Art lässt sich dieser Wert unter geeigneten Voraussetzungen an die Funktion \mathbf{x} (stetige Differenzierbarkeit der Funktionen $t \mapsto (x(t), y(t), z(t))$) gut durch $(t_1 - t_0)\mathbf{x}'(t_0)$, $\mathbf{x}'(t) = (x'(t), y'(t), z'(t))$ approximieren. Hängt auch $\mathbf{v} = \mathbf{v}(\mathbf{x})$ stetig von \mathbf{x} ab, so führt Summation über kleine Zeitintervalle $[0 = t_0, t_1], [t_1, t_2], \dots, [t_{n-1}, t_n = T]$ zu

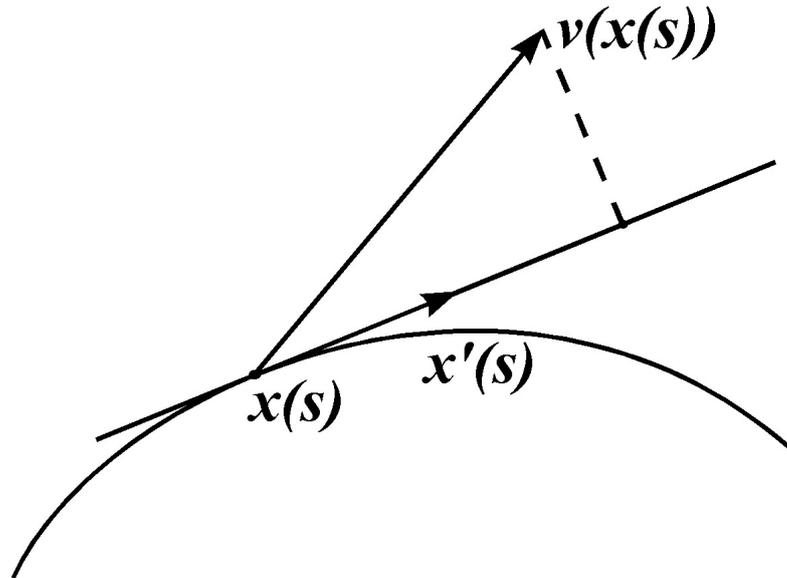
$$A \approx \sum_i \mathbf{v}(\mathbf{x}(t_i))\mathbf{x}'(t_i)(t_i - t_{i-1}).$$

So wie in 4.2.1 können wir die Summe als Riemannsumme auffassen, so dass im Grenzwert (Feinheit der Zerlegung gegen 0) gilt:

$$A = \int_0^T \mathbf{v}(\mathbf{x}(t))\mathbf{x}'(t) dt = \int_C \mathbf{v}(\mathbf{x}) d\mathbf{x}$$

In der etwas kürzeren Schreibweise des zweiten Integrals rechts bezeichnet C die Kurve (den Weg) entlang derer integriert wird. Das Integral auf der rechten Seite heißt **Kurvenintegral zweiter Art** des Vektorfeldes \mathbf{v} entlang der Kurve \mathbf{x} .

Eine graphische Darstellung zum Zeitpunkt s :



Übungsaufgabe 238. (T) Gegeben sind die Kurven $\mathbf{c}_n(t) = (t, 1-t^n, t^n)$ mit $n \in \mathbb{N}$ und $t \in [0, 1]$. Skizzieren Sie die Kurven und berechnen Sie explizit die Integrale $\int_{\mathbf{c}_n} \mathbf{v}(x) d\mathbf{x}$ für das Vektorfeld

$$\mathbf{v}(x, y, z) = \begin{pmatrix} y + z \\ x + z \\ x + y \end{pmatrix}.$$

Übungsaufgabe 239. (T) Gegeben ist die geschlossene Kurve C im \mathbb{R}^2 , welche sich aus C_1 und C_2 zusammensetzt, wobei C_1 die obere Hälfte des Kreises mit Radius 1 und Mittelpunkt $(0, 0)$ ist (d.h. jener Teil des Kreises, dessen Punkte positive y -Koordinate haben), und C_2 der Teil der x -Achse ist, der die Punkte $(-1, 0)$ und $(1, 0)$ verbindet. Berechnen Sie das Kurvenintegral $\int_C \mathbf{v}(\mathbf{x}) d\mathbf{x}$ für das Vektorfeld $\mathbf{v}(x, y) = \begin{pmatrix} x^2 y^2 \\ y \end{pmatrix}$.

Übungsaufgabe 240. (T) Vom Vektorfeld $\mathbf{V}(x, y, z) = \begin{pmatrix} 3x - 6y \\ 14yz^2 \\ 20xz^2 \end{pmatrix}$ berechne man entlang der Kurven $\gamma_1, \gamma_2, \gamma_3$ von $(0, 0, 0)$ nach $(1, 1, 1)$ die Kurvenintegrale $\int_{\gamma_n} \mathbf{V}(\mathbf{x}) d\mathbf{x}$, $n = 1, 2, 3$. Dabei ist $\gamma_1(t) = (t^2, t^2, t)$ mit $t \in [0, 1]$. γ_2 die gerade Verbindung von $(0, 0, 0)$ zu $(0, 1, 0)$, dann zu $(0, 1, 1)$ und schließlich zu $(1, 1, 1)$; γ_3 die gerade Verbindung von $(0, 0, 0)$ zu $(1, 1, 1)$.

Übungsaufgabe 241. (T) Wie Aufgabe 240 aber mit dem Vektorfeld $\mathbf{V}(x, y, z) = \begin{pmatrix} 3xy \\ -5z \\ 10x \end{pmatrix}$ und der Kurve $\gamma(t) = (t^2 + 1, 2t^2, t^3)$.

Übungsaufgabe 242. (T) Zeigen Sie, dass das Vektorfeld $\mathbf{v}(x, y) = \begin{pmatrix} 2xy+1 \\ x^2 \end{pmatrix}$ konservativ ist und berechnen Sie seine Potentialfunktion.

Berechnen Sie das Kurvenintegral $\int_C \mathbf{v} \, d\mathbf{x}$, wobei C den Kreisbogen $x^2 + y^2 = 4$, $x \geq 0$, $y \geq 0$ mit Anfangspunkt $(2, 0)$ und Endpunkt $(0, 2)$ bezeichne, direkt und mit Hilfe der Potentialfunktion.

Übungsaufgabe 243. (T) Gegeben ist das Vektorfeld $\mathbf{v}(x, y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$. Zeigen Sie, dass \mathbf{v} die Integrabilitätsbedingung erfüllt. Berechnen Sie $\int_\gamma \mathbf{v} \, d\mathbf{x}$, wobei γ der (gegen den Uhrzeigersinn durchlaufene) Kreis $x^2 + y^2 = 1$ ist. Wieso entsteht hier kein Widerspruch zur Wegunabhängigkeit von konservativen Vektorfeldern?

Übungsaufgabe 244. (T) Es ist nachzuweisen, dass durch $\mathbf{v}(x, y, z) = \begin{pmatrix} 2xy \\ x^2 + 2yz \\ y^2 + 1 \end{pmatrix}$ ein konservatives Vektorfeld gegeben ist. Es ist ein Potential zu berechnen und für eine von $(1, -2, 1)$ zu (x_0, y_0, z_0) führende Kurve C das Integral $\int_C \mathbf{v} \, d\mathbf{x}$ zu berechnen.

Übungsaufgabe 245. (T) Bestimmen Sie den Wert des Parameters λ so, dass das Vektorfeld $\mathbf{v}(x, y) = \left(\frac{\lambda}{x+y}, \frac{x}{y(x+y)} \right)$ konservativ wird. Berechnen Sie die zugehörige Potentialfunktion.

4.2.3 Wegunabhängigkeit von Kurvenintegralen in Potentialfeldern

Inhalt in Kurzfassung: Eine der bemerkenswertesten Eigenschaften von Potentialfeldern besteht in der Wegunabhängigkeit von Kurvenintegralen zweiter Art. Diese wird theoretisch wie auch anhand einiger Beispiele besprochen.

Es war eine wichtige Eigenschaft von Kurvenintegralen zweiter Art, nicht von der speziellen Parametrisierung abzuhängen. Eine noch stärkere Unabhängigkeitseigenschaften ergibt sich im Falle von Potentialfeldern. Dort hängen Kurvenintegrale nur mehr von Anfangs- und Endpunkt der Kurve ab und nicht vom Weg dazwischen. Man spricht von Wegunabhängigkeit:

Satz 4.2.3.1. Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^2$ oder \mathbb{R}^3 ein stetig differenzierbares Potentialfeld mit Potential P (also $P' = \mathbf{v}$). Ist C eine Kurve in G mit stückweise stetig differenzierbarer Parameterdarstellung $\mathbf{x} : [a, b] \rightarrow G$, dann gilt

$$(1) \quad \int_C \mathbf{v} \, d\mathbf{x} = P(\mathbf{x}(b)) - P(\mathbf{x}(a)).$$

Beweis. Wegen $\mathbf{v} = \text{grad}P$, der Kettenregel für Funktionen in mehreren Variablen und dem Hauptsatz der Differential- und Integralrechnung ist

$$\int_C \mathbf{v} \, d\mathbf{x} = \int_a^b \mathbf{v}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \, dt = \int_a^b \frac{d}{dt} P(\mathbf{x}(t)) \, dt = P(\mathbf{x}(b)) - P(\mathbf{x}(a)).$$

□

4 Integralrechnung in mehreren Variablen

Die Formel (1) sagt aus, dass ein Kurvenintegral in einem Potentialfeld nur von Anfangs- und Endpunkt, nicht aber vom tatsächlichen Verlauf der Kurve abhängt. Das drückt man gerne auch so aus: Das Kurvenintegral ist **wegunabhängig**. Speziell ist das Kurvenintegral über geschlossene Wege 0. Durch Umlauf um einen geschlossenen Weg in einem Potentialfeld läßt sich also keine Energie gewinnen. Eine weitere wichtige Folgerung des letzten Satzes lautet:

Satz (Bestimmung von Potentialen). *Weiß man, dass ein stetiges Vektorfeld $\mathbf{v} : G \rightarrow \mathbb{R}^2$ oder \mathbb{R}^3 ein Potentialfeld ist, dann erhält man ein Potential P auf folgende Weise:*

$$P(\mathbf{x}) = \int_{\mathbf{a}}^{\mathbf{x}} \mathbf{v} d\mathbf{y} \text{ für } \mathbf{x} \in G, \quad (\mathbf{a} \in G, \text{ fest})$$

wobei diese Schreibweise andeuten soll, dass das Integral ein Kurvenintegral über irgendeine Kurve in G ist, die eine stetig differenzierbare Parameterdarstellung hat und \mathbf{a} und \mathbf{x} verbindet.

Beispiel. Wir betrachten einen Punkt der Masse m in \mathbf{o} . Welche Energie E benötigt man, um eine Einheitsmasse vom Punkt $(1, 0, 0)$ in den Punkt $(1, 1, 1)$ zu bringen.

- (i) Wüsste man nicht, dass das Gravitationsfeld des Massenpunktes ein Potentialfeld ist, dann ist die Aufgabe unvollständig, man müsste noch die Verbindungskurve angeben.
(ii) Weiß man, dass unser Gravitationsfeld ein Potentialfeld ist, dann darf man irgendeine Verbindungskurve C wählen, z.B. die mit der Parameterdarstellung

$$\mathbf{x}(t) = (1-t) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ t \\ t \end{pmatrix}, \quad 0 \leq t \leq 1.$$

Dann ist

$$\begin{aligned} E &= - \int_C \gamma m \frac{-\mathbf{x}}{\|\mathbf{x}\|^3} d\mathbf{x} = \gamma m \int_0^1 \frac{\mathbf{x}(t) \cdot \dot{\mathbf{x}}(t)}{\|\mathbf{x}(t)\|^3} dt = \gamma m \int_0^1 \frac{2t}{(1+2t^2)^{3/2}} dt \\ &= \frac{\gamma m}{2} \int_0^1 \frac{4t dt}{(1+2t^2)^{3/2}} = \frac{\gamma m}{2} \int_0^2 \frac{du}{(1+u)^{3/2}} = \frac{\gamma m}{2} (1+u)^{-1/2} \cdot (-2) \Big|_0^2 \\ &= \gamma m \left(1 - \frac{1}{\sqrt{3}} \right). \end{aligned}$$

- (iii) Weiß man, dass das Gravitationsfeld die Potentialfunktion

$$P(\mathbf{x}) = \frac{\gamma m}{\|\mathbf{x}\|}$$

hat, dann ist

$$E = -(P((1, 1, 1)) - P(1, 0, 0)) = -\left(\frac{\gamma m}{\sqrt{3}} - \gamma m \right) = \gamma m \left(1 - \frac{1}{\sqrt{3}} \right).$$

Beispiel. Man berechne folgende Kurvenintegrale:

$$(i) \quad \int_C y^2 dx - x^2 dy, \quad C : \mathbf{x}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad 0 \leq t \leq \frac{\pi}{2},$$

$$\int_C = - \int_0^{\pi/2} (\sin^3 t + \cos^3 t) dt.$$

Die Differentialschreibweise im ersten Integral steht für $dx = x'(t)dt = -\sin(t)dt$ bzw. $dy = y'(t)dt = \cos(t)dt$, woraus sich der Rest ergibt. Beachtet man, dass

$$\int \sin^3 t dt = -\cos t + \frac{1}{3} \cos^3 t, \quad \int \cos^3 t dt = \sin t - \frac{1}{3} \sin^3 t$$

gilt, so folgt

$$\int_C = \cos t - \frac{1}{3} \cos^3 t - \sin t + \frac{1}{3} \sin^3 t \Big|_0^{\pi/2} = -\frac{4}{3}.$$

$$(ii) \quad \int_C y^2 dx - x^2 dy, \quad C : \mathbf{x}(t) = \begin{pmatrix} 1-t \\ t \end{pmatrix}, \quad 0 \leq t \leq 1.$$

$$\int_C = \int_0^1 (-2t^2 + 2t - 1) dt = \dots = -\frac{2}{3}.$$

Das Vektorfeld $\begin{pmatrix} y^2 \\ -x^2 \end{pmatrix}$ ist also **kein** Potentialfeld.

4.3 Oberflächenintegrale

Wenn auch nicht auf den ersten Blick sichtbar, so folgt der Abschnitt über Oberflächenintegrale einer analogen Struktur wie der über Kurvenintegrale. Diese Struktur ist geprägt von der Unterscheidung zwischen Oberflächenintegralen erster und zweiter Art. Von gleicher Art ist bei beiden Typen von Oberflächenintegralen der Integrationsbereich, nämlich eine durch zwei Parameter beschriebene Fläche im \mathbb{R}^3 . Durch die Parametrisierung wird die Integrationsaufgabe auf zweidimensionale Bereichsintegrale zurückgeführt. Die Ideen hinter dieser Zurückführung involvieren das Vektorprodukt und sind Gegenstand der einführenden Bemerkungen in 4.3.1. In 4.3.2 erfolgt dann die Anwendung auf die Integration von Skalarfeldern (Oberflächenintegrale erster Art), in 4.3.4 auf Vektorfelder (Oberflächenintegrale zweiter Art). Zu beiden Typen gibt es Unterabschnitte mit weiteren Erläuterungen und Illustrationen, die jedoch nicht zum Prüfungstoff zählen (4.3.3 und 4.3.5).

4.3.1 Vorbemerkungen zur Geometrie von Flächen

Inhalt in Kurzfassung: Als Vorbereitung auf die Oberflächenintegrale sowohl erster als auch zweiter Art beginnen wir mit einigen Bemerkungen zur Parametrisierung von (zweidimensionalen) Flächen im (dreidimensionalen) Raum. Besondere Beachtung verdient

dabei das Vektorprodukt, angewendet auf die Parametrisierung der Fläche.

Inhalt von Flächenstücken; Oberflächen

Es sei D beschränkt und abgeschlossen. Ist $f : D \rightarrow \mathbb{R}$ stetig, so nennt man

$$F = \{(x, y, z) : (x, y) \in D, z = f(x, y)\}$$

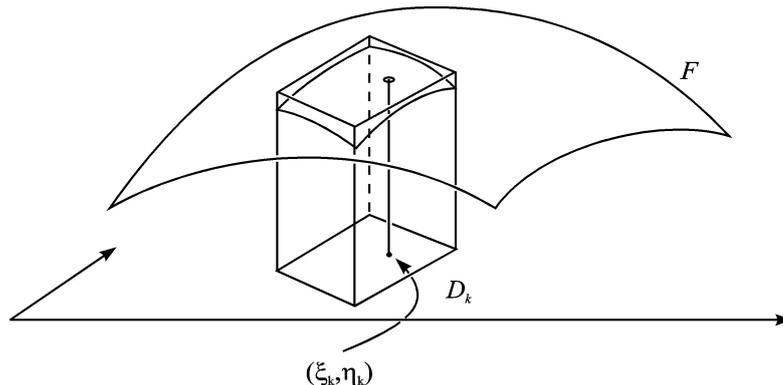
ein Flächenstück in expliziter Darstellung. Unter gewissen Einschränkungen für f kann man ihm einen Flächeninhalt zuordnen. Hat f stetige partielle Ableitungen, so hat F einen Inhalt $I(F) = |F|$, der sich in folgender Weise berechnen läßt:

$$(8) \quad I(F) = \iint_D \sqrt{1 + f_x(x, y)^2 + f_y(x, y)^2} dx dy.$$

Anschaulich kann man sich diese Formel auf folgende Weise klar machen: Wir betrachten einen Summanden einer Zwischensumme für dieses Integral:

$$\sqrt{1 + f_x(\xi_k, \eta_k)^2 + f_y(\xi_k, \eta_k)^2} |D_k|.$$

Das ist der Inhalt des in der Tangentialebene an F im Punkt $(\xi_k, \eta_k, f(\xi_k, \eta_k))$ und über D_k gelegenen Flächenstücks.



Dieses ebene Flächenstück nähert sehr gut den Teil von F an, der über D_k liegt.

Ist ein Flächenstück F in Parameterform gegeben, etwa

$$F = \{\mathbf{x}(u, v) : (u, v) \in D\} \quad \text{mit } \mathbf{x} = \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix},$$

und haben x, y, z stetige partielle Ableitungen, und ist

$$\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v) \neq \mathbf{o} \text{ auf } D, \quad \text{wobei } \mathbf{x}_u = \begin{pmatrix} x_u \\ y_u \\ z_u \end{pmatrix}, \quad \mathbf{x}_v = \begin{pmatrix} x_v \\ y_v \\ z_v \end{pmatrix}$$

(dieser Vektor ist ein Flächennormalenvektor im Punkt $\mathbf{x}(u, v)$), dann hat F einen Inhalt $I(F)$ und es gilt

$$(9) \quad I(F) = \iint_D \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| \, du \, dv.$$

(8) ist ein Sonderfall von (9).

Beispiel (Oberfläche und Volumen der Kugel).

Wir benützen die beiden folgenden Darstellungen der Oberfläche einer Kugel K mit Mittelpunkt \mathbf{o} und Radius R :

$$\mathbf{x} = \begin{pmatrix} R \sin \vartheta \cos \varphi \\ R \sin \vartheta \sin \varphi \\ R \cos \vartheta \end{pmatrix} \quad \begin{array}{l} \vartheta \text{ Poldistanz, } 0 \leq \vartheta \leq \pi, \\ \varphi \text{ geographische Länge, } 0 \leq \varphi \leq 2\pi, \end{array}$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ \sqrt{R^2 - (x^2 + y^2)} \end{pmatrix} \quad x^2 + y^2 \leq R^2.$$

Damit ist

$$O(K) = \iint_{\substack{0 \leq \vartheta \leq \pi \\ 0 \leq \varphi \leq 2\pi}} \|\mathbf{x}_\vartheta \times \mathbf{x}_\varphi\| \, d\vartheta \, d\varphi = \iint_{\substack{0 \leq \vartheta \leq \pi \\ 0 \leq \varphi \leq 2\pi}} R^2 \sin \vartheta \, d\vartheta \, d\varphi = R^2 \int_0^\pi \sin \vartheta \, d\vartheta \int_0^{2\pi} d\varphi = 4\pi R^2,$$

$$\begin{aligned} V(K) &= 2 \iint_{x^2 + y^2 \leq R^2} \sqrt{R^2 - (x^2 + y^2)} \, dx \, dy = 2 \iint_{\substack{0 \leq r \leq R \\ 0 \leq \varphi \leq 2\pi}} \sqrt{R^2 - r^2} \, r \, dr \, d\varphi = \\ &= -2\pi \int_0^R \sqrt{R^2 - r^2} (-2r) \, dr = -2\pi \left((R^2 - r^2)^{3/2} \cdot \frac{2}{3} \Big|_0^R \right) = \frac{4\pi}{3} R^3. \end{aligned}$$

4.3.2 Oberflächenintegrale erster Art (über Funktionen)

Inhalt in Kurzfassung: Bei Oberflächenintegralen erster Art wird, analog zu Kurvenintegralen erster Art, eine reellwertige Funktion (ein Skalarfeld) integriert, allerdings nicht über eine eindimensionale Kurve, sondern über eine zweidimensionale Fläche. Wieder spielen Parametrisierungen eine entscheidende Rolle, um solche Flächen überhaupt beschreiben zu können. Dadurch wird auch die Zurückführung auf zweidimensionale Bereichsintegrale möglich (so wie Kurvenintegrale auf eindimensionale Integrale zurückgeführt werden können). Dem Betrag der Ableitung der Parametrisierung einer Kurve entspricht nun der Betrag eines Vektorproduktes.

4 Integralrechnung in mehreren Variablen

Man stelle sich eine krumme Fläche F im Raum vor, auf der Masse verteilt ist. F werde durch eine Funktion

$$\mathbf{x} : D \rightarrow \mathbb{R}^3, \quad (u, v) \mapsto (x(u, v), y(u, v), z(u, v))$$

mit $D \subseteq \mathbb{R}^2$, $F := \mathbf{x}(D)$, beschrieben wird. Die Verteilung der Masse sei durch eine Dichtefunktion $f : F \rightarrow \mathbb{R}$ modelliert. Wir interessieren uns für die gesamte auf F befindliche Masse M . Auf einem kleinen Teilstück F_i von F befindet sich bei annähernd stetiger Verteilung f eine Gesamtmasse, die etwa dem Produkt $f(\mathbf{x}_i)\lambda_2(F_i)$, wobei $\mathbf{x}_i = \mathbf{x}(u_i, v_i)$ ein Punkt auf F_i ist und $\lambda_2(F_i)$ das zweidimensionale Flächenmaß von F_i . Ist \mathbf{x} stetig differenzierbar und $F_i = \mathbf{x}(Q_i)$ mit einem kleinen, achsenparallelen Quadrat $Q_i \subseteq D$, so unterscheidet sich F_i nur geringfügig von seiner linearen Approximation. Diese ist ein Parallelogramm, dessen erzeugende Vektoren die partiellen Ableitungen

$$\mathbf{x}_u(u_i, v_i) = \left(\frac{\partial x}{\partial u}(u_i, v_i), \frac{\partial y}{\partial u}(u_i, v_i), \frac{\partial z}{\partial u}(u_i, v_i) \right)$$

und

$$\mathbf{x}_v(u_i, v_i) = \left(\frac{\partial x}{\partial v}(u_i, v_i), \frac{\partial y}{\partial v}(u_i, v_i), \frac{\partial z}{\partial v}(u_i, v_i) \right)$$

von \mathbf{x} nach u bzw. v , jeweils an der Stelle (u_i, v_i) sind. Die Fläche eines Parallelogramms, das von zwei Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ aufgespannt wird, stimmt bekanntlich mit dem Betrag $\|\mathbf{a} \times \mathbf{b}\|$ des Vektorproduktes von \mathbf{a} und \mathbf{b} überein. Ähnliche Überlegungen zu Zerlegungen wie zum Kurvenintegral (nur jetzt des zweidimensionalen Bereichs D statt des eindimensionalen Zeitintervalls) und Grenzübergang für beliebig feine Zerlegungen führen wieder zu einer Darstellung als Integral. In der Lebesgueschen Notation gilt daher

$$M = \int_D f(\mathbf{x}(u, v)) \cdot \|\mathbf{x}_u \times \mathbf{x}_v(u, v)\| \, d\lambda_2(u, v),$$

aber auch folgende Schreibweisen sind üblich:

$$M = \int_F f \, d\mathbf{O} = \int \int_D f(\mathbf{x}(u, v)) \|\mathbf{x}_u \times \mathbf{x}_v(u, v)\| \, dudv$$

4.3.3 Oberflächenintegrale erster Art, Ergänzungen

Inhalt in Kurzfassung: Die folgenden Ergänzungen zu Theorie und Interpretation von Oberflächenintegralen erster Art sind nicht Prüfungsstoff, werden aber zur Lektüre empfohlen.

Es sei G ein Gebiet im \mathbb{R}^3 , D ein elementarer Bereich im \mathbb{R}^2 .

Definition

Es sei $f : G \rightarrow \mathbb{R}$. Ferner sei $\mathbf{x} : D \rightarrow G$ die Parameterdarstellung eines Flächenstücks F in G . Gibt es eine Zahl $O(f, F)$ mit folgender Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es ein

$\delta > 0$, so dass für jede Zerlegung D_1, \dots, D_n von D in elementare Bereiche mit Feinheit $< \delta$ für beliebige Wahl von Punkten $(\sigma_1, \tau_1) \in D_1, \dots, (\sigma_n, \tau_n) \in D_n$ gilt

$$\left| O(f, F) - \sum_{k=1}^n f(\mathbf{x}(\sigma_k, \tau_k)) |F_k| \right| < \varepsilon,$$

dann heißt $O(f, F)$ das **Oberflächenintegral** der Funktion f über das Flächenstück F . Dabei sind F_1, \dots, F_n die Flächenstücke mit den Parameterdarstellungen $\mathbf{x}|D_1, \dots, \mathbf{x}|D_n$. Übliche Symbole für $O(f, F)$ sind

$$\int_F f dO, \int_F f dS, \iint_F f dO, \dots$$

Es genügt also, die Funktion f auf dem Flächenstück F zu kennen.

Eindeutigkeit, Existenz und Berechnung

Ohne Schwierigkeiten sieht man, dass ein Oberflächenintegral eindeutig bestimmt ist — falls es überhaupt existiert. Es existiert erfreulicherweise in allen vernünftigen Fällen, wie das folgende Resultat zeigt.

Satz (Existenz und Rückführung auf Doppelintegrale). *Es sei $f : G \rightarrow \mathbb{R}^3$ stetig und*

$\mathbf{x} : D \rightarrow G$ Parameterdarstellung eines Flächenstücks F und \mathbf{x} habe stetige partielle Ableitungen und es sei $\mathbf{x}_u \times \mathbf{x}_v \neq \mathbf{o}$ auf D . Dann existiert das Oberflächenintegral von f über F und es gilt

$$\int_F f dO = \iint_D f(\mathbf{x}(u, v)) \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| du dv.$$

Hat F die spezielle Darstellung

$$z = g(x, y) \text{ mit } (x, y) \in D \quad (g : D \rightarrow \mathbb{R}, \text{ stetig}),$$

dann gilt

$$\int_F f dO = \iint_D f(x, y, g(x, y)) (1 + g_x(x, y)^2 + g_y(x, y)^2)^{1/2} dx dy.$$

Physikalische Deutung

Für die Definition des Oberflächenintegrals genügt es, f auf F zu kennen. Ist nun f (auf F) die Dichte einer Massenbelegung von F bezogen auf die Flächeneinheit, dann ist das Oberflächenintegral einfach die Gesamtmasse. Eine analoge Aussage gilt, wenn f die Ladungsdichte ist. Ist das Flächenstück F ein Teil der Erdoberfläche (Seehöhe) und f die Höhe (bezüglich der Seehöhe) des darüberliegenden Gebirges, so ist das Oberflächenintegral einfach das Gebirgsvolumen (über Seehöhe).

Flächen

Ohne dass wir darauf näher eingehen, halten wir fest, dass sich Flächen aus mehreren Flächenstücken zusammensetzen und dass das Oberflächenintegral über eine Fläche einfach die Summe der Oberflächenintegrale über die einzelnen Flächenstücke ist.

Beispiel (Schwerpunkt eines massenbelegten Flächenstücks). Es sei $\varrho(\mathbf{x})$ die Dichte der Massenbelegung eines Flächenstücks F und die Parameterdarstellung von F habe stetige partielle Ableitungen. Dann ist der (Massen-)Schwerpunkt von F gegeben durch

$$\frac{1}{\int_F \varrho(\mathbf{x}) dO} \begin{pmatrix} \int_F x \varrho(\mathbf{x}) dO \\ \int_F y \varrho(\mathbf{x}) dO \\ \int_F z \varrho(\mathbf{x}) dO \end{pmatrix}.$$

Beispiel. Das Gravitationspotential oder das elektrostatische Potential eines mit Masse bzw.

Ladung der Dichte $\varrho(\mathbf{x})$ belegten Flächenstücks F ist gegeben durch

$$P(\mathbf{x}) = \gamma \int_F \frac{\varrho(\mathbf{u})}{\|\mathbf{u} - \mathbf{x}\|} dO,$$

wobei γ eine Proportionalitätskonstante ist (Gravitationskonstante bzw. Coulombkonstante).

Beispiel. Man bestimme das Trägheitsmoment T einer mit Masse der Dichte 1 belegten Kugelfläche S vom Radius R bezüglich einer Achse durch den Mittelpunkt. O.B.d.A. darf man annehmen, dass \mathbf{o} Mittelpunkt und die z -Achse diese Achse ist. Dann gilt

$$T = 2 \int_F f dO \text{ mit } f(\mathbf{x}) = x^2 + y^2, \quad F : z = g(x, y) = (R^2 - (x^2 + y^2))^{1/2}.$$

Also ist

$$\begin{aligned} T &= 2 \iint_{x^2+y^2 \leq R^2} (x^2 + y^2) \left(1 + \frac{x^2 + y^2}{R^2 - (x^2 + y^2)}\right)^{1/2} dx dy = 2R \int_0^{2\pi} \int_0^R \frac{r^2}{\sqrt{R^2 - r^2}} r dr d\varphi \\ &= \frac{4\pi}{R} R R^3 \int_0^R \frac{(r/R)^3}{\sqrt{1 - (r/R)^2}} dr = 4\pi R^4 \int_0^1 \frac{s^3}{\sqrt{1 - s^2}} ds = 4\pi R^4 \int_0^{\pi/2} \frac{\sin^3 t}{\sqrt{1 - \sin^2 t}} \cos t dt \\ &= 4\pi R^4 \int_0^{\pi/2} \sin^3 t dt = \frac{8\pi R^4}{3}. \end{aligned}$$

4.3.4 Oberflächenintegrale zweiter Art (über Vektorfelder)

Inhalt in Kurzfassung: Oberflächenintegrale zweiter Art verhalten sich zu denen erster Art analog wie Kurvenintegrale zweiter zu denen erster Art, das heißt: Der Integrationsbereich ist eine zweidimensionale parametrisierte Fläche, Integrand ist aber nun ein

Vektorfeld, wobei an die Stelle des Produktes zweier Skalare das Skalarprodukt zweier Vektoren tritt: einer ist durch das Vektorfeld gegeben, der andere ist das Vektorprodukt der Ableitung der Parametrisierung nach den beiden Parametern (und nicht mehr dessen Betrag). Übliche physikalische Interpretation: Der gesamte Durchfluss des Vektorfeldes durch die Fläche.

Wieder gehen wir von einer krummen Fläche im Raum aus, die durch eine Funktion

$$\mathbf{x} : D \rightarrow \mathbb{R}^3, \quad (u, v) \mapsto (x(u, v), y(u, v), z(u, v))$$

mit $D \subseteq \mathbb{R}^2$, $F := \mathbf{x}(D)$, beschrieben wird. Im Unterschied zu Oberflächenintegralen erster Art, wo man an eine Massenverteilung auf einer Fläche denkt, empfiehlt sich bei jenen zweiter Art die Vorstellung eines Flusses im Raum. Die Aufgabe besteht darin, die Flüssigkeitsmenge zu bestimmen, die ein bestimmtes Flächenstück im Raum pro Zeiteinheit durchdringt. Wieder arbeitet man mit einer lokalen linearen Approximation (Tangentialebene) eines kleinen Teilflächenstücks $F_i := \mathbf{x}(D_i)$, wobei D_i ein kleines Teilrechteck des Parameterbereichs D ist. Diese Approximation ist ein Parallelogramm P_i , das wie oben von den partiellen Ableitungen $\mathbf{x}_u(u_i, v_i)$ und $\mathbf{x}_v(u_i, v_i)$ aufgespannt wird. Der Fluss ergibt sich jedoch aus der Intensität des Durchflusses, sprich aus der Projektion eines Vektors $\mathbf{v}(\mathbf{x}(u_i, v_i))$, mit dem wir Richtung und – über die Länge – die Intensität der Durchflusses an der Stelle $\mathbf{x}(u_i, v_i)$ modellieren, auf die Normale des Parallelogramms. Weil der Vektor $\mathbf{v}(\mathbf{x})$ von \mathbf{x} abhängt, liegt eigentlich eine Funktion $\mathbf{v} : D_{\mathbf{v}} \rightarrow \mathbb{R}^3$ mit $F \subseteq D_{\mathbf{v}} \subseteq \mathbb{R}^3$, ein sogenanntes **Vektorfeld** vor. Weil der Durchfluss (annähernd) linear von der Fläche des Parallelogramms abhängt, liefert das Skalarprodukt

$$\mathbf{v}(\mathbf{x}(u_i, v_i))(\mathbf{x}_u(u_i, v_i) \times \mathbf{x}_v(u_i, v_i))$$

gerade jenen Beitrag, der dem Flächenstück F_i entspricht. Das Volumen V der Flüssigkeit, die in einer Zeiteinheit F durchströmt ist also annähernd gegeben durch

$$V \approx \sum_i \mathbf{v}(\mathbf{x}(u_i, v_i))(\mathbf{x}_u(u_i, v_i) \times \mathbf{x}_v(u_i, v_i)) \lambda_2(D_i).$$

Approximationsargumente ähnlich den bisherigen an analoger Stelle führen (jedenfalls sofern \mathbf{x} stetig differenzierbar ist und \mathbf{v} stetig von \mathbf{x} abhängt), zur exakten Formel

$$V = \int_D \mathbf{v}(\mathbf{x}(u, v))(\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)) \, d\lambda_2(u, v).$$

Analog zu Oberflächenintegralen erster Art sind auch die Schreibweisen

$$V = \int_F \mathbf{v} \, d\mathbf{O} = \int_D \mathbf{v}(\mathbf{x}(u, v))(\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)) \, du \, dv$$

sind gebräuchlich.

Um Verwechslungen zu vermeiden, halte man sich nochmals den Unterschied zwischen Oberflächenintegralen erster und zweiter Art vor Augen. Im ersten Fall ist der Integrand das Produkt zweier Zahlenwerte (dem Wert der Funktion $f(\mathbf{x})$ mit der Länge des Vektorproduktes $\mathbf{x}_u \times \mathbf{x}_v$), im zweiten Fall das Skalarprodukt zweier Vektoren (des durch das Vektorfeld gegebenen Vektors $\mathbf{v}(\mathbf{x}(u, v))$ und des Vektorproduktes $\mathbf{x}_u \times \mathbf{x}_v$ selbst).

Übungsaufgabe 246. (T) Wie lautet das bei Oberflächenintegralen zweiter Art auftretende Oberflächenelement $d\mathbf{O}$ eines Zylinders mit konstantem Radius r ?

$$\mathbf{x}(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix} ?$$

Übungsaufgabe 247. (T) Wie lautet das bei Oberflächenintegralen zweiter Art auftretende Oberflächenelement $d\mathbf{O}$ einer Kugel mit konstantem Radius r ?

$$\mathbf{x}(r, \varphi, z) = \begin{pmatrix} r \cos \varphi \sin \vartheta \\ r \sin \varphi \sin \vartheta \\ r \cos \vartheta \end{pmatrix} ?$$

Übungsaufgabe 248. (T) Gegeben ist der Kegel mit Spitze im Ursprung, Höhe 1 und halbem Öffnungswinkel $\pi/4$, außerdem das (konstante) Vektorfeld

$$\mathbf{v}(x, y, z) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -1 \right).$$

Berechnen Sie den Fluß von V durch die Mantelfläche und durch die Deckfläche des Kegels und fertigen Sie eine aussagekräftige Skizze an.

Übungsaufgabe 249. (T) Gegeben ist das Tetraeder T mit Eckpunkten $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$. sowie das Vektorfeld

$$\mathbf{v}(\mathbf{x}) := \begin{pmatrix} 2y \\ 2x \\ xyz \end{pmatrix}.$$

Berechnen Sie das Oberflächenintegral zweiter Art $\iint_{\partial T} \mathbf{v} d\mathbf{O}$. ∂T bezeichnet hier die Oberfläche des Tetraeders.

Übungsaufgabe 250. (T) Gegeben ist die Halbkugel (östliche Hemisphäre)

$$H := \{(x, y, z) : x^2 + y^2 + z^2 \leq 1, y \geq 0\}$$

sowie das Vektorfeld

$$\mathbf{v}(x, y, z) := \begin{pmatrix} x \\ y \\ z^2 \end{pmatrix}.$$

Berechnen Sie das Oberflächenintegral zweiter Art $\iint_{\partial H} \mathbf{v} d\mathbf{O}$. ∂H bezeichnet hier die Oberfläche der Halbkugel.

Übungsaufgabe 251. (T) Gegeben sind zwei Kugeln K_1 und K_2 (Mittelpunkt im Ursprung) mit Radien $R_1 < R_2$. Es bezeichne K_1^2 jenes Gebiet (Kugelschale), welches aus

Punkten besteht, die zwischen beiden Kugeln liegen, d.h. jene \mathbf{x} mit $R_1 < \|\mathbf{x}\| < R_2$. Berechnen Sie

$$\int \int_{\partial K_1^2} \frac{\mathbf{x}}{\|\mathbf{x}\|^3} d\mathbf{O}.$$

Hinweis: Achten Sie hier auf die Orientierung des Normalvektors und verwenden Sie dass für Kugeln $d\mathbf{O}$ parallel zu \mathbf{x} ist.

4.3.5 Oberflächenintegrale zweiter Art, Ergänzungen

Inhalt in Kurzfassung: Wie schon bei den Oberflächenintegralen erster Art folgen auch auf jene zweiter Art (4.3.4) nun noch einige Ergänzungen und Aspekte, die nicht Prüfungsstoff sind, dem Verständnis aber durchaus dienen können.

Definition

Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein Vektorfeld und $\mathbf{x} : D \rightarrow G$ die Parameterdarstellung eines Flächenstücks F in G . Gibt es eine Zahl $O(\mathbf{v}, F)$ mit der folgenden Eigenschaft: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für jede Zerlegung D_1, \dots, D_n von D in elementare Bereiche mit Feinheit $< \delta$ für beliebige Wahl von Punkten $(\sigma_1, \tau_1) \in D_1, \dots, (\sigma_n, \tau_n) \in D_n$ gilt

$$(1) \quad \left| O(\mathbf{v}, F) - \sum_{k=1}^n \mathbf{v}(\mathbf{x}(\sigma_k, \tau_k)) \cdot \mathbf{n}(\sigma_k, \tau_k) |F_k| \right| < \varepsilon,$$

dann heißt $O(\mathbf{v}, F)$ das **Oberflächenintegral** von \mathbf{v} über F .

Dabei ist

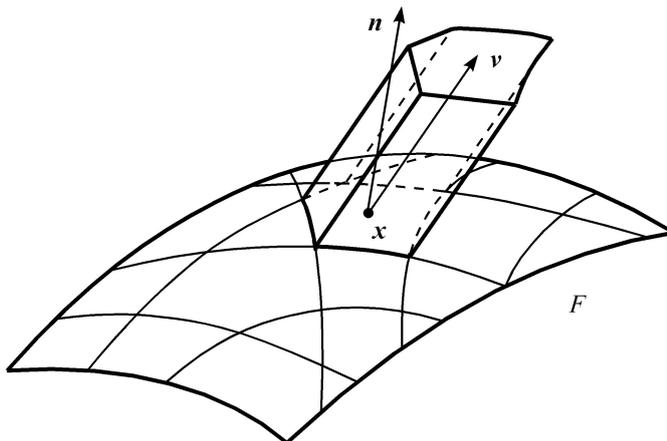
$$\mathbf{n}(u, v) = \frac{\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)}{\|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\|} \text{ für } (u, v) \in D$$

der Normaleneinheitsvektor von F in $\mathbf{x}(u, v)$ und F_1, \dots, F_n sind die Flächenstücke mit den

Parameterdarstellungen $\mathbf{x}|_{D_1}, \dots, \mathbf{x}|_{D_n}$. Symbole

$$\int_F \mathbf{v} d\mathbf{O}, \int_F \mathbf{v} d\mathbf{S}, \iint_F \mathbf{v} d\mathbf{O}, \dots$$

Es genügt also, das Vektorfeld \mathbf{v} auf dem Flächenstück F zu kennen.



Eindeutigkeit, Existenz und Berechnung

Existiert das Oberflächenintegral, so ist es eindeutig. Für die Praxis wichtig ist folgender

Satz (Existenz und Rückführung auf Doppelintegrale). *Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein stetiges Vektorfeld und $\mathbf{x} : D \rightarrow G$ sei eine Parameterdarstellung eines Flächenstücks F , \mathbf{x} habe stetige partielle Ableitungen und es sei $\mathbf{x}_u \times \mathbf{x}_v \neq \mathbf{0}$ auf F . Dann existiert das Oberflächenintegral von \mathbf{v} über F und es gilt*

$$\int_F \mathbf{v} d\mathbf{O} = \iint_D \mathbf{v}(\mathbf{x}(u, v)) \cdot \mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v) du dv.$$

Physikalische Deutung

Wir nehmen an, dass \mathbf{v} das stationäre Geschwindigkeitsfeld der Strömung einer inkompressiblen Flüssigkeit der Dichte 1 ist. Wir betrachten das Flächenstück F_k . Die durch F_k hindurchtretende

Flüssigkeit hat annähernd die Geschwindigkeit $\mathbf{v}(\mathbf{x}(\sigma_k, \tau_k))$. Die in einer Sekunde durch F_k hindurchtretende Flüssigkeit erfüllt also annähernd ein schräges Prisma mit Grundfläche F_k und Deckfläche $F_k + \mathbf{v}(\mathbf{x}(\sigma_k, \tau_k))$. Sein Volumen ist also annähernd Höhe mal Grundfläche, also

$$\mathbf{v}(\mathbf{x}(\sigma_k, \tau_k)) \cdot \mathbf{n}(\sigma_k, \tau_k) |F_k|.$$

Es folgt, dass die pro Sekunde durch F hindurchtretende Flüssigkeit annähernd gleich ist

$$\sum_{k=1}^n \mathbf{v}(\mathbf{x}(\sigma_k, \tau_k)) \cdot \mathbf{n}(\sigma_k, \tau_k) |F_k|.$$

Beachtet man die Definition des Oberflächenintegrals und insbesondere (1), so erhält man: *Das Oberflächenintegral*

$$\int_F \mathbf{v} d\mathbf{O}$$

ist gleich der pro Sekunde durch F hindurchtretenden Nettomenge an Flüssigkeit.

Ähnliche Aussagen erhält man bei Wärmeströmungen und elektrischen und magnetischen Feldern.

Beispiel (Gesetz von Faraday). Es seien $\mathbf{E}(\mathbf{x}, t)$ und $\mathbf{H}(\mathbf{x}, t)$ zusammengehörige elektrische und magnetische Felder im Vakuum. Ist F ein Flächenstück, das von einer einfach geschlossenen Kurve C berandet wird, dann gilt (bei passender Wahl des Einheitensystems)

$$\int_C \mathbf{E} d\mathbf{x} = -\frac{d}{dt} \iint_F \mathbf{H} d\mathbf{O},$$

daher ist also die „Gesamtspannung“ längs C ist gleich der negativen zeitlichen Änderung des „magnetischen Flusses“ durch F . Wäre C ein Draht, so würde in C ein Strom fließen, der proportional zur Gesamtspannung ist.

Beispiel (Gesetz von Gauß). Es sei $\mathbf{E}(\mathbf{x})$ ein stationäres elektrisches Feld im Vakuum, hervorgerufen durch gewisse Ladungen. Ferner sei F eine geschlossene Fläche mit einem Innen- und einem Außengebiet (Beispiele: Kugel, Torus). Dann ist

$$\iint_F \mathbf{E} d\mathbf{O}$$

die „Gesamtladung“ im Innengebiet von F .

Beispiel. Es sei

$$\mathbf{v}(\mathbf{x}) = \begin{pmatrix} 1 \\ x \\ z \end{pmatrix}.$$

Man berechne den Fluss pro Sekunde aus der Halbkugel

$$x^2 + y^2 + z^2 \leq 1, \quad z \geq 0.$$

Ihr Rand setzt sich aus zwei Flächenstücken zusammen.

$$F_1 : \mathbf{x} = \begin{pmatrix} u \\ v \\ \sqrt{1-u^2-v^2} \end{pmatrix}, \quad u^2 + v^2 \leq 1, \quad F_2 : \mathbf{x} = \begin{pmatrix} u \\ v \\ 0 \end{pmatrix}, \quad u^2 + v^2 \leq 1.$$

4 Integralrechnung in mehreren Variablen

Nettodurchfluss pro Sekunde: (1) – (2) mit

$$\begin{aligned}(1) &= \iint_{F_1} \mathbf{v} d\mathbf{O} = \iint_{u^2+v^2 \leq 1} \mathbf{v} \cdot \mathbf{x}_u \times \mathbf{x}_v dudv = \left| \mathbf{x}_u \times \mathbf{x}_v = \begin{pmatrix} \frac{u}{\sqrt{1-u^2-v^2}} \\ \frac{v}{\sqrt{1-u^2-v^2}} \\ 1 \end{pmatrix} \right| \\ &= \iint_{u^2+v^2 \leq 1} \left(\frac{u}{\sqrt{1-u^2-v^2}} + \frac{uv}{\sqrt{1-u^2-v^2}} + \sqrt{1-u^2-v^2} \right) dudv \\ &= \int_0^{2\pi} \int_0^1 \left(\frac{r^2 \cos \varphi}{\sqrt{1-r^2}} + \frac{r^3 \cos \varphi \sin \varphi}{\sqrt{1-r^2}} + r\sqrt{1-r^2} \right) dr d\varphi \\ &= 0 + 0 - 2\pi \int_0^1 ((1-r^2)^{3/2})' \frac{2}{3} \cdot \frac{1}{2} dr = \frac{2\pi}{3}. \\ (2) &= \iint_{F_2} \mathbf{v} d\mathbf{O} = \iint_{u^2+v^2 \leq 1} 0 dudv = 0.\end{aligned}$$

Achtung: Bei einer geschlossenen Fläche, die sich aus mehreren Flächenstücken zusammensetzt, müssen die Parameterdarstellungen der einzelnen Flächenstücke so gewählt werden, dass die Normalvektoren stets in das Äußere weisen.

4.4 Vektoranalysis und Integralsätze

In der Vektoranalysis stehen gewisse Skalar- und Vektorfelder in Zentrum, die sich aus den partiellen Ableitungen gegebener Felder ergeben. Schon bekannt ist das Gradientenfeld $\text{grad } P$ eines vorgegebenen Skalarfeldes P (für Potential). Dabei handelt es sich um das aus den partiellen Ableitungen von P als Komponenten gebildete Vektorfeld. Es ist Gegenstand von 4.4.1. Die Rotation $\text{rot } \mathbf{v}$ eines Vektorfeldes \mathbf{v} ist für dreidimensionales \mathbf{v} selbst wieder ein Vektorfeld, im Zweidimensionalen ein Skalarfeld. Definition und geometrische Deutung der Rotation sind Gegenstand von 4.4.2. Die Divergenz $\text{div } \mathbf{v}$ eines Vektorfeldes \mathbf{v} , siehe 4.4.3, entsteht als Summe gewisser partieller Ableitungen und ist somit eine Skalarfeld. Durch die Integralsätze werden diese Begriffe zueinander in Beziehung gesetzt. Die gemeinsame Struktur der Integralsätze wird in 4.4.4 herausgearbeitet. Im konkreten werden sie in den nachfolgenden Unterabschnitten 4.4.5 (Satz von Green), 4.4.6 (Satz von Stokes) und 4.4.7 (Satz von Gauß) behandelt. Der letzte Unterabschnitt 4.4.8 enthält Ergänzungen, die nicht zum Prüfungsstoff gehören aber hilfreiche Illustrationen des Stoffes darstellen.

Im gesamten Abschnitt bezeichnet G ein Gebiet, d.h. eine offene und zusammenhängende Menge im \mathbb{R}^2 oder im \mathbb{R}^3 . Meist fungiert G als Definitionsbereich eines Skalar- oder Vektorfeldes.

4.4.1 Gradient

Inhalt in Kurzfassung: Dieser Unterabschnitt bringt wenig Neues. Vor allem handelt es sich um eine Wiederholung des Begriffs des Gradientenfeldes als Vektorfeld, das durch Ableitung eines Skalarfeldes zustandekommt. Neu ist die Notation mit dem Nabla-Operator ∇ .

Wir beginnen mit einigen Wiederholungen und Ergänzungen über **Gradientenfelder**: Hat das Skalarfeld $P : G \rightarrow \mathbb{R}$ stetige partielle Ableitungen, dann nennt man das Vektorfeld

$$\text{grad } P(\mathbf{x}) = \begin{pmatrix} P_x(\mathbf{x}) \\ P_y(\mathbf{x}) \\ P_z(\mathbf{x}) \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in G$$

das **Potential-** oder **Gradientenfeld** von P . Das Skalarfeld P heißt dann das **Potential** des Gradientenfeldes. Mit dem **Nabla(-operator)** genannten Symbol

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix}$$

schreibt man das Gradientenfeld auch gerne in der Form

$$\text{grad } P = \nabla P.$$

Wir haben früher ein Kriterium angegeben, mit dem man überprüfen kann, ob ein Vektorfeld ein Gradientenfeld ist. Mit dem folgenden Resultat fassen wir noch weitere wichtige Eigenschaften des Gradienten zusammen.

Hat das Skalarfeld $P : G \rightarrow \mathbb{R}$ stetige partielle Ableitungen und es sei $\text{grad } P \neq \mathbf{0}$ auf G , dann stellt die Gleichung

$$P = \text{const}$$

eine sogenannte **Potentialfläche** dar. An jeder Stelle $\mathbf{a} \in G$ ist $\text{grad } P(\mathbf{a})$ ein Normalvektor der Fläche (oder Kurve)

$$P(\mathbf{x}) = P(\mathbf{a}).$$

$\text{grad } P(\mathbf{a})$ weist von \mathbf{a} aus in die Richtung, in der P am stärksten zunimmt (also in die Richtung mit maximaler Richtungsableitung).

Beispiel (Wärmestrom). Es sei $T(\mathbf{x}, t)$ die Temperatur in einem Körper an der Stelle \mathbf{x} zum Zeitpunkt t . Dann fließt bekanntlich Wärme aus heißeren in kältere Bereiche. Genauer gilt für den Wärmestrom \mathbf{w}

$$\mathbf{w}(\mathbf{x}, t) = -l \text{grad } T(\mathbf{x}, t) \quad (l = l(\mathbf{x}) \text{ Wärmeleitfähigkeit an der Stelle } \mathbf{x}).$$

Übungsaufgabe 252. (*T*) Es sei f eine radialsymmetrische Funktion, d.h. $f(\vec{r}) = \phi(r)$ für eine Funktion $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ und $r = \|\vec{r}\|$. Berechnen Sie den Gradienten von f allgemein und für die Spezialfälle $\phi_i(r) = r^i$ für $i = -1, 0, 1$.

4.4.2 Rotation

Inhalt in Kurzfassung: Für die Dimensionen $n = 2$ und $n = 3$ beschreibt die Rotation $\operatorname{rot} \mathbf{v}$ eines n -dimensionalen Vektorfeldes $\mathbf{v} : G \rightarrow \mathbb{R}^n$, $G \subseteq \mathbb{R}^n$ an einer Stelle \mathbf{x} anschaulich, wie stark sich \mathbf{v} in der Nähe von \mathbf{x} um diesen Punkt \mathbf{x} dreht. Weil Drehungen in der Ebene ($n = 2$) um einen Punkt durch einen einzigen Parameter (die Winkelgeschwindigkeit), im Raum ($n = 3$) durch drei Parameter (zwei für die Rotationsachse, einen für die Winkelgeschwindigkeit) beschrieben werden, sind die Werte von $\operatorname{rot} \mathbf{v}$ Zahlen im Fall $n = 2$ und Vektoren im \mathbb{R}^3 im Fall $n = 3$. Somit ist $\operatorname{rot} \mathbf{v} : G \rightarrow \mathbb{R}$ für $n = 2$ ein Skalarfeld, während $\operatorname{rot} \mathbf{v} : G \rightarrow \mathbb{R}^3$ für $n = 3$ ein Vektorfeld ist.

So wie der Gradient ist auch die Rotation eine Größe, die sich aus partiellen Ableitungen zusammensetzt, allerdings angewendet auf ein Vektorfeld $\mathbf{v} : G \rightarrow \mathbb{R}^n$ mit $G \subseteq \mathbb{R}^n$ ($n = 2, 3$) und kein Skalarfeld. Wir schreiben

$$\mathbf{v} : G \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$$

für ebene Vektorfelder ($n = 2$) und

$$\mathbf{v} : G \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} u(x, y, z) \\ v(x, y, z) \\ w(x, y, z) \end{pmatrix}$$

für Vektorfelder im \mathbb{R}^3 .

Und zwar soll die Rotation $\operatorname{rot} \mathbf{v}$ in einem Punkt \mathbf{x} angeben, wie sehr sich \mathbf{v} in der Umgebung von \mathbf{x} um \mathbf{x} (entgegen dem Uhrzeigersinn) dreht. Hier besteht zwischen den Dimensionen $n = 2$ und $n = 3$ ein wesentlicher Unterschied.

Für $n = 2$, also in der Ebene, ist nur eine geometrische Ausrichtung von Rotation denkbar. Entsprechend ist nur ein eindimensionaler Parameter nötig, der die Intensität der Rotation beschreibt. Wie man sich durch eine Zeichnung schnell klar macht, wird eine Drehung gegen den Uhrzeigersinn umso intensiver, je stärker die y -Komponente des Vektorfeldes, also $v(x, y)$ in Richtung der x -Achse zunimmt, andererseits je stärker die x -Komponente des Vektorfeldes, also $u(x, y)$ entlang der y -Achse in negative Richtung weist. Entsprechend lautet die Definition für ein differenzierbares Vektorfeld $\mathbf{v} : G \rightarrow \mathbb{R}^2$ mit $G \subseteq \mathbb{R}^2$:

$$\operatorname{rot} \mathbf{v} := v_x - u_y.$$

Dass diese Definition tatsächlich sehr gut die intendierte geometrische Deutung zulässt, wird durch die Rotationsversion des Integralsatzes von Green (4.4.5.1) bestätigt werden.

Für $n = 3$, d.h. für dreidimensionale Vektorfelder muss eine sinnvolle Definition der Rotation mehr Information enthalten. Wie wir bereits wissen, lässt sich eine Rotation im Raum durch drei Parameter beschreiben, zum Beispiel einen Vektor im \mathbb{R}^3 , dessen Richtungsvektor die Rotationsachse angibt und dessen Länge die Rotationsgeschwindigkeit. Das Vorzeichen korrespondiert mit der Orientierung der Rotation. Es erweist sich, dass folgende Definition genau dieser Bedeutung entspricht:

Sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ mit $G \subseteq \mathbb{R}^3$ ein differenzierbares Vektorfeld mit. Dann ist die **Rotation** von \mathbf{v} , Symbole $\operatorname{rot} \mathbf{v}$ oder $\operatorname{curl} \mathbf{v}$, das Vektorfeld

$$\operatorname{rot} \mathbf{v} = \nabla \times \mathbf{v} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} w_y - v_z \\ u_z - w_x \\ v_x - u_y \end{pmatrix}.$$

Gilt

$$\operatorname{rot} \mathbf{v} = \mathbf{0},$$

so nennt man \mathbf{v} **rotationsfrei** oder **wirbelfrei**.

Eine Heuristik dafür, dass, analog zur Rotation für $n = 2$, auch diese Definition die intendierte Bedeutung als Rotation im Raum hat, ist komplizierter. Eine überzeugende Bestätigung ergibt sich wieder durch einen Integralsatz, nämlich jenen von Stokes (siehe 4.4.6.1).

Beispiel (Maxwellsche Gleichungen). Zwischen einem elektrischen Feld $\mathbf{E}(\mathbf{x}, t)$ und dem zugehörigen Magnetfeld $\mathbf{H}(\mathbf{x}, t)$ herrschen im Vakuum die Beziehungen

$$\operatorname{rot} \mathbf{H} = \frac{\partial \mathbf{E}}{\partial t}, \quad \operatorname{rot} \mathbf{E} = -\frac{\partial \mathbf{H}}{\partial t}.$$

Beispiel. Wir betrachten das Geschwindigkeitsfeld eines mit konstanter Winkelgeschwindigkeit ω um eine feste Achse durch \mathbf{o} rotierenden Körpers:

$$\mathbf{v}(\mathbf{x}) = \boldsymbol{\omega} \times \mathbf{x} \quad (\|\boldsymbol{\omega}\| = \omega).$$

Gesucht ist $\operatorname{rot} \mathbf{v}$. Da

$$\boldsymbol{\omega} \times \mathbf{x} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z\omega_2 - y\omega_3 \\ x\omega_3 - z\omega_1 \\ y\omega_1 - x\omega_2 \end{pmatrix}$$

erhalten wir

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) = \nabla \times (\boldsymbol{\omega} \times \mathbf{x}) = \begin{pmatrix} \omega_1 + \omega_1 \\ \omega_2 + \omega_2 \\ \omega_3 + \omega_3 \end{pmatrix} = 2\boldsymbol{\omega}.$$

Physikalische Deutung

Es sei \mathbf{v} das Geschwindigkeitsfeld einer stationären Strömung einer inkompressiblen Flüssigkeit. Um $\operatorname{rot} \mathbf{v}$ an einer Stelle \mathbf{x} zu bestimmen, bringt man ein Schaufelrädchen im Punkt \mathbf{x} an und bringt die Achse in die Lage, in der es mit maximaler Winkelgeschwindigkeit rotiert. Dann weist die Achse in Richtung des Vektors $\operatorname{rot} \mathbf{v}(\mathbf{x})$ (Rechte-Hand-Regel) und $\|\operatorname{rot} \mathbf{v}(\mathbf{x})\|$ ist gleich der doppelten Winkelgeschwindigkeit.

Eine weitere Deutung der Rotation wird sich aus dem Integralsatz von Stokes ergeben.

Potentialfelder

Aus dem früher besprochenen Kriterium für Potentialfelder ergibt sich Folgendes: Ein Vektorfeld auf einem einfach zusammenhängenden Gebiet mit stetigen partiellen Ableitungen ist genau dann ein Potentialfeld, wenn es rotationsfrei ist.

Übungsaufgabe 253. (T) Es sei \vec{V} ein radialsymmetrisches Vektorfeld, d.h. $\vec{V}(\vec{r}) = \phi(r)\vec{e}_r$ für eine Funktion $\phi: \mathbb{R}^+ \rightarrow \mathbb{R}^+$, wobei $\vec{e}_r = \frac{\vec{r}}{r}$ und $r = \|\vec{r}\|$. Zeigen Sie dass \vec{V} stets wirbelfrei ist und ein Potential besitzt.

Hinweis: Verwenden Sie die Rechenregel $\nabla \times (f\vec{A}) = f(\nabla \times \vec{A}) + (\nabla f) \times \vec{A}$.

Übungsaufgabe 254. (E) Sei \mathbf{c} ein fester Vektor im \mathbb{R}^3 . Bestimmen Sie alle Vektorfelder \mathbf{V} auf \mathbb{R}^3 mit $\text{rot } \mathbf{V} = \mathbf{c}$. Das Feld \mathbf{V} nennt man dann ein „Vektorpotential“ von \mathbf{c} .

Anleitung: Finden Sie zunächst irgendein Vektorfeld mit $\text{rot } \mathbf{V} = \mathbf{c}$. Überlegen Sie sich nun, dass, wenn \mathbf{W} ebenfalls ein Vektorpotential ist, das Feld $\mathbf{V} - \mathbf{W}$ die Integrabilitätsbedingungen erfüllt.

4.4.3 Divergenz

Inhalt in Kurzfassung: Die Divergenz $\text{div } \mathbf{v}$ eines Vektorfeldes $\mathbf{v}: G \rightarrow \mathbb{R}^n$, $G \subseteq \mathbb{R}^n$, an einer Stelle \mathbf{x} beschreibt für beide Dimensionen $n = 2$ und $n = 3$ anschaulich, wie stark der Fluss entlang der Koordinatenachsen zunimmt und berechnet sich demnach als die Summe von zwei bzw. drei partiellen Ableitungen. Somit ist $\text{div } \mathbf{v}: G \rightarrow \mathbb{R}$ sowohl für $n = 2$ als auch für $n = 3$ ein Skalarfeld.

Die dritte Größe, die sich aus partiellen Ableitungen eines Feldes ergibt und die wir hier besprechen wollen, ist die Divergenz div . Wie die Rotation bezieht sie sich auf ein Vektorfeld $\mathbf{v}: G \rightarrow \mathbb{R}^n$ mit $G \subseteq \mathbb{R}^n$ ($n = 2, 3$). Und zwar soll $\text{div } \mathbf{v}$ an einer Stelle \mathbf{x} angeben, wie stark die Strömung in positive Richtung der Koordinatenachsen zunimmt. Mit der bisherigen Notation

$$\mathbf{v}: G \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}$$

beziehungsweise

$$\mathbf{v}: G \rightarrow \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} u(x, y, z) \\ v(x, y, z) \\ w(x, y, z) \end{pmatrix}$$

setzt sich $\text{div } \mathbf{v}$ also (für differenzierbares \mathbf{v}) additiv aus den partiellen Ableitungen u_x , v_y und, für $n = 3$, w_z zusammen. Somit ist $\text{div } \mathbf{v}: G \rightarrow \mathbb{R}$ selbst ein Skalarfeld, nämlich

$$\text{div } \mathbf{v} := u_x + v_y = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$$

beziehungsweise

$$\text{div } \mathbf{v} := u_x + v_y + w_z = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$$

im Fall $n = 2$ bzw. $n = 3$.

Unter Benutzung des Nablaoperators schreibt man dafür auch

$$\nabla \cdot \mathbf{v}.$$

Man nennt ein Vektorfeld \mathbf{v} mit stetigen partiellen Ableitungen **quellfrei**, falls gilt

$$\operatorname{div} \mathbf{v} = 0.$$

Physikalische Deutungen ergeben sich wieder aus Integralsätzen, nämlich für $n = 2$ durch den Satz von Green in der Divergenzversion 4.4.5.2 bzw. für $n = 3$ durch den Divergenzatz von Gauß 4.4.7.1. Dazu aber noch einige weitere Erläuterungen:

Es sei zuerst \mathbf{v} das Geschwindigkeitsfeld einer stationären Gasströmung. Schreibt man einem Punkt \mathbf{x} einen kleinen Bereich zum Zeitpunkt $t = 0$ um, und wird der Bereich von der Gasströmung mitgenommen, so ändert sich sein Volumen $V(t)$ mit der Zeit t . Es ist

$$\frac{V'(0)}{V(0)} \sim \operatorname{div} \mathbf{v}(\mathbf{x}),$$

wobei die Übereinstimmung umso genauer ist, je kleiner der Bereich ursprünglich war.

Eine zweite Interpretation lautet wie folgt: Es sei \mathbf{v} das Geschwindigkeitsfeld einer stationären Strömung einer inkompressiblen Flüssigkeit. Schreibt man einem Punkt \mathbf{x} eine kleine geschlossene Fläche um, dann ist $\operatorname{div} \mathbf{v}(\mathbf{x})$ annähernd die Differenz der ausströmenden und einströmenden Flüssigkeit pro Zeiteinheit dividiert durch das Volumen des Bereichs. Da die Flüssigkeit inkompressibel ist, muss in dem Bereich Flüssigkeit erzeugt oder vernichtet werden — je nachdem ob diese Differenz positiv oder negativ ist. Man spricht von Quellen bzw. Senken. Die Divergenz gibt also die Differenz der Menge der erzeugten und der Menge der vernichteten Flüssigkeit pro Volumeneinheit an. Man nennt sie daher auch gerne **Quelldichte**; vgl. den Integralsatz von Gauß.

Beispiel (Gravitationsfeld). Die Kraft, die eine kontinuierliche Massenverteilung auf eine Einheitspunktmasse in \mathbf{x} ausübt, ist nach dem Gravitationsgesetz gleich

$$\mathbf{f}(\mathbf{x}) = \gamma \int \int \int \frac{\varrho(\mathbf{u})(\mathbf{u} - \mathbf{x})}{\|\mathbf{u} - \mathbf{x}\|^3} \, dudvdw.$$

(Das Integral ist komponentenweise zu verstehen.)

Es folgt aus der Leibnizschen Regel für Parameterintegrale:

$$\begin{aligned} \operatorname{div} \mathbf{f}(\mathbf{x}) &= \gamma \frac{\partial}{\partial x} \int \int \int \frac{\varrho(\mathbf{u})(u - x)}{((u - x)^2 + (v - y)^2 + (w - z)^2)^{3/2}} \, dudvdw + \dots \\ &= \gamma \int \int \int \frac{\varrho(\mathbf{u})(-1)\|\mathbf{u} - \mathbf{x}\|^3 - \varrho(\mathbf{u})(u - x)\frac{3}{2}\|\mathbf{u} - \mathbf{x}\|2(u - x)(-1)}{\|\mathbf{u} - \mathbf{x}\|^6} \, dudvdw + \dots \\ &= -\gamma \int \int \int \frac{\varrho(\mathbf{u})(3\|\mathbf{u} - \mathbf{x}\|^3 - 3\|\mathbf{u} - \mathbf{x}\|((u - x)^2 + (v - y)^2 + (w - z)^2))}{\|\mathbf{u} - \mathbf{x}\|^6} \, dudvdw = 0. \end{aligned}$$

4 Integralrechnung in mehreren Variablen

D.h. das Gravitationsfeld (im Äußeren der es erzeugenden Massenverteilung) ist quellfrei. Für das Potential P des Gravitationsfeldes \mathbf{f} gilt also

$$\operatorname{div} \operatorname{grad} P = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} + \frac{\partial^2 P}{\partial z^2} = \Delta P = 0.$$

Übungsaufgabe 255. (T^*) Es sei \vec{V} ein radialsymmetrisches Vektorfeld, d.h. $\vec{V}(\vec{r}) = \phi(r)\vec{e}_r$ für eine Funktion $\phi: \mathbb{R}^+ \rightarrow \mathbb{R}^+$, wobei $\vec{e}_r = \frac{\vec{r}}{r}$ und $r = \|\vec{r}\|$. Berechnen Sie die Divergenz von f allgemein und für die Spezialfälle $\phi_i(r) = r^i$ für $i = -2, 1, 2$.

Hinweis: Verwenden Sie die Rechenregel $\nabla \cdot (f\vec{A}) = f(\nabla \cdot \vec{A}) + (\nabla f) \cdot \vec{A}$.

Übungsaufgabe 256. (T) Bestimmen Sie die Konstante μ so, dass das Vektorfeld $\mathbf{W}_\mu(x, y, z) = (1 + \mu x, x, 1)$ quellfrei ist, also $\operatorname{div} \mathbf{W}_\mu = 0$ erfüllt. Geben Sie alle Vektorpotentiale an, d.h. alle Vektorfelder \mathbf{A}_μ mit $\operatorname{rot} \mathbf{A}_\mu = \mathbf{W}_\mu$ an.

Identitäten für Vektoroperationen, Zusammenfassung:

Für grad , rot , div gibt es eine Reihe von Identitäten: Es seien $f, g: G \rightarrow \mathbb{R}$ Skalarfelder und $\mathbf{v}, \mathbf{w}: G \rightarrow \mathbb{R}^3$ Vektorfelder mit stetigen partiellen Ableitungen erster oder zweiter Ordnung, je nach Erfordernis. Dann gilt

- $\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f$
- $\operatorname{rot}(f\mathbf{v}) = \operatorname{grad} f \times \mathbf{v} + f \operatorname{rot} \mathbf{v}$
- $\operatorname{div}(f\mathbf{v}) = \operatorname{grad} f \cdot \mathbf{v} + f \operatorname{div} \mathbf{v}$
- $\operatorname{div}(\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot \operatorname{rot} \mathbf{v} - \mathbf{v} \cdot \operatorname{rot} \mathbf{w}$
- $\operatorname{rot}(\operatorname{grad} f) = \mathbf{o}$
- $\operatorname{div}(\operatorname{rot} \mathbf{v}) = 0$
- $\operatorname{div}(\operatorname{grad} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \Delta f$ (Laplace-Operator)

Übungsaufgabe 257. (P) Bestätigen Sie durch explizites Nachrechnen die beiden Regeln

$$\nabla \times (\nabla f) = 0, \quad \nabla \cdot (\nabla \times \vec{A}) = 0.$$

Dabei sind $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ und $\vec{A}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ als stetig differenzierbar vorausgesetzt.

Übungsaufgabe 258. (P) Berechnen Sie allgemein für ein Skalarfeld $\phi: \mathbb{R}^3 \rightarrow \mathbb{R}$ und ein Vektorfeld $\mathbf{V}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$\operatorname{div}(\nabla \phi), \operatorname{rot}(\nabla \phi), \operatorname{div}(\operatorname{rot} \mathbf{V}), \operatorname{rot}(\phi \mathbf{V}), \operatorname{rot} \operatorname{div}(\phi \mathbf{V}).$$

Bestätigen Sie Ihre Ergebnisse anhand der konkreten Beispiele $\phi(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ und $\mathbf{V}(x, y, z) = (x, y, z)$ und fertigen Sie jeweils eine aussagekräftige Skizze an!

Übungsaufgabe 259. (P) Rechnen Sie folgende Identität für Vektorfelder $\mathbf{u} = (u_1, u_2, u_3)$ nach:

$$\operatorname{rot}(\operatorname{rot}(\mathbf{u})) = \operatorname{grad}(\operatorname{div}(\mathbf{u})) - \Delta(\mathbf{u}),$$

wobei unter $\Delta(\mathbf{u})$ der Vektor $(\Delta u_1, \Delta u_2, \Delta u_3)$ zu verstehen ist. $\Delta = \nabla \cdot \nabla$ ist der Laplace-Operator.

4.4.4 Gemeinsamkeiten der Integralsätze

Inhalt in Kurzfassung: Die Integralsätze von Green, Stokes und Gauß lassen sich – so wie auch der bereits aus Mathematik 1 vertraute Hauptsatz der Differential- und Integralrechnung – als Spezialfälle eines allgemeinen Satzes auffassen, der ebenfalls nach Stokes benannt ist. Leider ist es hier nicht möglich, die zu seinem vollen Verständnis erforderliche Theorie der Integration von Differentialformen in der entsprechenden Allgemeinheit zu entwickeln. Es ist aber möglich, die sehr einprägsame formale Struktur dieses allgemeinen Satzes anzugeben. In Verbindung mit dem Hauptsatz der Differential- und Integralrechnung fällt es dann leichter, sich die Aussagen der klassischen Integralsätze einzuprägen. Diese werden dann noch in den nachfolgenden Unterabschnitten im Einzelnen zu besprechen sein.

Im Zentrum der nachfolgenden drei Unterabschnitte werden die Integralsätze von Green, Stokes und Gauß stehen. Sie lassen sich alle als Analoga zum Hauptsatz der Differential- und Integralrechnung auffassen. Um das zu verstehen, ist es allerdings sinnvoll, den Hauptsatz in etwas ungewohnter Weise zu interpretieren:

Ist F eine Stammfunktion der stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$, so lautet die übliche Formulierung des Hauptsatzes:

$$\int_a^b f(x) dx = F(b) - F(a)$$

Das Ungewohnte besteht nun darin, die Differenz rechts als Integral von F über den Rand $\partial[a, b] = \{a, b\}$ des Integrationsbereichs $[a, b]$ zu verstehen. Und zwar fassen wir den Rand des (von a nach b orientierten) eindimensionalen Intervalls $[a, b]$ als eine 0-dimensionale aber geordnete Menge $\{a < b\}$ mit Anfangspunkt a und Endpunkt b auf. Wir definieren das Integral über eine solche 0-dimensionale, zweielementige geordnete Menge als

$$\int_{\partial[a,b]} F = \int_{\{a < b\}} F := F(b) - F(a),$$

also: Funktionswert am Endpunkt minus Funktionswert am Anfangspunkt. Der Hauptsatz erhält somit die Gestalt

$$\int_{[a,b]} f dx = \int_{\partial[a,b]} F.$$

Die allgemeine Form der Integralsätze ergibt sich durch Abstraktion: Das Integrationsintervall $[a, b]$ wird durch einen allgemeinen (eventuell auch höherdimensionalen) Bereich B mit Rand ∂B ersetzt und die Funktion F durch ein Objekt ω , das man **Differentialform**

nennt und zu dem es auch eine Ableitung $d\omega$ gibt, im Hauptsatz $dF = F' dx = f dx$. Die resultierende Form des Hauptsatzes ist der **allgemeine Satz von Stokes** (im Gegensatz zum klassischen Satz von Stokes 4.4.6.1):

Satz 4.4.4.1.

$$\int_B d\omega = \int_{\partial B} \omega$$

In Worten: Das Bereichsintegral über B der Ableitung $d\omega$ stimmt mit dem Integral von ω über den Rand von B überein.

Auch wenn die Bedeutung von ω hier etwas dunkel bleiben muss, so lohnen doch einige Erläuterungen zum allgemeinen Satz von Stokes:

Wichtig ist zunächst, dass die Bereiche und ihre Ränder vorzeichenbehaftet sind: Bei einem 0-dimensionalen zweielementigen Bereich hängt das Vorzeichen von der Reihenfolge ab (Anfangspunkt negativ, Endpunkt positiv – wie bei Vektoren), bei einem Intervall kommt es auf seine Orientierung an (entsprechend der Regel $\int_a^b = -\int_b^a$), bei der Parametrisierung von Kurven und den daraus resultierenden Kurvenintegralen auf den Sinn der Durchlaufung, ähnlich bei Oberflächenintegralen auf die Reihenfolge der Koordinaten in der Parametrisierung (Rechtsschraubenregel) etc.

Mit beträchtlichem begrifflichen Aufwand lässt sich eine Theorie der Integration von Differentialformen entwickeln, deren Hauptergebnis eben Satz 4.4.4.1 ist, wobei ω darin so allgemein gewählt werden kann, dass alle hier behandelten Integralsätze Spezialfälle sind:

- Der Hauptsatz der Differential- und Integralrechnung ergibt sich als Spezialfall in der oben bereits beschriebenen Weise.
- Im Satz von Green ist B ein ebener Bereich, der von einer Kurve (als ∂B) berandet wird. Die Rolle von ω übernimmt das Vektorfeld \mathbf{v} (4.4.5.1) bzw. das dazu normale Vektorfeld (4.4.5.2), und die Ableitung ist $\text{rot } \mathbf{v}$ bzw. $\text{div } \mathbf{v}$.
- Im klassischen Satz von Stokes (4.4.6.1) ist B eine zweidimensionale Fläche im dreidimensionalen Raum, die durch eine Kurve im Raum als ∂B berandet wird. Als Ableitung eines dreidimensionalen Vektorfeldes (das für ω steht und das entlang der Kurve integriert wird) fungiert $\text{rot } \mathbf{v}$.
- Im Satz von Gauß (4.4.7.1) ist B ein dreidimensionaler Bereich B mit zweidimensionalem Rand ∂B . Wieder übernimmt ein Vektorfeld \mathbf{v} die Rolle von ω und $\text{div } \mathbf{v}$ die der Ableitung $d\omega$.

In jedem Fall ist die Dimension des Randes ∂B um 1 niedriger als die Dimension von B . Doch nun zu den Integralsätzen im Einzelnen.

4.4.5 Integralsatz von Green im \mathbb{R}^2

Inhalt in Kurzfassung: Der Satz von Green bezieht sich auf stetig differenzierbare zweidimensionale Vektorfelder \mathbf{v} auf einem ebenen Bereich D , der durch eine stückweise stetig differenzierbare einfache (d.h. sich selbst nicht überschneidende) Kurve C begrenzt wird, die im mathematisch positiven Sinn (d.h. bei üblicher Beschriftung der Achsen des Koordinatensystems gegen den Uhrzeigersinn) durchlaufen wird. In seiner Rotationsvariante besagt der Satz von Green, dass das Kurvenintegral zweiter Art über \mathbf{v} entlang C denselben Wert liefert wie das Bereichsintegral über $\text{rot } \mathbf{v}$. Daraus lässt sich sehr schnell auch die Divergenzvariante des Satzes von Green ableiten (wie auch umgekehrt): Das Kurvenintegral über das zu \mathbf{v} normale Vektorfeld stimmt mit dem Bereichsintegral über die $\text{div } \mathbf{v}$ überein.

In diesem Abschnitt sei G durchwegs ein Gebiet im \mathbb{R}^2 . Der **Satz von Green** in der ersten Version (Rotationsversion) lautet:

Satz 4.4.5.1. *Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^2$ ein Vektorfeld mit stetigen partiellen Ableitungen und C eine einfach geschlossene Kurve, die samt dem von ihr berandeten, beschränkten und abgeschlossenen elementaren Bereich D auf G liegt und eine stückweise stetig differenzierbare Parameterdarstellung $\mathbf{x} : [a, b] \rightarrow G$ besitzt. C sei im mathematisch positiven Sinn durchlaufen. Dann gilt*

$$\int_C \mathbf{v} \, d\mathbf{x} = \iint_D \text{rot } \mathbf{v} \, dx dy = \iint_D \underbrace{\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right)}_{= \text{rot } \mathbf{v}} \, dx dy. \quad (4.1)$$

Bevor wir auf Folgerungen eingehen, geben wir eine Begründung von (4.1), wobei wir uns auf den Fall beschränken, dass C von jeder achsenparallelen Geraden in einem Intervall oder in zwei Punkten geschnitten wird. Da jeder elementare Bereich D in solche Bereiche zerlegt werden kann, und Kurvenintegrale über entgegengesetzt durchlaufene Kurvenstücke entgegengesetzt gleich sind, folgt daraus der allgemeine Satz.

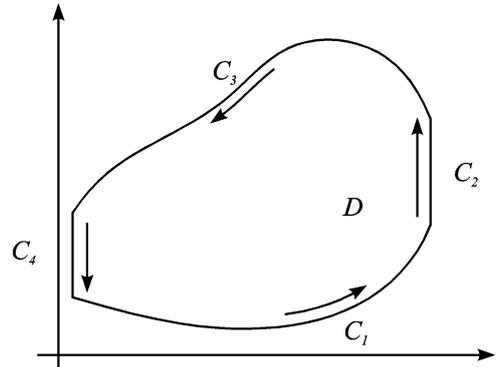
Wir berechnen

$$\int_C \mathbf{v} \, d\mathbf{x} = \int_C \begin{pmatrix} u \\ v \end{pmatrix} d\mathbf{x} = \int_C \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} + \int_C \begin{pmatrix} 0 \\ v \end{pmatrix} d\mathbf{x}.$$

Dazu zerlegen wir C in vier Teilkurven C_1, C_2, C_3 und C_4 und wählen folgende stückweise stetig differenzierbare Parameterdarstellungen:

4 Integralrechnung in mehreren Variablen

$$\begin{aligned}
 C_1 : \mathbf{x}(s) &= \begin{pmatrix} s \\ \gamma(s) \end{pmatrix}, & a \leq s \leq b, \\
 C_2 : \mathbf{x}(s) &= \begin{pmatrix} b \\ s \end{pmatrix}, & \gamma(b) \leq s \leq \delta(b), \\
 -C_3 : \mathbf{x}(s) &= \begin{pmatrix} s \\ \delta(s) \end{pmatrix}, & a \leq s \leq b, \\
 -C_4 : \mathbf{x}(s) &= \begin{pmatrix} a \\ s \end{pmatrix}, & \gamma(a) \leq s \leq \delta(a).
 \end{aligned}$$



Damit erhalten wir

$$\begin{aligned}
 \int_C \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} &= \int_{C_1} \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} + \int_{C_2} \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} + \int_{C_3} \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} + \int_{C_4} \begin{pmatrix} u \\ 0 \end{pmatrix} d\mathbf{x} \\
 &= \int_a^b u(s, \gamma(s)) ds + 0 - \int_a^b u(s, \delta(s)) ds - 0 = \\
 &= - \int_a^b \left(\int_{\gamma(s)}^{\delta(s)} \frac{\partial u}{\partial y}(s, t) dt \right) ds = - \iint_D \frac{\partial u}{\partial y} dx dy
 \end{aligned}$$

und

$$\int_C \begin{pmatrix} 0 \\ v \end{pmatrix} d\mathbf{x} = \int_a^b v(s, \gamma(s)) \gamma'(s) - v(s, \delta(s)) \delta'(s) ds + \int_{\gamma(b)}^{\delta(b)} v(b, s) ds - \int_{\gamma(a)}^{\delta(a)} v(a, s) ds.$$

Nach der Leibnizschen Regel für die Differentiation von Parameterintegralen gilt aber auch

$$\begin{aligned}
 \iint_D \frac{\partial v}{\partial x} dx dy &= \int_a^b \int_{\gamma(s)}^{\delta(s)} \frac{\partial v}{\partial s}(s, y) dy ds \\
 &= \int_a^b \frac{d}{ds} \int_{\gamma(s)}^{\delta(s)} v(s, y) dy ds + \int_a^b v(s, \gamma(s)) \gamma'(s) - v(s, \delta(s)) \delta'(s) ds
 \end{aligned}$$

und daher

$$\int_C \begin{pmatrix} 0 \\ v \end{pmatrix} d\mathbf{x} = \iint_D \frac{\partial v}{\partial x} dx dy.$$

Addition liefert die gewünschte Aussage.

Beispiel. Es sei C die positiv durchlaufene Einheitskreislinie. Nach dem Integralsatz von Green gilt dann

$$\int_C \begin{pmatrix} xy \\ x \end{pmatrix} d\mathbf{x} = \iint_{x^2+y^2 \leq 1} (1-x) dx dy = \pi.$$

Übungsaufgabe 260. (T) Berechnen Sie das Kurvenintegral $\int_C \mathbf{V}(\mathbf{x}(s)) ds$, wobei

$$\mathbf{V}(x, y) = \begin{pmatrix} x - y \\ 1 - x \end{pmatrix}$$

und C der im positiven Sinne durchlaufene Umfang des Dreiecks mit den Eckpunkten $(1, 0)$, $(0, 1)$ und $(0, 0)$ ist.

Übungsaufgabe 261. (T) Bestätigen Sie den Satz von Green für das Vektorfeld

$$\mathbf{V}(x, y) = \begin{pmatrix} x^2 y^2 \\ y \end{pmatrix}$$

und die Kurve C , welche der im positiven Sinne durchlaufene Rand des Halbkreises mit Mittelpunkt $(0, 0)$, Radius 2 und nichtnegativer y -Koordinate ist.

Übungsaufgabe 262. (E) Formulieren und überprüfen Sie den Satz von Green für Kreise mit Mittelpunkt im Ursprung!

Hinweis: Es ist hilfreich Polarkoordinaten zu verwenden und dann die Vorgehensweise für Rechtecke zu imitieren!

Ergänzungen und Deutung

Ist die Parameterdarstellung \mathbf{x} von C so beschaffen, dass $\|\dot{\mathbf{x}}(t)\| = 1$ ist, für alle $t \in [a, b]$, dann ist

$$\int_a^b \mathbf{v}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) dt,$$

also gleich dem Integral über die Tangentialkomponente von \mathbf{v} längs C , d.h. das Integral ist die **Zirkulation** von \mathbf{v} längs C . Nach (4.1) und dem Mittelwertsatz für Doppelintegrale gilt also

$$\text{Zirkulation} = \iint_D \text{rot } \mathbf{v} dx dy = |D| \text{rot } \mathbf{v}(\mathbf{a}), \quad \mathbf{a} \in D \text{ passend.}$$

4 Integralrechnung in mehreren Variablen

(Hier wie an einigen anderen Stellen, ist $|D|$ eine alternative Notation für das Flächenmaß $\lambda_2(D)$ von D . Eine andere ist $I(D)$.) Die Rotation von \mathbf{v} kann also als Zirkulation pro Flächeneinheit oder als **Zirkulationsdichte** gedeutet werden.

Wendet man den Integralsatz von Green auf das Vektorfeld

$$\begin{pmatrix} -v \\ u \end{pmatrix}$$

an, erhält man statt (4.1) die folgende Version des **Satzes von Green**:

Satz 4.4.5.2. *Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^2$ ein Vektorfeld mit stetigen partiellen Ableitungen und C eine einfach geschlossene Kurve, die samt dem von ihr berandeten, beschränkten und abgeschlossenen elementaren Bereich D auf G liegt und eine stückweise stetig differenzierbare Parameterdarstellung $\mathbf{x} : [a, b] \rightarrow G$, $t \mapsto (x(t), y(t))$ besitzt. C sei im mathematisch positiven Sinn durchlaufen. Dann gilt*

$$\int_a^b \mathbf{v}(\mathbf{x}(t)) \cdot \mathbf{n}(t) dt = \iint_D \underbrace{\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)}_{\text{div } \mathbf{v}} dx dy, \quad (4.2)$$

wobei $\mathbf{n}(t) = (\dot{y}(t), -\dot{x}(t))^T$ den Normalvektor an die Kurve C im Punkt $\mathbf{x}(t)$ bezeichnet.

Ist die Parameterdarstellung so, dass $\|\dot{\mathbf{x}}(t)\| = 1$ ist für alle $t \in [a, b]$, dann ist auch $\|\mathbf{n}(t)\| = 1$. Da wir annehmen, dass C im positiven Sinne durchlaufen wird, ist $\mathbf{n}(t)$ der nach außen weisende Normaleneinheitsvektor von C und

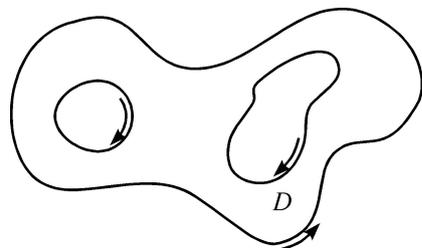
$$\int_a^b \mathbf{v}(\mathbf{x}(t)) \cdot \mathbf{n}(t) dt$$

ist dann das Integral über die Normalkomponente von \mathbf{v} längs C , d.h. der sogenannten **Fluss** von \mathbf{v} durch C . Damit gilt nach (4.2) und dem Mittelwertsatz für Doppelintegrale

$$\text{Fluss} = \iint_D \text{div } \mathbf{v} dx dy = |D| \text{div } \mathbf{v}(\mathbf{a}), \quad \mathbf{a} \in D \text{ passend.}$$

Die Divergenz von \mathbf{v} ist also der Fluss pro Flächeneinheit, oft auch **Quelldichte** genannt.

Zum Schluss merken wir an, dass der Integralsatz von Green auch für sehr viel allgemeinere Bereiche gilt, etwa für solche der folgenden Form:



Anwendungen:

Leibnizsche Sektorformel

Wendet man den Integralsatz von Green auf das Vektorfeld $\mathbf{v} = (y, -x)^T$ an, so folgt

$$\int_C \mathbf{v} d\mathbf{x} = \int_a^b (y(t) \dot{x}(t) - x(t) \dot{y}(t)) dt = \iint_D (-1 - 1) dx dy = -2|D| = -2\lambda_2(D).$$

Wir erhalten also die Leibnizsche Sektorformel

$$\lambda_2(D) = |D| = \frac{1}{2} \left| \int_a^b (\dot{x} y - x \dot{y}) dt \right|.$$

Integrabilitätsbedingung

Wir nehmen an, dass das Vektorfeld \mathbf{v} mit stetigen partiellen Ableitungen die Eigenschaft hat, wegunabhängige Kurvenintegrale zu besitzen. Dann gilt für jede einfach geschlossene Kurve C , die samt ihrem Inneren auf G liegt,

$$\int_C \mathbf{v} d\mathbf{x} = 0. \quad (4.3)$$

Also ist nach dem Integralsatz von Green

$$\int_C \mathbf{v} d\mathbf{x} = \iint_D (v_x - u_y) dx dy = 0, \quad (4.4)$$

wobei D der abgeschlossene beschränkte Bereich mit Rand C ist. Wir zeigen, dass daraus die Integrabilitätsbedingung

$$v_x - u_y = 0 \quad (4.5)$$

folgt. Wäre an einer Stelle $\mathbf{a} \in G$ nämlich $v_x(\mathbf{a}) - u_y(\mathbf{a}) \neq 0$, etwa > 0 , so folgt aus der Stetigkeit von v_x, u_y auch $v_x(\mathbf{x}) - u_y(\mathbf{x}) > 0$ in einer Umgebung von \mathbf{a} . Nun legt man in diese Umgebung eine einfach geschlossene Kurve C mit den obigen Eigenschaften. Dann ist die rechte Seite in (4.4) positiv, während die linke nach (4.3) verschwindet. Widerspruch. Damit gilt (4.5).

Laplaceoperator und Normalableitung

Es seien G, C, D wie oben, für die Parameterdarstellung \mathbf{x} von C gelte aber $\|\dot{\mathbf{x}}(t)\| = 1$ für $t \in [a, b]$. Ferner sei $f : G \rightarrow \mathbb{R}$ ein Skalarfeld mit stetigen 2. partiellen Ableitungen. Dann gilt

$$\begin{aligned} \iint_D \Delta f dx dy &= \iint_D (f_{xx} + f_{yy}) dx dy = \int_C \begin{pmatrix} -f_y \\ f_x \end{pmatrix} d\mathbf{x} = \int_a^b \begin{pmatrix} -f_y(\mathbf{x}(t)) \\ f_x(\mathbf{x}(t)) \end{pmatrix} \cdot \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} dt \\ &= \int_a^b \begin{pmatrix} f_x(\mathbf{x}(t)) \\ f_y(\mathbf{x}(t)) \end{pmatrix} \begin{pmatrix} \dot{y}(t) \\ -\dot{x}(t) \end{pmatrix} dt = \int_a^b \frac{\partial f}{\partial \mathbf{n}}(\mathbf{x}(t)) dt. \end{aligned}$$

Greensche Formeln

Es seien G, C, D wie oben und $\mathbf{v} : G \rightarrow \mathbb{R}^2$ ein Vektorfeld mit stetigen partiellen Ableitungen bis zur 2. Ordnung. Wie eben folgt aus

$$(uv_x)_x = u_x v_x + uv_{xx}, \quad (uv_y)_y = u_y v_y + uv_{yy}$$

die Beziehung

$$\iint_D (u_x v_x + uv_{xx} + u_y v_y + uv_{yy}) \, dx dy = \int_C u \begin{pmatrix} -v_y \\ v_x \end{pmatrix} d\mathbf{x},$$

oder

$$\iint_D \text{grad } u \cdot \text{grad } v \, dx dy + \iint_D u \Delta v \, dx dy = \int_a^b u \frac{\partial v}{\partial \mathbf{n}} dt,$$

die **erste Greensche Formel**. Vertauscht man u mit v , erhält man eine analoge Formel und Subtraktion liefert die **zweite Greensche Formel**

$$\iint_D (u \Delta v - v \Delta u) \, dx dy = \int_a^b \left(u \frac{\partial v}{\partial \mathbf{n}} - v \frac{\partial u}{\partial \mathbf{n}} \right) dt.$$

4.4.6 Integralsatz von Stokes

Inhalt in Kurzfassung: Der Satz von Stokes ist das dreidimensionale Analogon der Rotationsvariante der Satzes von Green und besagt, dass das Kurvenintegral über ein dreidimensionales stetig differenzierbares Vektorfeld entlang einer stückweise stetig differenzierbaren Kurve C , die ein Flächenstück F im Raum berandet, übereinstimmt mit dem Oberflächenintegral über $\text{rot } \mathbf{v}$. (Das Vorzeichen ist im Sinne der Rechtsschraubenregel zu bilden.)

Es sei G ein Gebiet im \mathbb{R}^3 und D ein elementarer Bereich im \mathbb{R}^2 . Der **Integralsatz von Stokes** lautet:

Satz 4.4.6.1. *Sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein Vektorfeld mit stetigen partiellen Ableitungen, $F \subset G$ ein Flächenstück mit stetig differenzierbarer Parameterdarstellung und C die Randkurve von F mit stückweise stetig differenzierbarer Parameterdarstellung. Wird C so durchlaufen, dass F „links“ liegt und der Umlaufsinn der Kurve zusammen mit der Normalenrichtung von F eine Rechtsschraubung ergibt, dann gilt*

$$\int_C \mathbf{v} \, d\mathbf{x} = \iint_F \text{rot } \mathbf{v} \, d\mathbf{O}. \quad (4.6)$$

Deutung

Ist \mathbf{v} das Geschwindigkeitsfeld einer stationären Strömung einer inkompressiblen Flüssigkeit in G , dann ist das linke Integral in (4.6) die Zirkulation längs C und das rechte etwa

$|F|$ mal der Normalkomponente von $\operatorname{rot} \mathbf{v}$ auf F . D.h. die Normalkomponente von $\operatorname{rot} \mathbf{v}$ könnte man als Zirkulationsdichte der Flüssigkeitsströmung in F deuten.

Ist \mathbf{v} ein Kraftfeld, so stellt das linke Integral in (4.6) die Arbeit dar, die geleistet wird, wenn man einen der Kraft unterworfenen Punkt um C herum bewegt. Nach Stokes läßt sich also diese Arbeit als das Integral über F der Normalkomponente der Rotation ausdrücken.

Anwendung:

Integrabilitätsbedingungen

Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein Vektorfeld mit stetigen partiellen Ableitungen und wegunabhängigen Kurvenintegralen. Dann ist die Integrabilitätsbedingung

$$\operatorname{rot} \mathbf{v} = \mathbf{o}$$

erfüllt. Wäre an einer Stelle $\mathbf{a} \in G$ nämlich $\operatorname{rot} \mathbf{v}(\mathbf{a}) \neq \mathbf{o}$, so könnte man ein Flächenstück F durch \mathbf{a} finden mit

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) > 0 \text{ für } \mathbf{x} \in F,$$

in Widerspruch zum Satz von Stokes.

Übungsaufgabe 263. (T) Eine turbulente Strömung besitzt das Geschwindigkeitsfeld

$$\vec{v}(x, y, z) = (x^2 y^2, -z, y^2 + \frac{1}{x-z})^T.$$

Berechnen Sie die Wirbelstärke $\operatorname{rot} \vec{v}$, sowie das Kurvenintegral von \vec{v} entlang der Kurve C mit

$$(x(t), y(t), z(t)) = (3 \cos t, 3 \sin t, 3 \cos t - 1)^T, \quad t \in [0, 2\pi]. \quad (4.7)$$

Übungsaufgabe 264. (T) Gegeben ist wieder das Geschwindigkeitsfeld

$$\vec{v}(x, y, z) = (x^2 y^2, -z, y^2 + \frac{1}{x-z})^T. \quad (4.8)$$

Die Kurve C mit

$$(x(t), y(t), z(t)) = (3 \cos t, 3 \sin t, 3 \cos t - 1)^T, \quad t \in [0, 2\pi]. \quad (4.9)$$

entsteht als Rand der Schnittfläche des Zylinders Z um die z -Achse mit Radius $r \leq 3$ mit der Ebene $E : x - z = 1$. Bestätigen Sie den Integralsatz von Stokes indem Sie das Oberflächenintegral von \vec{v} über die Schnittfläche ausführen.

Übungsaufgabe 265. (T) Berechnen Sie das Integral $\int_C \mathbf{V}(\mathbf{x}(s)) ds$ wobei

$$\mathbf{V}(x, y, z) = \begin{pmatrix} z - y \\ x - z \\ y - x \end{pmatrix}$$

und C der im positiven Sinne durchlaufene Rand des Dreiecks mit den Eckpunkten $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$.

4.4.7 Integralsatz von Gauß

Inhalt in Kurzfassung: So wie der Satz von Stokes als Übertragung der Rotationsversion des Satzes von Green von zwei in drei Dimensionen verstanden werden kann, ist der Satz von Gauß die dreidimensionale Variante der Divergenzversion. Er besagt, dass das Oberflächenintegral über ein stetig differenzierbares Vektorfeld \mathbf{v} über die Oberfläche eines dreidimensionalen Bereich mit dem Bereichsintegral über $\operatorname{div} \mathbf{v}$ übereinstimmt.

Es sei G ein Gebiet im \mathbb{R}^3 . Der **Integralsatz von Gauß**, genannt auch **Divergenzsatz** (vor allem in der englischsprachigen Literatur) lautet:

Satz 4.4.7.1. *Sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein Vektorfeld mit stetigen partiellen Ableitungen und F eine geschlossene Fläche in G , die samt ihrem Innenbereich einen beschränkten, abgeschlossenen, elementaren Bereich D in G bildet und sich aus endlich vielen Flächenstücken zusammensetzt mit stetig differenzierbaren Parameterdarstellungen. Weisen die Normalenrichtungen an alle diese Flächenstücke nach außen, dann gilt*

$$\iint_F \mathbf{v} \, d\mathbf{O} = \iiint_D \operatorname{div} \mathbf{v} \, dx \, dy \, dz. \quad (4.10)$$

Der Beweis verläuft ähnlich wie der Beweis des Integralsatzes von Green im \mathbb{R}^2 , die technischen Einzelheiten sind aber schwieriger.

Physikalische Deutung

Die hier besprochene Deutung des Gaußschen Integralsatzes wird auch die frühere Deutung der Divergenz klar machen. Es sei \mathbf{v} das stationäre Geschwindigkeitsfeld der Strömung einer inkompressiblen Flüssigkeit. Dann ist das linke Integral in (4.10) die Differenz von ausströmender und einströmender Flüssigkeit pro Zeiteinheit. (Wir nehmen an, dass die Flächennormalvektoren $\mathbf{x}_u \times \mathbf{x}_v$ nach außen weisen.) Da die Flüssigkeit inkompressibel ist, darf es zu keiner Vermehrung oder Verminderung von Flüssigkeit kommen. Es muss daher in D Flüssigkeit erzeugt oder vernichtet werden. Man spricht von **Quellen** und **Senken**. Der Überschuss von ausströmender und einströmender Flüssigkeit heißt **Gesamtergiebigkeit**. Sie ist positiv oder negativ, je nachdem die Quellen oder Senken überwiegen. Die Gesamtergiebigkeit von D durch das Volumen von D nennt man **mittlere Ergiebigkeit**. Zieht man D auf einen Punkt \mathbf{x} zusammen, erhält man im Grenzwert die **spezifische Ergiebigkeit** in \mathbf{x} oder die **Quelldichte**. Aus dem Mittelwertsatz für Volumsintegrale folgt

$$\operatorname{div} \mathbf{v}(\mathbf{x}) = \text{Quelldichte in } \mathbf{x}.$$

Übungsaufgabe 266. (T) *Bestätigen Sie den Satz von Gauß am Beispiel des Kegels der Höhe $H = 1$ um die (positive) z -Achse mit Spitze im Ursprung und Öffnungswinkel 45° und des Vektorfelds*

$$\mathbf{V}(x, y, z) = \begin{pmatrix} x^2 + z \\ y + z^2 \\ y^2 + x \end{pmatrix}$$

Übungsaufgabe 267. (T) Gegeben ist das Tetraeder T mit Eckpunkten $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ und $(0, 0, 1)$ sowie das Vektorfeld

$$\mathbf{V}(x, y, z) := \begin{pmatrix} 2y \\ 2x \\ xyz \end{pmatrix}.$$

Berechnen Sie das Oberflächenintegral $\iint_{\partial T} \mathbf{V} d\mathbf{O}$.

Übungsaufgabe 268. (T*) Wie die vorige Aufgabe, aber für die vier einzelnen Seitenflächen.

Übungsaufgabe 269. (T) Gegeben ist die Halbkugel (östliche Hemisphäre) $H := \{(x, y, z) : x^2 + y^2 + z^2 \leq 1, y \geq 0\}$ sowie das Vektorfeld

$$\mathbf{V}(x, y, z) := \begin{pmatrix} x \\ y \\ z^2 \end{pmatrix}.$$

Berechnen Sie das Oberflächenintegral $\iint_{\partial H} \mathbf{V} d\mathbf{O}$.

Übungsaufgabe 270. (T) Wie die vorige Aufgabe, aber für die Halbsphäre (d.h. ohne Deckfläche).

Übungsaufgabe 271. (T*) Gegeben sind zwei Kugeln K_1 und K_2 (Mittelpunkt im Ursprung) mit Radien $R_1 < R_2$. Es bezeichne K_1^2 jenes Gebiet (Kugelschale), welches aus Punkten besteht, die zwischen beiden Kugeln liegen, d.h. jene \mathbf{x} mit $R_1 < \|\mathbf{x}\| < R_2$. Berechnen Sie

1. $\int \partial K_1^2 \frac{\mathbf{x}}{\|\mathbf{x}\|^3} d\mathbf{O}$ mithilfe des Gauß'schen Satzes.

Wenn Sie motiviert sind, versuchen Sie bitte sich freiwillig an folgender physikalischer Anwendung dieser Beobachtung:

- 2.* Das von einer in \mathbf{y} angebrachten Punktladung der Stärke σ in \mathbf{x} hervorgerufene Potential ist $V(\mathbf{x}) = 8.99 \times 10^9 \frac{\sigma}{\|\mathbf{x}-\mathbf{y}\|}$. Das Potential einer mit Ladungsdichte σ belegten Hohlkugel ∂K ist dementsprechend

$$V(\mathbf{x}) = 8.99 \times 10^9 \int_{\partial K} \frac{\sigma}{\|\mathbf{x}-\mathbf{y}\|} dO(\mathbf{y}).$$

Berechnen Sie $V(\mathbf{x})$ für Punkte \mathbf{x} innerhalb der Hohlkugel!

Ergebnis: $V(\mathbf{x}) = -8.99 \times 10^9 \times 4\pi\sigma R$ im Inneren der Hohlkugel, d.h. dieser Bereich ist wegen $F = -\nabla V$ kräftefrei!

4 Integralrechnung in mehreren Variablen

- 3.* Das von einer Hohlkugel ∂K mit Radius R und Flächendichte σ hervorgerufene Gravitationspotential ist

$$V_{\text{hohl}}(\mathbf{x}) = -6.67 \times 10^{-11} \int_{\partial K} \frac{\sigma}{\|\mathbf{x} - \mathbf{y}\|^2} dO(\mathbf{y}),$$

das von einer Vollkugel K mit Radius R und Massendichte ρ hervorgerufene Potential

$$V_{\text{voll}}(\mathbf{x}) = -6.67 \times 10^{-11} \int_K \frac{\rho}{\|\mathbf{x} - \mathbf{y}\|^2} d\mathbf{y},$$

und das von einer im Ursprung sitzenden Punktmasse M hervorgerufene Potential ist

$$V_{\text{punkt}}(\mathbf{x}) = -6.67 \times 10^{-11} \frac{M}{\|\mathbf{x} - \mathbf{y}\|^2}.$$

Wenn die Massen der Kugeln bzw. der Punktmasse gleich sind, so stimmen für \mathbf{x} mit $\|\mathbf{x}\| > R$ auch die Potentiale $V_{\text{hohl}}(\mathbf{x})$, $V_{\text{voll}}(\mathbf{x})$ und $V_{\text{punkt}}(\mathbf{x})$ überein.

4.4.8 Ergänzungen zu den Integralsätzen

Inhalt in Kurzfassung: Wieder folgen Ergänzungen und Illustrationen zu diesem Abschnitt, die nicht Prüfungsstoff sind, deren Lektüre aber das Verständnis des Stoffes festigen kann.

Für manche Anwendungen ist folgendes einfache Resultat nützlich. (Ähnliche Ergebnisse spielen in zahlreichen Teilen der Analysis eine Rolle.)

Satz (Eindeutigkeit für Oberflächenintegrale). *Es sei $\mathbf{v} : G \rightarrow \mathbb{R}^3$ ein Vektorfeld mit stetigen partiellen Ableitungen. Ist das Oberflächenintegral über jede geschlossene Fläche F (wie im Integralsatz beschrieben) null, so gilt*

$$\operatorname{div} \mathbf{v} = 0 \text{ auf } G.$$

Greensche Formeln

Es seien G, F, D wie oben. Ferner haben $u, v : G \rightarrow \mathbb{R}$ stetige partielle Ableitungen bis zur 2. Ordnung. Wir wenden den Gaußschen Integralsatz auf das Vektorfeld

$$\mathbf{v} = u \operatorname{grad} v$$

an. Beachtet man die Identität

$$\operatorname{div} (u \operatorname{grad} v) = \operatorname{grad} u \cdot \operatorname{grad} v + u \Delta v,$$

so folgt die **erste Greensche Formel**

$$\iiint_D (\text{grad } u \cdot \text{grad } v + u \Delta v) \, dx dy dz = \iint_F u \text{grad } v \, d\mathbf{O} = \iint_F u \frac{\partial v}{\partial \mathbf{n}} \, dO.$$

Vertauscht man u und v und subtrahiert, so erhält man die **zweite Greensche Formel**

$$\iiint_D (u \Delta v - v \Delta u) \, dx dy dz = \iint_F \left(u \frac{\partial v}{\partial \mathbf{n}} - v \frac{\partial u}{\partial \mathbf{n}} \right) \, dO.$$

Die Wärmeleitungsgleichung:

Es sei $T(\mathbf{x}, t)$ die Temperatur in einem Körper K im Punkt \mathbf{x} zum Zeitpunkt t . Dann gilt für den Wärmefluss \mathbf{w}

$$\mathbf{w}(\mathbf{x}, t) = -l \text{grad } T \quad (l > 0, \text{ fest}). \quad (4.11)$$

Wir nehmen dabei an, dass der Wärmeleitungskoeffizient l konstant ist. Es sei nun D ein elementarer Bereich im Inneren von K , der von einer geschlossenen Fläche F berandet wird (so dass die Voraussetzungen des Gaußschen Integralsatzes erfüllt sind). Die durch F pro Zeiteinheit ausströmende (Netto-)Wärmemenge ist einerseits gleich

$$\iint_F \mathbf{w} \, d\mathbf{O}. \quad (4.12)$$

Beachtet man andererseits, dass die Gesamtwärmemenge in D gleich ist

$$\iiint_D \sigma \varrho T \, dx dy dz, \quad (\sigma > 0 \text{ spezifische Wärme, } \varrho > 0 \text{ Dichte})$$

so ist die durch F pro Zeiteinheit ausströmende Wärmemenge

$$-\frac{d}{dt} \iiint_D \sigma \varrho T \, dx dy dz. \quad (4.13)$$

Die Ausdrücke (4.12) und (4.13) müssen also übereinstimmen. Aus dem Gaußschen Integralsatz, (4.11) und der Leibnizschen Regel für die Differentiation von Parameterintegralen erhält man

$$\begin{aligned} \iint_F \mathbf{w} \, d\mathbf{O} &= -l \iiint_D \text{div grad } T \, dx dy dz = -l \iiint_D \Delta T \, dx dy dz \\ &= -\frac{d}{dt} \iiint_D \sigma \varrho T \, dx dy dz = -\iiint_D \sigma \varrho \frac{\partial T}{\partial t} \, dx dy dz \end{aligned}$$

und damit

$$\iiint_D \left(\frac{\partial T}{\partial t} - \frac{l}{\sigma \varrho} \Delta T \right) \, dx dy dz = 0.$$

4 Integralrechnung in mehreren Variablen

Das gilt für alle Bereiche D im Inneren von K , also erhalten wir die Wärmeleitungsgleichung

$$\frac{\partial T}{\partial t} = \frac{l}{\sigma \varrho} \Delta T. \quad (4.14)$$

Die Kontinuitätsgleichung der Hydrodynamik:

Es sei $\mathbf{v}(\mathbf{x}, t)$ die Geschwindigkeit einer Flüssigkeitsströmung ohne Quellen und Senken und $\varrho(\mathbf{x}, t)$ die Dichte an der Stelle $\mathbf{x} \in G$ zum Zeitpunkt t . Es seien D, F wie im Gaußschen Integralsatz. Die Änderungsrate der Flüssigkeitsmenge in D ist dann einerseits gegeben durch

$$\frac{d}{dt} \iiint_D \varrho(\mathbf{x}, t) dx dy dz = \iiint_D \frac{\partial \varrho}{\partial t} dx dy dz$$

und andererseits durch

$$- \iint_F \varrho \mathbf{v} d\mathbf{O} = - \iiint_D \operatorname{div}(\varrho \mathbf{v}) dx dy dz.$$

Also ist

$$\iiint_D \left(\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) \right) dx dy dz = 0.$$

Da D beliebig war, erhalten wir die Kontinuitätsgleichung

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(\varrho \mathbf{v}) = 0.$$

Im Spezialfall einer inkompressiblen Flüssigkeit, wenn also $\varrho = \text{const}$ ist, vereinfacht sie sich zu

$$\operatorname{div} \mathbf{v} = 0.$$

5 Stochastik

Die ursprüngliche Bedeutung des altgriechischen Wortes, aus dem sich unser Wort „Stochastik“ ableitet, ist in etwa „Kunst des Ratens“. Nach heutigem Verständnis umfasst die Stochastik jene Teilgebiete der Mathematik, die sich mit Wahrscheinlichkeit und Zufall befassen, also Wahrscheinlichkeitstheorie und Statistik. Der begriffliche Apparat dafür ist durchaus beachtlich. Entsprechend umfangreich ist 5.1, der erste Abschnitt dieses Kapitels, der sich mit den Grundbegriffen beschäftigt. Genannt seien an dieser Stelle: Ereignisse, Wahrscheinlichkeitsräume und -verteilungen, Verteilungs- und Dichtefunktionen, Zufallsgrößen, Erwartungswert und Varianz, bedingte Wahrscheinlichkeit und Unabhängigkeit von Ereignissen und Zufallsgrößen. Erst in 5.2 kommen wir zu tiefer liegenden Ergebnissen. Hervorgehoben seien vor allem das Gesetz der großen Zahlen und der Zentrale Grenzwertsatz. An diesen theoretischen Resultaten setzt dann die Statistik in 5.3 an. Ihr Ziel ist es, aus empirischen Daten Rückschlüsse auf Wahrscheinlichkeitsverteilungen zu ziehen. Ihre Methoden, von denen wir Punktschätzer, Konfidenzintervalle und das Testen von Hypothesen besprechen werden, fußen auf der Wahrscheinlichkeitstheorie der ersten beiden Abschnitte dieses Kapitels.

5.1 Grundbegriffe

In diesem recht langen Abschnitt werden Grundbegriffe der Stochastik entwickelt. Dazu muss beachtliche begriffliche Arbeit geleistet, aber noch relativ wenig gerechnet und bewiesen werden. Die Grenze zum nächsten Abschnitt mit den Hauptsätzen der Stochastik wird dann gezogen, wenn es nicht mehr primär um die Prägung der richtigen Begriffe geht, sondern sobald nennenswerte Rechnungen und Beweise erforderlich werden. Die Struktur des Abschnitts erwächst aus dem Anliegen, die zahlreichen Begriffe passend aufeinander aufzubauen. Das bedingt die folgende Reihenfolge.

Wir beginnen in 5.1.1 mit der Abstraktion einer naiven Vorstellung der Wahrscheinlichkeit von Ereignissen und abstrahieren zu den Begriffen Wahrscheinlichkeitsmaß und Wahrscheinlichkeitsraum. Einige besonders typische Beispiele dazu werden in 5.1.2 besprochen. In 5.1.3 führt die Unterscheidung diskreter und reeller Wahrscheinlichkeitsräume bzw. Verteilungen u.a. zu den Begriffen Verteilungsfunktion und Dichte. Von größtem Interesse sind in der Stochastik Zufallsgrößen und ihre Verteilungen. Sie werden in 5.1.4 eingeführt. Ihre wichtigsten Kenngrößen sind Erwartungswert (5.1.5) und Varianz(5.1.6). Ein weiteres Standbein für alles Weitere bilden die Begriffe der bedingten Wahrscheinlichkeit und noch mehr der Unabhängigkeit von Ereignissen oder auch von Zufallsgrößen (5.1.7). Eine elementare, aber lehrreiche Aussage über bedingte Wahrscheinlichkeiten ist der Satz von Bayes, dem 5.1.8 gewidmet ist.

5.1.1 Wahrscheinlichkeit und Wahrscheinlichkeitsräume

Inhalt in Kurzfassung: Ungeachtet der Schwierigkeit, den Begriff der Wahrscheinlichkeit philosophisch überzeugend zu fassen, gelingt ein mathematisch sehr befriedigender Aufbau der Wahrscheinlichkeitstheorie mit dem von Kolmogorow stammenden Ansatz. Er basiert auf der Maßtheorie, wie sie in Mathematik 1 skizziert wurde. Als Trägermenge fungiert die Menge Ω aller sogenannten Elementarereignisse. Gewissen Teilmengen A davon (sie bilden laut Definition des Maßbegriffs eine σ -Algebra) weist eine Funktion \mathbb{W} ihre Wahrscheinlichkeit $\mathbb{W}(A)$ (eine Zahl aus $[0, 1]$) derart zu, dass \mathbb{W} ein zu $\mathbb{W}(\Omega) = 1$ normiertes Maß ist.

Philosophisch ist der Begriff der Wahrscheinlichkeit ein äußerst schwieriger. Entsprechend tauchten im Lauf der Geschichte die unterschiedlichsten Definitionen auf. Auch die Mathematik erhebt nicht den Anspruch, definieren zu können, was Wahrscheinlichkeit ihrem innersten Wesen nach sei. Umso beeindruckender ist es, dass der sehr einfache Ansatz des russischen Mathematikers Andrei Nikolajewitsch Kolmogorow (1903-1987), die Wahrscheinlichkeitstheorie und damit die gesamte Stochastik auf der Maßtheorie aufzubauen, kaum Wünsche offen lässt. Und zwar erweist es sich für eine mathematisch zufriedenstellende Theorie als vollkommen ausreichend festzuhalten, wie man mit Wahrscheinlichkeiten rechnet. Dieser Ansatz und die darauf aufbauende Theorie soll nun in den Grundzügen entwickelt werden. Dabei kürzen wir das Wort „Wahrscheinlichkeit“ insbesondere in längeren Wortverbindungen oft einfach mit dem Buchstaben \mathbb{W} ab.

In einem gewissen Kontext mit Ungewissheit sei Ω die Menge aller denkbaren Ereignisse, der sogenannten **Elementarereignisse**. Ein Elementarereignis zeichnet sich gegenüber anderen Ereignissen dadurch aus, dass es sich nicht mehr sinnvoll in kleinere Ereignisse zerlegen lässt. Im Gegensatz dazu fassen wir ein beliebiges (nicht notwendig elementares) Ereignis als eine Menge von Elementarereignissen, also als eine Teilmenge A von Ω auf. A kann leer sein, aus einem Element bestehen oder auch aus (eventuell unendlich) vielen. Ohne den Zahlenwert der Wahrscheinlichkeit für A kennen zu müssen, schreiben wir dafür $\mathbb{W}(A)$. Jedenfalls handelt es sich um eine reelle Zahl zwischen 0 und 1. Ereignisse mit Wahrscheinlichkeit 1 heißen **fast sicher**. (Weniger gebräuchlich ist es, Ereignisse mit Wahrscheinlichkeit 0 **fast unmöglich** zu nennen.) Die Wahrscheinlichkeit \mathbb{W} ist also eine Abbildung mit $\mathbb{W} : A \mapsto \mathbb{W}(A)$ für gewisse $A \subseteq \Omega$. In manchen Fällen sind alle $A \subseteq \Omega$ zugelassen. Dann ist \mathbb{W} auf der gesamten Potenzmenge $\mathcal{P}(\Omega)$, der Menge *aller* Teilmengen von Ω definiert. In anderen Fällen muss man sich auf ein echtes Teilsystem $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ beschränken. Die Elemente von \mathcal{A} heißen auch die **messbaren** Ereignisse.

Wir haben damit geklärt, dass (in einem gegebenen Kontext) die Wahrscheinlichkeit eine Abbildung $\mathbb{W} : \mathcal{A} \rightarrow [0, 1]$ mit $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ ist, wobei Ω die Menge der Elementarereignisse bezeichnet. Somit geht es nur noch darum, welche Eigenschaften man von einer Funktion $\mathbb{W} : \mathcal{A} \rightarrow [0, 1]$ verlangen soll, damit man sinnvoll von Wahrscheinlichkeit sprechen kann. Sicher wird man $\mathbb{W}(\emptyset) = 0$ fordern, weil die leere Menge \emptyset überhaupt kein Ereignis enthält, also sicher kein Element von \emptyset eintreten wird. Insbesondere verlangt

man damit auch $\emptyset \in \mathcal{A}$. Analog verlangt man $\mathbb{W}(\Omega) = 1$ (und folglich $\Omega \in \mathcal{A}$), weil mit Sicherheit, insbesondere also mit Wahrscheinlichkeit 1, irgendeines der denkbaren Ereignisse und somit ein Element von Ω eintreten wird. Neben diesen beiden (sehr trivialen) Forderungen ist die entscheidende, aber ebenfalls selbstverständliche Eigenschaft, die man von Wahrscheinlichkeit verlangt, die Additivität:

$$\mathbb{W}(A \cup B) = \mathbb{W}(A) + \mathbb{W}(B)$$

für zwei beliebige einander ausschließende Ereignisse A und B . Die Interpretation liegt auf der Hand: Die Vereinigung $A \cup B$ bezeichnet das Ereignis, dass ein Elementarereignis eintritt, das in der Vereinigung von A und B liegt, mit anderen Worten: $A \cup B$ tritt ein, wenn von den Ereignissen A und B wenigstens eines eintritt. Haben A und B leeren Schnitt $A \cap B = \emptyset$, so bedeutet $A \cup B$ „entweder A oder B “. Die Wahrscheinlichkeit dafür muss in diesem Fall die Summe der Wahrscheinlichkeiten $\mathbb{W}(A)$ und $\mathbb{W}(B)$ sein.

Es zeigt sich, dass man, um eine stärkere Theorie zu erhalten, diese (endliche) Additivität sogar zur σ -Additivität verschärfen kann, ohne sich damit Probleme einzuhandeln. Dabei kommen nicht nur zwei, sondern eine (abzählbar) unendliche Folge einander ausschließender Ereignisse (d.h. paarweise disjunkter Mengen) A_1, A_2, \dots vor:

$$\mathbb{W}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{W}(A_n)$$

Damit diese Vereinigungen sowie einfache mengentheoretische Operationen mit Ereignissen gebildet werden können, verlangt man, dass mit messbaren Ereignissen A_n , $n \in \mathbb{N}$, stets auch die Vereinigung $\bigcup_{n=1}^{\infty} A_n$ ein messbares Ereignis ist, also in \mathcal{A} liegt, außerdem Komplemente (d.h. $A \in \mathcal{A}$ impliziert $\Omega \setminus A \in \mathcal{A}$) und damit auch Durchschnitte

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \Omega \setminus (\Omega \setminus A_n) = \Omega \setminus \bigcup_{n=1}^{\infty} (\Omega \setminus A_n).$$

Wegen $\emptyset, \Omega \in \mathcal{A}$ (wie schon vorher festgehalten) muss es sich bei \mathcal{A} also um eine σ -Algebra handeln und bei \mathbb{W} um ein Maß, so wie wir es bereits in Mathematik 1 definiert haben. Man darf sich also die Bestimmung von Wahrscheinlichkeiten ähnlich vorstellen wie die Messung von Längen, Flächen und Volumina. Im Zuge dieses Kapitels werden wir sehen, dass mit diesem Zugang Stochastik tatsächlich nicht nur sinnvoll betrieben werden kann, sondern dass dabei sogar sehr starke Ergebnisse folgen.

Der Übersicht halber fassen wir zusammen:

Definition 5.1.1.1. Ein **Wahrscheinlichkeitsraum** (*W-Raum*) ist ein Tripel $(\Omega, \mathcal{A}, \mathbb{W})$ mit folgenden Eigenschaften:

1. Ω ist eine nichtleere Menge.
2. \mathcal{A} ist eine σ -Algebra auf Ω . Das bedeutet explizit:
 - a) $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ mit $\emptyset \in \mathcal{A}$ und $\Omega \in \mathcal{A}$.

- b) $A \in \mathcal{A}$ impliziert $\Omega \setminus A \in \mathcal{A}$.
 c) $A_1, A_2, \dots \in \mathcal{A}$ impliziert $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

3. \mathbb{W} ist ein sogenanntes **Wahrscheinlichkeitsmaß**, d.h. ein auf 1 normiertes Maß auf \mathcal{A} . Das bedeutet explizit:

- a) $\mathbb{W} : \mathcal{A} \rightarrow [0, 1]$ mit $\mathbb{W}(\emptyset) = 0$ und $\mathbb{W}(\Omega) = 1$
 b) \mathbb{W} ist **σ -additiv**, d.h.: Aus $A_1, A_2, \dots \in \mathcal{A}$ und $A_i \cap A_j = \emptyset$ für alle $i \neq j \in \{1, 2, \dots\}$ folgt

$$\mathbb{W}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{W}(A_n).$$

Die Elemente von Ω heißen **Elementarereignisse**. Teilmengen A von Ω heißen **Ereignisse**, jene mit $A \in \mathcal{A}$ heißen **messbare Ereignisse**. Der Wert $\mathbb{W}(A)$ für $A \in \mathcal{A}$ heißt **Wahrscheinlichkeit** des (messbaren) Ereignisses A .

Die σ -Additivität impliziert auch die endliche Additivität: Für endlich viele paarweise disjunkte messbare Mengen A_1, \dots, A_n gilt

$$\mathbb{W}\left(\bigcup_{i=1}^n A_i\right) = \mathbb{W}(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n \mathbb{W}(A_i).$$

Zum Beweis hat man lediglich $A_{n+1} = A_{n+2} = \dots = \emptyset$ zu setzen und die σ -Additivität anzuwenden.

5.1.2 W-Räume und Zufallsexperimente, grundlegende Beispiele

Inhalt in Kurzfassung: Das allgemeine Konzept des W-Raums aus 5.1.1 wird an Beispielen illustriert wie: Münzwurf, einmaliges Würfeln, mehrmaliger Münzwurf, unendlich wiederholter Münzwurf, das Einheitsintervall als W-Raum.

Wir wollen nun einige sehr typische Beispiele behandeln, die zeigen, wie W-Räume konkret als mathematische Modelle von zufälligen Abläufen (Zufallsexperimenten) dienen können.

Münzwurf: Besteht das Zufallsexperiment aus dem Werfen einer Münze, so ist

$$\Omega = \{\text{Kopf}, \text{Zahl}\} = \{K, Z\}.$$

Sei $p = \mathbb{W}(K) \in [0, 1]^1$ die Wahrscheinlichkeit Kopf zu werfen. Dann folgt aus der Additivität wegen

$$1 = \mathbb{W}(\Omega) = \mathbb{W}(\{K\}) + \mathbb{W}(\{Z\}) = p + \mathbb{W}(Z)$$

¹Eigentlich sollte es an dieser Stelle $\mathbb{W}(\{K\})$ heißen, weil der Buchstabe K hier ja für das Elementarereignis „Kopf“ steht, das ein Element und keine Teilmenge von Ω ist. Wir verwenden für Elementarereignisse aber oft die kürzere Schreibweise ohne Mengenklammer.

sofort $\mathbb{W}(Z) = 1 - p$. Im W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$, der unser Zufallsexperiment modelliert, ist also $\mathcal{A} = \mathcal{P}(\Omega) = \{\emptyset, \{K\}, \{Z\}, \Omega\}$ und $\mathbb{W} : \mathcal{A} \rightarrow [0, 1]$ vollständig gegeben durch

$$\mathbb{W}(\emptyset) = 0, \quad \mathbb{W}(\{K\}) = p, \quad \mathbb{W}(\{Z\}) = 1 - p, \quad \mathbb{W}(\Omega) = 1.$$

Ist $p = 1 - p = \frac{1}{2}$, so spricht man auch von einer fairen Münze.

Würfeln, Laplace-Experiment: Besteht das Zufallsexperiment aus dem Werfen eines Würfels, so bestehen die Elementarereignisse im Würfeln einer bestimmten Augenzahl von 1 bis 6. Obwohl diese Ereignisse streng genommen von den Zahlen (und auch von den Ziffern) 1 bis 6 zu unterscheiden sind, schreiben wir der einfacheren Notation halber trotzdem

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Ist der Würfel fair (d.h. tritt jede der sechs möglichen Augenzahlen mit derselben Wahrscheinlichkeit $\frac{1}{6}$ auf), so ist im W-Raum $(\Omega, \mathcal{P}(\Omega), \mathbb{W})$, der dieses Zufallsexperiment modelliert,

$$\mathbb{W}(A) = \frac{|A|}{6}.$$

Weil jedes der endlich vielen (in diesem Fall sechs) Elementarereignisse aus Ω dieselbe Wahrscheinlichkeit erhält, spricht man auch von **diskreter Gleichverteilung** auf Ω . Abgesehen von den Elementarereignissen können wir auch für kompliziertere Ereignisse die Wahrscheinlichkeit bestimmen, etwa: $\mathbb{W}(G) = \mathbb{W}(U) = \mathbb{W}(P) = \frac{1}{2}$ für die Mengen $G = \{2, 4, 6\}$ der geraden, $U = \{1, 3, 5\}$ der ungeraden Augenzahlen und $P = \{2, 3, 5\}$ der Augenzahlen, die Primzahlen sind; oder $\mathbb{W}(U \cap P) = \frac{1}{3}$, $\mathbb{W}(G \cap P) = \frac{1}{6}$, $\mathbb{W}(U \cup P) = \frac{2}{3}$, $\mathbb{W}(G \cup P) = \frac{5}{6}$.

Im allgemeinen Fall diskreter Gleichverteilung ergibt sich die Wahrscheinlichkeit eines Ereignisses A nach der Formel

$$\mathbb{W}(A) = \frac{|A|}{|\Omega|},$$

die der Merkregel „günstige durch mögliche Fälle“ entspricht. Man spricht in dieser Situation auch von einem **Laplace-Experiment**.

Als Warnung sei hier betont, dass bei der Modellbildung die Annahme, dass alle Elementarereignisse gleich wahrscheinlich sind, nicht selbstverständlich ist, sondern einer Rechtfertigung bedarf. Oft ist es auf den ersten Blick gar nicht klar, welche Elementarereignisse es sind, für die diese Annahme zutrifft. Als einfaches Beispiel denke man an den Wurf mit zwei fairen Würfeln. Hier sind es die insgesamt 36 möglichen Paare $(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)$, denen man sinnvoller Weise gleiche Wahrscheinlichkeit, also $\frac{1}{36}$ beimisst, auch wenn man sich beispielsweise für die Summe der Augenzahlen interessiert, die folglich unterschiedliche Wahrscheinlichkeiten haben, von $\mathbb{W}(2) = \mathbb{W}(12) = \frac{1}{36}$ bis $\mathbb{W}(7) = \frac{6}{36} = \frac{1}{6}$. Bei der Ermittlung der Anzahlen von möglichen bzw. häufigen Fällen sind häufig kombinatorische Formeln nützlich, woran man die Bedeutung der Kombinatorik als Hilfsmittel in der Wahrscheinlichkeitstheorie ermisst. Wir werden etwas weiter unten nochmals darauf zurückkommen. Sind im aktuellen Beispiel mit zwei Würfeln überdies Zweifel an der Fairness der Würfel angebracht,

so verliert die Maxime „günstige durch mögliche Fälle“ gänzlich ihre Grundlage.

Endliche Versuchswiederholungen: Ein einzelner Münzwurf (analog für Würfel) wäre noch kaum hinreichender Grund, um einen W -Raum zur Modellierung heranzuziehen. Wahrscheinlichkeit wird erst dann manifest, wenn viele Ereignisse vorliegen, an denen man die Häufigkeiten der verschiedenen möglichen Versuchsausgänge studieren kann. Nehmen wir also an, dass eine Münze nicht nur einmal geworfen wird, sondern n -mal, $n \in \mathbb{N}$. Für Kopf und Zahl schreiben wir nun 0 und 1. Ein Versuchsergebnis (Elementarereignis) ist dann gegeben durch eine 0-1-Folge (a_1, \dots, a_n) , $a_i \in \{0, 1\}$ der Länge n . Ist die Münze fair, so erhält jedes dieser 2^n möglichen Elementarereignisse dieselbe Wahrscheinlichkeit, nämlich 2^{-n} (diskrete Gleichverteilung wie beim einmaligen Würfeln). Der W -Raum $(\Omega, \mathcal{P}(\Omega), \mathbb{W})$ zur Modellierung dieses Zufallsexperiments ist also gegeben durch die Menge $\Omega = \{0, 1\}^n$ aller 0-1-Folgen der Länge n , die σ -Algebra $\mathcal{A} = \mathcal{P}(\Omega)$, die alle Teilmengen von Ω enthält und, ähnlich wie beim einmaligen Münzwurf, die Wahrscheinlichkeit \mathbb{W} mit

$$\mathbb{W}(A) = \frac{|A|}{2^n}.$$

Darin wurde stillschweigend die Annahme gemacht, dass Münzwürfe einander nicht beeinflussen. Man spricht von **Unabhängigkeit** der einzelnen Münzwürfe. Dieser Begriff ist für die gesamte Stochastik zentral. Wir werden noch ausführlicher auf ihn zurückkommen, und er wird uns durch das ganze Kapitel begleiten. Beim Münzwurf bedeutet Unabhängigkeit, dass für den $i+1$ -ten Münzwurf die Wahrscheinlichkeiten für Kopf und Zahl nicht davon abhängen, wie die ersten i Münzwürfe ausgegangen sind. In diesem Beispiel scheint diese Annahme selbstverständlich zu sein. In vielen Anwendungen muss man diese Annahme aber sorgfältig prüfen.

Unendliche Versuchswiederholungen: Zwar ist es real nicht möglich, eine Münze unendlich oft zu werfen. Die W -Theorie entfaltet ihre Macht aber erst dann, wenn man sehr wohl auch unendliche Versuchsreihen ins Auge fasst. Die Hauptsätze der Stochastik, die wir später ausführlich besprechen werden, geben dafür eindrucksvolles Zeugnis. In Verallgemeinerung der endlichen Versuchsreihe wollen wir nun also einen W -Raum betrachten, wo die Elementarereignisse unendliche Folgen $\mathbf{a} = (a_n)$ mit $a_n \in \{0, 1\}$ sind. In diesem Fall erweist es sich als unmöglich, allen Teilmengen $A \subseteq \Omega$ eine Wahrscheinlichkeit $\mathbb{W}(A)$ zuzuweisen. Der Grund ist von ähnlicher Art wie für die in Mathematik 1 besprochene Unmöglichkeit, allen Teilmengen der reellen Zahlen ein Längenmaß zuzuordnen. (Dass diese beiden Probleme sehr eng zusammenhängen, wird plausibel, wenn man daran denkt, dass jede unendliche 0-1-Folge als Binärdarstellung einer Zahl des Einheitsintervalls aufgefasst werden kann. Wir werden im nächsten Beispiel auf diese Analogie zurückkommen.) Man kann aber zum Beispiel solche Mengen unendlicher Folgen betrachten, die durch Bedingungen an einen endlichen Anfangsabschnitt der Folgen bestimmt sind. Zum Beispiel könnte man die Menge A_{010} all jener Folgen $\mathbf{a} = (a_n)$ betrachten, die mit 010 beginnen, wo also $a_1 = 0$, $a_2 = 1$ und $a_3 = 0$ gilt. Da bei einer fairen Münze die Wahrscheinlichkeit, bei den ersten Versuchen genau diese Werte zu

werfen, ein Achtel ist, sollte $\mathbb{W}(A_{010}) = \frac{1}{8}$ gelten. Das Entsprechende kann man natürlich mit jeder endlichen Folge als Anfangsabschnitt machen. Tatsächlich ergibt sich mit Hilfe dieser Forderungen ein W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ der gewünschten Art, wo die σ -Algebra \mathcal{A} zwar nicht alle $A \in \mathcal{P}(\Omega)$ enthält, jedenfalls aber A_{010} und sämtliche analog gebauten Mengen und, wegen der Abgeschlossenheit bezüglich der Bildung von Komplementen sowie abzählbarer Vereinigungen und Schnitten, viele, viele Mengen darüber hinaus. Das ist reicht vollkommen aus für die W-Theorie.

Das Einheitsintervall als W-Raum: Wir betrachten nun das Einheitsintervall $[0, 1]$ und das darauf eingeschränkte Lebesguemaß $\lambda_{[0,1]}$. Es ist definiert auf den in $[0, 1]$ enthaltenen Lebesgue-messbaren Mengen und stimmt dort mit dem Lebesguemaß (Längenmaß) auf ganz \mathbb{R} überein. Dann liegt ein W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ vor mit $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{A}_{[0,1]}$ (die σ -Algebra der in $[0, 1]$ enthaltenen Lebesgue-messbaren Mengen) und $\mathbb{W} = \lambda_{[0,1]}$ (das auf die in $[0, 1]$ enthaltenen messbaren Mengen eingeschränkte Lebesguemaß). Man kann sich diesen Raum als Modell für ein Zufallsexperiment vorstellen, wo das Einheitsintervall als eindimensionale Zielscheibe für einen vom Zufall beherrschten Wurf fungiert. A priori kann jeder Punkt des Intervalls getroffen werden. Die Wahrscheinlichkeit, in ein bestimmtes Teilintervall zu treffen, ist gleich der Länge des Intervalls. Man spricht deshalb auch von der **stetigen Gleichverteilung** auf $[0, 1]$. Stellt man sich die Punkte des Einheitsintervalls in Binärdarstellung vor, so ist ein Elementarereignis durch eine unendliche 0-1-Folge gegeben, wie im Beispiel zuvor.² Der Menge A_{010} von dort entspricht dann die Menge jener reellen Zahlen mit Binärdarstellung $0,010\dots$. Das ist das Intervall $[\frac{1}{4}, \frac{3}{8}]$ mit der Länge $\frac{1}{8}$, was mit der Wahrscheinlichkeit $\mathbb{W}(A_{010}) = \frac{1}{8}$ übereinstimmt. In diesem Sinn entspricht die stetige Gleichverteilung auf $[0, 1]$ im wahrscheinlichkeitstheoretischen Sinn also dem unendlich wiederholten Münzwurf. Das erklärt auch die Bemerkung beim vorangegangenen Beispiel, dass die Probleme mit nicht messbaren Mengen sich von den reellen Zahlen auf unendliche Folgenräume übertragen.

Aus diesen Beispielen wird plausibel, dass die Wirksamkeit der Wahrscheinlichkeitstheorie bei solchen Experimenten besonders deutlich wird, wo mehrere zufällige Prozesse hintereinander stattfinden. Typisch sind daher Situationen mit Stichproben aus mehreren einzelnen Ereignissen, die man sich oft als zufällige „Ziehungen“ von Elementen aus einer Grundmenge vorstellt. Für weitere Ziehungen kann man die bereits gezogenen Objekte wieder zulassen oder nicht. Je nachdem spricht man vom *Ziehen mit Zurücklegen* oder *Ziehen ohne Zurücklegen*. Außerdem nennen wir die Stichprobe *geordnet* bzw. *ungeordnet*, je nachdem, ob die Reihenfolge der gezogenen Elemente eine Bedeutung hat oder nicht. Die resultierenden Fragestellungen laufen oft auf kombinatorische Probleme hinaus, von denen wir einige schon im ersten Kapitel von Mathematik 1 behandelt haben. Die Ausgangssituation ist dabei, dass k Elemente aus einer Grundmenge der Größe

²Wir sehen davon ab, dass manche Zahlen durch zwei verschiedene Binärfolgen gegeben sein können, etwa $0,010000000\dots = 0,001111111\dots = \frac{1}{4}$. Diese Schwierigkeit löst sich dadurch auf, dass es nur abzählbar viele solche Zahlen gibt, jede davon als Einermenge vom Maß 0, was sich wegen der σ -Additivität wieder nur zu einer Menge vom Maß 0 aufaddiert. Rein maßtheoretisch unterscheiden sich die beiden Sichtweisen also nicht nennenswert.

n ausgewählt werden. Zu unterscheiden sind:

- *Geordnete Stichproben mit Zurücklegen*
Mögliche Versuchsausgänge: n^k , entspricht der Anzahl der Funktionen von einer k - in eine n -elementige Menge.
- *Geordnete Stichproben ohne Zurücklegen*
Mögliche Versuchsausgänge: $\frac{n!}{(n-k)!}$, entspricht der Anzahl der injektiven Funktionen von einer k - in eine n -elementige Menge (für $k \leq n$).
- *Ungeordnete Stichproben ohne Zurücklegen*
Mögliche Versuchsausgänge: $\binom{n}{k}$, entspricht der Anzahl der k -elementigen Teilmengen einer n -elementigen Menge (für $k \leq n$).
- *Ungeordnete Stichproben mit Zurücklegen*
Mögliche Versuchsausgänge: $\binom{n+k-1}{k}$, entspricht der Anzahl der k -elementigen Multi-Teilmengen einer n -elementigen Menge (die Teilmengen dürfen hier Elemente „mehrfach“ enthalten).

Übungsaufgabe 272. (T) Lösen Sie folgende Probleme mit den geeigneten Grundformeln der Kombinatorik:

1. *Wieviele Möglichkeiten gibt es einen Lottoschein (6 aus 45) auszufüllen?*
2. *Wieviele Möglichkeiten gibt es einen Lottoschein (6 aus 45) auszufüllen und keine einzige Zahl richtig anzukreuzen?*
3. *Wieviele Möglichkeiten gibt es einen Lottoschein (6 aus 45) auszufüllen und genau eine Zahl richtig anzukreuzen?*
4. *Warum wird in Österreich beim Lotto 6 aus 45 in Deutschland aber 6 aus 49 gespielt?*

Übungsaufgabe 273. (T) Lösen Sie folgende Probleme mit den geeigneten Grundformeln der Kombinatorik:

1. *Wie viele Möglichkeiten gibt es einen Totoschein (12 Tipps mit 1,2,X) auszufüllen?*
2. *Wie viele Möglichkeiten gibt es einen Totoschein (12 Tipps mit 1,2,X) auszufüllen und keinen Tipp richtig zu haben?*
3. *Wie viele Möglichkeiten gibt es einen Totoschein (12 Tipps mit 1,2,X) auszufüllen und alle Tipps richtig zu haben?*

Übungsaufgabe 274. (T) Adam und Beate reservieren unabhängig voneinander einen Platz im Speisewagen des Orient-Express. Es gibt dort 5 Tische zu je 4 Plätzen. Die Zuteilung von Reservierung zu Sitzplatz erfolgt zufällig, d.h. wir nehmen ein Laplace-Experiment an. Wie wahrscheinlich ist es, dass Adam und Beate am selben Tisch zu sitzen kommen?

Übungsaufgabe 275. (T) Auf einer Party befinden sich n Personen. Wir nehmen an, dass die Geburtstage dieser Personen zufällig und mit gleichen Wahrscheinlichkeiten auf die 365 Tage eines Jahres (kein Schaltjahr!) verteilt sind. Wie groß muss n mindestens sein, damit die Wahrscheinlichkeit, dass sich auf der Party mindestens zwei Personen mit gleichem Geburtstag befinden größer als 50% ist?

Anleitung: Schreiben Sie zuerst die Gegenwahrscheinlichkeit, d.h. die Wahrscheinlichkeit dass alle an unterschiedlichen Tagen Geburtstag haben an, und zücken Sie erst dann ihren Taschenrechner!

Übungsaufgabe 276. (T)

1. Geben Sie die Wahrscheinlichkeit an, dass bei gleichzeitigem Werfen von zwei fairen fünfseitigen Würfeln (W_5) die Gesamtaugenanzahl $k \in \{2, 3, \dots, 10\}$ beträgt.
2. Geben Sie die Wahrscheinlichkeit an, dass bei gleichzeitigem Werfen von drei fairen dreiseitigen Würfeln (W_3) die Gesamtaugenanzahl $k \in \{3, 4, \dots, 9\}$ beträgt.
3. Geben Sie die Wahrscheinlichkeit an, dass bei gleichzeitigem Werfen von zwei fairen sechsseitigen Würfeln (W_6) die Gesamtaugenanzahl $k \in \{2, 3, \dots, 12\}$ beträgt.

Übungsaufgabe 277. (E) Würden Sie folgender Aussage zustimmen?

„Eine zufällige gewählte natürliche Zahl $n \geq 1$ ist mit 50 Prozent Wahrscheinlichkeit gerade und mit 50 Prozent Wahrscheinlichkeit ungerade.“

Begründen Sie ausführlich Ihre Antwort!

Übungsaufgabe 278. (T) Führen Sie die Menge Ω aller Elementarereignisse des Zufallsexperiments Würfeln mit einem achtseitigen Würfel (W_8) an. Geben Sie außerdem

1. zwei Ereignisse A und B in Ω an, die nicht unvereinbar sind. Beschreiben Sie die Menge $A \cap B$.
2. Geben Sie ein weiteres Ereignis C in Ω an und beschreiben Sie die Menge $A \cup C$
3. Geben Sie zwei unvereinbare Ereignisse E und F in Ω an, sodass das Ereignis $E \cup F$ der ganze Raum ist.

Übungsaufgabe 279. (T) Das Experiment bestehe aus dem Werfen eines sechsseitigen Würfels (W_6) und einer Münze (Kopf oder Zahl). Zeigt die Münze Zahl, so wird die doppelte Augenzahl des Würfels notiert, bei Kopf nur die einfache.

1. Geben Sie einen geeigneten Ergebnisraum Ω an.
2. Wie groß ist, unter der Laplace-Annahme, die Wahrscheinlichkeit, dass eine gerade Zahl notiert wird?

5.1.3 Diskrete und reelle Verteilungen, Verteilungsfunktionen und Dichten

Inhalt in Kurzfassung: Diskrete W-Räume haben endliche oder abzählbar unendliche Trägermenge, und alle Teilmengen sind messbar. Das gilt in der Regel nicht für reelle W-Verteilungen. Reelle Verteilungen lassen sich auch durch Verteilungsfunktionen vollständig beschreiben, die durch drei Eigenschaften charakterisiert sind: Grenzwert 0 und 1 für $x \rightarrow \pm\infty$, monoton wachsend und rechtsstetig. Jede Verteilungsfunktion lässt sich auch durch ihre α -Quantile beschreiben, jene Stellen, wo sie, grob gesagt, den Wert α überschreitet. Spezielle Quantile sind Median ($\alpha = \frac{1}{2}$) und die beiden Quartile ($\alpha = \frac{1}{4}$ und $\alpha = \frac{3}{4}$). Diskrete reelle Verteilungen ergeben sich über diskrete Verteilungen auf Teilmengen von \mathbb{R} . Konträr dazu sind absolut stetige Verteilungen auf \mathbb{R} , die sich mittels Integralen über eine Dichtefunktion beschreiben lassen. Ist die Dichtefunktion stetig, so ergibt sich aus dem Hauptsatz der Differential- und Integralrechnung, dass die Verteilungsfunktion differenzierbar ist und ihre Ableitung mit der Dichte übereinstimmt.

Die Komplikation, dass in einem W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ die σ -Algebra \mathcal{A} nicht mit der gesamten Potenzmenge $\mathcal{P}(\Omega)$ übereinstimmen muss, ist zwar bei vielen W-Räumen unausweichlich, doch gibt es auch viele interessante Fälle mit $\mathcal{A} = \mathcal{P}(\Omega)$. Insbesondere liegt diese Situation vor, wenn Ω endlich oder abzählbar unendlich ist, und alle einelementigen Mengen $\{\omega\}$ mit $\omega \in \Omega$ messbar sind. In 5.1.2 waren die Beispiele mit ein- oder mehrmaligem, nicht aber unendlich wiederholtem Münzwurf oder Würfeln von dieser Art. Die folgende Definition macht deutlich, worauf es dabei ankommt:

Definition 5.1.3.1. *Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{W})$ heißt **diskret**, wenn Ω endlich oder abzählbar unendlich ist und $\mathcal{A} = \mathcal{P}(\Omega)$ alle Teilmengen von Ω enthält.*

Sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein diskreter W-Raum, also $\Omega = \{\omega_1, \dots, \omega_n\}$ endlich oder $\Omega = \{\omega_n : n \in \mathbb{N}\}$ abzählbar unendlich, jeweils mit $\omega_i \neq \omega_j$ für $i \neq j$. Alle einelementigen Mengen sind messbar, also sind die Wahrscheinlichkeiten $p_n := \mathbb{W}(\{\omega_n\})$ definiert. Wenn letzteres der Fall ist, schreiben wir (nicht nur für diskrete W-Räume), etwas ungenau, wie auch schon an früherer Stelle $p_n = \mathbb{W}(\omega_n)$. Wie aus Normierung $\mathbb{W}(\Omega) = 1$ und σ -Additivität folgt, gilt in diskreten W-Räumen mit obiger Notation

$$\sum_i p_i = \mathbb{W}(\Omega) = 1.$$

Dabei erstreckt sich die Summation im endlichen Fall über die Indizes $i = 1, \dots, n$, im unendlichen Fall liegt eine (absolut konvergente) unendliche Reihe vor. Eine beliebige Menge $A \subseteq \Omega$ hat, wieder wegen der σ -Additivität, die Wahrscheinlichkeit

$$\mathbb{W}(A) = \mathbb{W}\left(\bigcup_{i: \omega_i \in A} \{\omega_i\}\right) = \sum_{i: \omega_i \in A} p_i.$$

In vielen interessanten Fällen mit überabzählbar unendlichem Ω (zwei davon haben wir bereits in 5.1.2 behandelt, nämlich unendlich wiederholter Münzwurf und stetige Gleichverteilung auf dem Einheitsintervall) ist die Einschränkung auf eine Teil- σ -Algebra

\mathcal{A} von $\mathcal{P}(\Omega)$ notwendig. Insbesondere gilt das für $\Omega = \mathbb{R}$, einen Fall den wir wegen seiner herausragenden Bedeutung ebenfalls durch eine eigene Definition hervorheben wollen.

Definition 5.1.3.2. Ein Wahrscheinlichkeitsraum heißt **reell**, wenn er von der Form $(\mathbb{R}, \mathcal{A}, \mathbb{W})$ ist, wobei \mathcal{A} wenigstens alle Intervalle (offen, halboffen, abgeschlossen, beschränkt, unbeschränkt) enthält. In diesem Fall heißt \mathbb{W} eine **Wahrscheinlichkeitsverteilung** (*W-Verteilung*) auf \mathbb{R} oder auch eine **reelle Wahrscheinlichkeitsverteilung**. Kürzer sagt man dazu oft auch nur **Verteilung**.

Da bei reellen Verteilungen jede Menge der Form

$$(-\infty, x] = \bigcup_{k \in \mathbb{Z}: k \leq x} (k, x]$$

mit $x \in \mathbb{R}$ messbar ist, kann man jeder W-Verteilung auf \mathbb{R} ihre sogenannte Verteilungsfunktion zuordnen:

Definition 5.1.3.3. Ist \mathbb{W} eine reelle W-Verteilung, so heißt die reelle Funktion $F_{\mathbb{W}}$, definiert durch

$$F_{\mathbb{W}}(x) := \mathbb{W}((-\infty, x]) \quad \text{für alle } x \in \mathbb{R},$$

die **Verteilungsfunktion** von \mathbb{W} .

Man überlegt sich leicht, dass für eine Verteilungsfunktion $F = F_{\mathbb{W}}$ stets gilt:

1. $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$.
2. F ist auf ganz \mathbb{R} monoton wachsend.
3. F ist rechtsstetig, d.h. $\lim_{x \rightarrow x_0^+} F(x) = F(x_0)$ für alle $x_0 \in \mathbb{R}$.

Übungsaufgabe 280. (E) Begründen Sie diese drei Eigenschaften. Hinweis: Die σ -Additivität von \mathbb{W} fließt sowohl in der ersten als auch in der dritten Eigenschaft ein.

Umgekehrt ist jede Funktion F mit diesen drei Eigenschaften die Verteilungsfunktion einer (eindeutig bestimmten) W-Verteilung $\mathbb{W} = \mathbb{W}_F$ mit

$$\mathbb{W}((a, b]) = F(b) - F(a)$$

für alle $a < b \in \mathbb{R}$. Somit enthalten reelle W-Verteilungen und Verteilungsfunktionen dieselbe Information und können durch das jeweils andere Objekt ersetzt werden, wann immer es bequem erscheint.

Insbesondere wissen wir über eine Verteilung alles, wenn wir zu jedem $\alpha \in [0, 1]$ wissen, wo die monoton wachsende Verteilungsfunktion den Wert α überschreitet. Man definiert:

Definition 5.1.3.4. Eine Zahl $x_\alpha \in \mathbb{R}$ heißt **α -Quantil** für die Verteilung \mathbb{W} mit Verteilungsfunktion F , wenn $F(x) \leq \alpha$ für alle $x < x_\alpha$ gilt und $F(x) \geq \alpha$ für alle $x > x_\alpha$. (Manchmal bezeichnet man auch die Menge aller $x_\alpha \in \mathbb{R}$ mit dieser Eigenschaft als das α -Quantil der Verteilung.)

Speziell heißt ein α -Quantile für $\alpha = \frac{1}{2}$ auch **Median** der Verteilung, für $\alpha \in \{\frac{1}{4}, \frac{3}{4}\}$ spricht man auch von einem **Quartil**.

Jedes $x \in \mathbb{R}$ ist offensichtlich α -Quantil für $\alpha := F(x)$. Umgekehrt gibt es für jedes α mit $0 < \alpha < 1$ mindestens ein α -Quantil: Wegen $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$ gibt es nämlich $x_1 < x_2 \in \mathbb{R}$ mit $F(x_1) < \alpha < F(x_2)$. Wegen der Monotonie von F hat die Menge aller $x \in \mathbb{R}$ mit $\alpha \leq F(x)$ ein Infimum x_α , das (wie man sich schnell überlegt) ein α -Quantil sein muss.

Übungsaufgabe 281. (E) Führen Sie diese Überlegung aus.

Ist F an dieser Stelle nicht nur rechtsstetig, sondern beidseitig stetig, so gilt sogar $F(x_\alpha) = \alpha$. Ist F an dieser Stelle nicht stetig mit

$$\alpha_0 := \lim_{x \rightarrow x_\alpha^-} F(x) < \alpha_1 := \lim_{x \rightarrow x_\alpha^+} F(x) = F(x_\alpha),$$

so ist x_α gleichzeitig α' -Quantil für alle $\alpha' \in [\alpha_0, \alpha_1]$.

Ist F auf einem Intervall von (a, b) konstant mit Wert α , dann wegen der Rechtsstetigkeit auch auf $[a, b)$, während $F(b) > \alpha$ möglich ist. Dennoch sind alle $x \in [a, b]$ dann α -Quantile.

Ist die Trägermenge Ω eines diskreten W-Raums $(\Omega, \mathcal{A}, \mathbb{W})$ eine Teilmenge von \mathbb{R} , so können wir aus diesem Raum sehr leicht einen reellen W-Raum $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mathbb{W}_{\mathbb{R}})$ machen, wo sogar *alle* Teilmengen der (überabzählbaren) Menge \mathbb{R} messbar sind. Für beliebiges $A \subseteq \mathbb{R}$ haben wir als Wahrscheinlichkeit lediglich die Summe aller Wahrscheinlichkeiten von Elementen $\omega \in \Omega$ zu nehmen, die in A liegen. Man spricht von diskreten reellen Verteilungen:

Definition 5.1.3.5. Eine diskrete (reelle) Verteilung $\mathbb{W}_{\mathbb{R}}$ liegt vor, wenn es einen diskreten W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ mit $\Omega \subseteq \mathbb{R}$ gibt, aus dem der W-Raum $(\mathbb{R}, \mathcal{P}(\mathbb{R}), \mathbb{W}_{\mathbb{R}})$ gemäß

$$\mathbb{W}_{\mathbb{R}}(A) := \sum_{\omega \in \Omega: \omega \in A} \mathbb{W}(\omega)$$

entsteht.

Einen starken Gegensatz zu diskreten bilden solche reelle Verteilungen, wie sie im folgenden Satz beschrieben werden.

Satz 5.1.3.6. Sei $\rho : \mathbb{R} \rightarrow \mathbb{R}$ mit $\rho(x) \geq 0$ für alle $x \in \mathbb{R}$ eine reelle Funktion mit

$$\int_{-\infty}^{\infty} \rho(x) dx = \int_{\mathbb{R}} \rho d\lambda = 1.$$

(Wie in Mathematik 1 bezeichnet λ hier das Lebesguemaß.) Dann wird durch

$$\mathbb{W}_{\rho}(A) := \int_A \rho(x) dx = \int_A \rho d\lambda$$

für alle $A \in \mathcal{A}$ (\mathcal{A} bezeichnet die σ -Algebra der Lebesgue-messbaren Mengen $A \subseteq \mathbb{R}$) ein W-Raum $(\mathbb{R}, \mathcal{A}, \mathbb{W}_{\rho})$ definiert. Die Funktion ρ heißt die **Dichtefunktion** oder **Dichte** der Verteilung \mathbb{W}_{ρ} . Jedes W-Maß auf \mathbb{R} , zu dem es eine Dichte gibt, heißt **absolut stetig** (bezüglich des Lebesguemaßes).

Ein Beispiel eines absolut stetigen Maßes haben wir schon kennen gelernt, nämlich die stetige Gleichverteilung auf $[0, 1]$. Als Dichte kann man die Indikatorfunktion $\rho := \mathbf{1}_{[0,1]}$ des Einheitsintervalls mit $\rho(x) := 1$ für $x \in [0, 1]$ und $\rho(x) = 0$ sonst nehmen. Für ein beliebiges Intervall $I = [a, b]$ mit $a < b$ definiert die Dichtefunktion $\rho := \frac{1}{b-a} \mathbf{1}_{[a,b]}$, also $\rho_I(x) := \frac{1}{b-a}$ für $x \in I$ und $\rho_I(x) = 0$ sonst, die **stetige Gleichverteilung auf $[a, b]$** .

Man beachte, dass man Dichtefunktionen ρ an einzelnen Punkten oder sogar auf beliebigen λ -Nullmengen umdefinieren kann, ohne dass sich dadurch die W -Verteilung \mathbb{W}_ρ ändert. Deshalb sind Dichtefunktionen nie eindeutig bestimmt. Zu einer gegebenen Verteilung \mathbb{W} kann es aber höchstens eine stetige Dichtefunktion ρ mit $\mathbb{W}_\rho = \mathbb{W}$ geben. Gibt es eine solche, so spielt diese natürlich eine bevorzugte Rolle. Das wichtigste Beispiel, das wir später noch ausführlich behandeln werden, ist die Dichtefunktion

$$\rho(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

der (Standard-) Normalverteilung.

Wir erinnern uns nochmals an den Zusammenhang zwischen Verteilung \mathbb{W} und der ihr zugeordneten Verteilungsfunktion F . Ist überdies ρ eine Dichte von \mathbb{W} , so bedeutet das

$$F(x) = \int_{-\infty}^x \rho(x) dx.$$

Ist ρ stetig, so folgt aus dem Hauptsatz der Differential- und Integralrechnung, dass F differenzierbar ist mit $F' = \rho$. Also:

Satz 5.1.3.7. *Hat eine W -Verteilung \mathbb{W} eine stetige Dichtefunktion ρ und ist F die Verteilungsfunktion von \mathbb{W} , so gilt $F' = \rho$.*

Erst recht sind Verteilungsfunktionen von absolut stetigen Verteilungen stetig. Die Umkehrung gilt nicht. Man kann zum Beispiel sogenannte **singuläre Maße** konstruieren, deren Verteilungsfunktion F stetig und auch fast überall, d.h. überall außerhalb einer geeigneten λ -Nullmenge N , differenzierbar ist mit $F'(x) = 0$ für alle $x \in \mathbb{R} \setminus N$, auf N aber nicht differenzierbar. Folglich haben singuläre Maße keine Dichte. Das klassische Beispiel eines solchen singulären Maßes ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 282. *(E) Man überlege sich, dass es ein W -Maß μ auf dem Einheitsintervall gibt wie folgt. Bezeichne M die Menge aller $x \in [0, 1]$, die eine Darstellung der Form $x = \sum_{n=1}^{\infty} \frac{2a_n}{3^n}$ mit $a_n \in \{0, 1\}$ haben. Für solche x sei $\varphi(x) := \sum_{n=1}^{\infty} \frac{a_n}{2^n}$ und für jedes Intervall $I \subseteq [0, 1]$ gelte $\mu(I) = \lambda(\varphi(I))$. Dadurch ist μ als W -Maß im Wesentlichen sogar eindeutig bestimmt. Dieses μ hat eine stetige Verteilungsfunktion, ist aber singulär.*

5.1.4 Zufallsgrößen und ihre Verteilung

Inhalt in Kurzfassung: Zufallsgrößen sind messbare Abbildungen X von einem W -Raum in die reellen Zahlen. Die Messbarkeit ist genau jene Bedingung, die sicherstellt, dass

einem Ereignis der Form „ $X(\omega)$ fällt in ein Intervall I “ eine wohldefinierte Wahrscheinlichkeit zukommt. Damit lässt sich die Verteilungsfunktion von X definieren, wodurch wiederum eine reelle Verteilung bestimmt ist, die die Verteilung von X heißt.

Zufallsgrößen oder auch **Zufallsvariable** sind jene Abbildungen, die den entscheidenden Zusammenhang herstellen zwischen Elementarereignissen und Messgrößen, für die man sich interessiert. Ein sehr einfaches Beispiel wäre die Zufallsgröße X , die dem Ereignis ω_k , eine bestimmte Augenzahl k zu würfeln, eben diese Zahl $X(\omega_k) = k$ zuordnet. Deutlicher als an diesem trivialen Beispiel wird der Sinn des Begriffs der Zufallsgröße, wenn man nicht nur einmal, sondern n -mal würfelt und sich für die Summe der gewürfelten Augenzahlen interessiert. Die entsprechende Zufallsgröße X ordnet dann einer Folge (k_1, \dots, k_n) von Augenzahlen die Summe $\sum_{i=1}^n k_i$ zu. Allgemein darf man sich im Zusammenhang mit Zufallsgrößen irgendeine Messdaten $X(\omega)$ vorstellen, die sich aus einem zufälligen Experiment mit Elementarereignis ω ergeben.

Das Wesentliche an einer Zufallsgröße X besteht darin, dass es eine wohldefinierte Wahrscheinlichkeit gibt, dass der Wert von X in ein beliebig vorgegebenes Intervall I fällt. Das heißt, dass die Menge $\{\omega \in \Omega : X(\omega) \in I\}$ messbar, d.h. ein Element der σ -Algebra \mathcal{A} im zugrundeliegenden W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ ist. Die allgemeine Definition lautet also:

Definition 5.1.4.1. *Es sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein W-Raum. Eine Funktion $X : \Omega \rightarrow \mathbb{R}$ heißt **Zufallsgröße** oder **Zufallsvariable** auf $(\Omega, \mathcal{A}, \mathbb{W})$, wenn sie messbar ist, d.h. wenn für alle reellen Intervalle I (offen, halboffen, abgeschlossen, beschränkt, unbeschränkt) das Urbild*

$$X^{(-1)}(I) = \{\omega \in \Omega : X(\omega) \in I\} = (X \in I) \in \mathcal{A}$$

messbar ist.

So wie die Schreibweise „ $(X \in I)$ “ in der Formel bezeichnet generell eine Aussage mit einer Zufallsvariablen (hier X) zwischen einem Paar von Klammern *die Menge aller $\omega \in \Omega$, auf die die Anwendung der Zufallsgröße zu einer wahren Aussage führt*. Für die Wahrscheinlichkeit zum Beispiel des Ereignisses $(a < X < b)$, dass der Wert von X zwischen a und b liegt schreibt man oft noch einfacher $\mathbb{W}(a < X < b)$ etc.

Beispiele: Auf einem beliebigen W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ ist für jedes $A \in \mathcal{A}$ die Indikatorfunktion $X := \mathbf{1}_A$ mit $\omega \mapsto 1$ für $\omega \in A$ und $\omega \mapsto 0$ für $\omega \in \Omega \setminus A$ messbar und somit, als reellwertige Funktion, eine Zufallsgröße. Ist der W-Raum diskret, so sind alle Teilmengen $A \subseteq \Omega$ und somit alle Urbilder unter irgendeiner Funktion auf X messbar. Deshalb sind auf diskreten W-Räumen alle Funktionen $X : \Omega \rightarrow \mathbb{R}$ auch Zufallsgrößen.

Generell hat die Messbarkeitsbedingung in der Definition einer Zufallsgröße X zur Folge, dass die Funktion

$$F_X(x) := \mathbb{W}(X \in (-\infty, x])$$

wohldefiniert ist. Sie hat alle Eigenschaften einer Verteilungsfunktion, also gibt es eine zugehörige W-Verteilung, die Verteilung von X :

Definition 5.1.4.2. Sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein W -Raum und X eine darauf definierte Zufallsgröße. Dann heißt die W -Verteilung $\mathbb{W} = \mathbb{W}_{F_X}$ mit Verteilungsfunktion

$$F_X(x) := \mathbb{W}(X \in (-\infty, x])$$

die **Verteilung von X** .

Man beachte, dass zu jeder W -Verteilung $\mathbb{W}_{\mathbb{R}}$ auf \mathbb{R} die identische Abbildung

$$X : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x$$

eine Zufallsgröße mit der Verteilung $\mathbb{W}_{\mathbb{R}}$ ist. Es gibt also zu jeder vorgegebenen Verteilung auf \mathbb{R} auch eine Zufallsgröße mit dieser Verteilung.

Eine der wichtigsten Aufgaben der Stochastik besteht darin, die Verteilung von Zufallsgrößen zu bestimmen. In der W -Theorie dominieren allgemeine, theoretische Überlegungen, mit denen man sehr starke Ergebnisse erzielen kann. In der angewandten Statistik versucht man – auf Basis der W -Theorie – aus empirischen Daten Rückschlüsse auf konkrete W -Verteilungen zu ziehen. Bevor wir diesen Fragen weiter nachgehen, wollen wir noch die zwei wichtigsten Kenngrößen von Verteilungen und somit von Zufallsgrößen kennenlernen: Erwartungswert und Varianz.

5.1.5 Erwartungswert

Inhalt in Kurzfassung: Der Erwartungswert lässt sich intuitiv als mittlerer oder durchschnittlicher Wert einer Verteilung oder einer Zufallsgröße interpretieren. Bei diskreten Verteilungen handelt es sich beim Erwartungswert um das mit den Wahrscheinlichkeiten gewichtete Mittel der auftretenden Werte. Bei Verteilungen mit Dichte lässt sich der Erwartungswert als reelles Integral darstellen. Verwendet man Integrale bezüglich beliebigen Maßen, so lassen sich beide Fälle in einer gemeinsamen Schreibweise behandeln. Das hat bei der weiteren Entfaltung der W -Theorie sehr große Vorteile.

Unter dem **Mittel-** oder **Erwartungswert** $\mathbb{E}(\mathbb{W})$ einer W -Verteilung \mathbb{W} auf \mathbb{R} hat man sich, so wie es die Bezeichnung suggeriert, einen Wert vorzustellen, den eine Zufallsgröße mit Verteilung \mathbb{W} „im Mittel“, d.h. durchschnittlich annimmt. Im Fall einer diskreten Gleichverteilung auf einer endlichen Menge $\Omega = \{x_1, \dots, x_n\} \subseteq \mathbb{R}$ ist damit das arithmetische Mittel der Werte gemeint, also $\mathbb{E}(\mathbb{W}) = \frac{1}{n} \sum_{i=1}^n x_i$. Sind auf dieser Menge die Wahrscheinlichkeiten nicht alle $= \frac{1}{n}$, sondern $\mathbb{W}(x_i) = p_i \geq 0$ mit $\sum_{i=1}^n p_i = 1$, so ist das arithmetische Mittel durch das gewichtete Mittel

$$\mathbb{E}(\mathbb{W}) = \sum_{i=1}^n p_i x_i$$

zu ersetzen. Wir erinnern uns an Mathematik 1 und die Integration nach einem Maß. In unserem Fall liegt das Maß \mathbb{W} vor. Weil es auf der endlichen Menge Ω konzentriert ist,

wird ein Integral über eine Funktion $f : \Omega \rightarrow \mathbb{R}$ zur Summe:

$$\int_{\Omega} f d\mathbb{W} = \sum_{i=1}^n p_i f(x_i).$$

Also ist der Erwartungswert gleich dem Integral über die Funktion $f(x) := x$, als Formel:

$$\mathbb{E}(\mathbb{W}) = \int_{\Omega} x d\mathbb{W}(x) = \int_{\mathbb{R}} x d\mathbb{W}(x).$$

Die letzte Gleichheit gilt deshalb, weil die gesamte W -Masse von \mathbb{W} auf der Teilmenge $\Omega \subseteq \mathbb{R}$ konzentriert ist.

Diese Formel als Umschreibung einer endlichen Summe mag ungewohnt erscheinen. Ihr Nutzen wird deutlich, wenn wir als \mathbb{W} zum Beispiel die stetige Gleichverteilung auf dem Intervall $[a, b]$ betrachten, wo der Erwartungswert das arithmetische Mittel $\frac{a+b}{2}$ der Intervallgrenzen sein sollte. Für diese Verteilung steht die Dichtefunktion $\rho(x) = \frac{1}{b-a}$ für $x \in [a, b]$ und $\rho(x) = 0$ zur Verfügung. Ein Integral über eine reelle Funktion f bezüglich dieser Verteilung \mathbb{W} nimmt die Form

$$\int_{\mathbb{R}} f d\mathbb{W} = \int_{\mathbb{R}} f(x)\rho(x) dx$$

an, was für die Funktion $f(x) = x$ mit Stammfunktion $F(x) = \frac{x^2}{2}$ zu

$$\int_a^b \frac{x}{b-a} dx = \frac{1}{b-a}(F(b) - F(a)) = \frac{b^2 - a^2}{2(b-a)} = \frac{(b+a)(b-a)}{2(b-a)} = \frac{a+b}{2}$$

wird. Wir haben also tatsächlich den erwünschten Erwartungswert erhalten.

Solche Berechnungen sind kein Spezifikum der Gleichverteilung, sondern für allgemeine Verteilungen \mathbb{W} sinnvoll. Man definiert daher:

Definition 5.1.5.1. Sei \mathbb{W} eine W -Verteilung auf \mathbb{R} , für die das Integral

$$\mathbb{E}(\mathbb{W}) := \int_{\mathbb{R}} x d\mathbb{W}(x)$$

als reelle Zahl existiert. Dann heißt dieser Wert **Erwartungswert** oder auch **Mittelwert** von \mathbb{W} .

Ist X eine Zufallsgröße auf einem beliebigen W -Raum mit Verteilung \mathbb{W} und existiert $\mathbb{E}(\mathbb{W})$, so nennt man diesen Wert auch **Erwartungswert** oder **Mittelwert** von X und schreibt dafür $\mathbb{E}(X)$.

Sei die Zufallsgröße X auf dem W -Raum $(\Omega, \mathcal{A}, \mathbb{W}_{\Omega})$ definiert und $\mathbb{W}_{\mathbb{R}}$ die Verteilung von X auf \mathbb{R} . Dann lässt sich schreiben:

$$\mathbb{E}(X) := \int_{\Omega} X d\mathbb{W}_{\Omega} = \int_{\Omega} X(\omega) d\mathbb{W}_{\Omega}(\omega) = \int_{\mathbb{R}} x d\mathbb{W}_{\mathbb{R}}(x) = \mathbb{E}(\mathbb{W}_{\mathbb{R}})$$

Wir werden den Unterscheidungen, die diesen verschiedenen Schreibweisen zugrunde liegen, nicht sehr viel Aufmerksamkeit schenken. Wertvoll sind diese Möglichkeiten vor

allem für eine einheitliche Behandlung aller Situationen. In den meisten Fällen liegen aber Verteilungen vor, die entweder rein diskret sind, oder für die es eine Dichtefunktion gibt. Für diskrete Verteilungen können wir genauso mit Summen arbeiten:

$$\mathbb{E}(X) = \sum_i p_i x_i,$$

wenn X die Werte $x_i \in \mathbb{R}$ mit Wahrscheinlichkeit p_i annimmt. Für Verteilungen mit Dichtefunktion ρ wiederum gilt:

$$\mathbb{E}(X) = \int_{\mathbb{R}} \rho(x) \cdot x \, dx$$

Man beachte die Analogie der beiden Formeln, wobei die Dichtefunktion ρ an die Stelle der p_i tritt. Die Definition 5.1.5.1 erfasst beide Situationen (und noch viele andere) mit Hilfe des gemeinsamen Konzeptes der Integration nach einem Maß.

Beispiele zum Erwartungswert:

- (a) Sei \mathbb{W} das Wahrscheinlichkeitsmaß in einem Laplace Experiment auf $\{1, \dots, n\}$. Dann gilt

$$\mathbb{E}(\mathbb{W}) = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + \dots + n \cdot \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n i = \frac{n+1}{2}.$$

- (b) Bezeichnet X die Indikatorfunktion von $A \in \mathcal{A}$ über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{W})$, so ist

$$\mathbb{E}X = 0 \cdot \mathbb{W}(\{X = 0\}) + 1 \cdot \mathbb{W}(\{X = 1\}) = \mathbb{W}(A).$$

- (c) Nehme die Zufallsgröße X jeden Wert 2^n , $n = 1, 2, \dots$, mit der Wahrscheinlichkeit 2^{-n} an. Dann ist der Erwartungswert

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} 2^{-n} \cdot 2^n = \sum_{n=1}^{\infty} 1 = \infty$$

unendlich.

- (d) Wir nehmen in (c) die Wahrscheinlichkeiten $2^{-(n+1)}$ statt 2^{-n} , dafür sowohl für den positiven Wert 2^n als auch für den negativen Wert -2^n . Dann wird der Erwartungswert

$$\mathbb{E}(X) = \sum_{n=1}^{\infty} 2^{-(n+1)} \cdot 2^n - \sum_{n=1}^{\infty} 2^{-(n+1)} \cdot 2^n = \infty - \infty$$

überhaupt sinnlos. (Man erinnere sich an die Definition des Lebesgue-Integrals mit den Fallunterscheidungen bezüglich $\pm\infty$.) Eine Zufallsgröße muss also keinen Erwartungswert haben.

5.1.6 Varianz

Inhalt in Kurzfassung: Die Varianz einer Verteilung oder einer Zufallsgröße wird als der Erwartungswert des Quadrats der Abweichung vom Erwartungswert definiert. Sie ist ein Maß für die Streuung um den Mittelwert.

War der Erwartungswert eine Kenngröße für den mittleren Wert einer \mathbb{W} -Verteilung bzw. Zufallsgröße, so beschreibt die Varianz, wie stark die Verteilung um diesen Mittelwert streut. Aus technischen Gründen, die später klar werden, erweist es sich als zweckmäßig, das Quadrat der Abweichung zu messen. Man definiert also:

Definition 5.1.6.1. Sei \mathbb{W} eine \mathbb{W} -Verteilung auf \mathbb{R} , für die $\mu := \mathbb{E}(\mathbb{W})$ existiert. Existiert auch das Integral

$$\mathbb{V}(\mathbb{W}) := \int_{\mathbb{R}} (x - \mu)^2 d\mathbb{W}(x),$$

so heißt dieser Wert $\mathbb{V}(\mathbb{W})$ die **Varianz** von \mathbb{W} .

Ist X eine Zufallsgröße mit einer Verteilung \mathbb{W} , für die $\mathbb{V}(\mathbb{W})$ existiert, so heißt dieser Wert auch die Varianz von X und wird auch mit $\mathbb{V}(X)$ bezeichnet.

Die Quadratwurzel der Varianz nennt man auch **Standardabweichung** einer Verteilung bzw. einer Zufallsgröße.

Offenbar lässt sich die Varianz $\mathbb{V}(X)$ einer Zufallsgröße X auch als Erwartungswert der Verteilung der Zufallsgröße $(X - \mathbb{E}(X))^2$ schreiben:

$$\mathbb{V}(X) = \mathbb{E}(X - \mathbb{E}(X))^2$$

Hat X eine diskrete Verteilung mit Wahrscheinlichkeiten $p_i = \mathbb{W}(x_i)$ und Erwartungswert $\mathbb{E}(X) = \mu$, so bedeutet das:

$$\mathbb{V}(X) = \sum_i p_i (x_i - \mu)^2$$

Für eine Zufallsgröße mit Erwartungswert μ , zu deren Verteilung es eine Dichte ρ gibt, gilt

$$\mathbb{V}(X) = \int_{\mathbb{R}} (x - \mu)^2 \rho(x) dx.$$

Hätten wir in Definition 5.1.6.1 statt des Quadrats den Absolutbetrag der Abweichung vom Mittelwert genommen, so hätten wir die **mittlere absolute Abweichung**

$$\int_{\mathbb{R}} |x - \mu| d\mathbb{W}(x),$$

erhalten.

5.1.7 Bedingte Wahrscheinlichkeit, Unabhängigkeit und Produkträume

Inhalt in Kurzfassung: Die bedingte Wahrscheinlichkeit $\mathbb{W}(A|B)$ eines Ereignisses A unter der Bedingung B gibt die Wahrscheinlichkeit von A unter der Voraussetzung an, dass B eingetreten ist. Diese Wahrscheinlichkeit ist der Quotient $\frac{\mathbb{W}(A \cap B)}{\mathbb{W}(B)}$. Für Ereignisse, die keinen Einfluss aufeinander haben, stimmen $\mathbb{W}(A)$ und $\mathbb{W}(A|B)$ überein. Das bedenkend nennt man sie unabhängig, wenn $\mathbb{W}(A \cap B) = \mathbb{W}(A) \cdot \mathbb{W}(B)$ gilt. Für Familien von mehr als zwei Ereignissen muss man Unabhängigkeit stärker formulieren, entsprechend für Zufallsgrößen statt Ereignissen. Typische Beispiele für unabhängige Zufallsgrößen treten bei mehrmals hintereinander ausgeführten Zufallsexperimenten auf, sofern die einzelnen Experimente keinen Einfluss aufeinander haben. Zur Modellierung unabhängiger Zufallsgrößen im Rahmen eines einzigen \mathbb{W} -Raumes eignet sich der Produktraum.

Die bedingte Wahrscheinlichkeit eines Ereignisses A unter der Bedingung B soll die Wahrscheinlichkeit von A angeben, wenn bereits bekannt ist, dass B eingetreten ist. Wir wollen das illustrieren, indem wir den \mathbb{W} -Raum mit der diskreten Gleichverteilung auf $\Omega = \{1, 2, 3, 4, 5, 6\}$ für einmaliges Würfeln heranziehen. Bezeichne A und B die Ereignisse, eine Primzahl bzw. eine ungerade Zahl zu würfeln. Wissen wir bereits, dass B eingetreten ist, so geht es nur mehr darum, mit welcher Wahrscheinlichkeit ein beliebiges Element aus $B = \{1, 3, 5\}$ eine Primzahl ist. Darunter sind die günstigen Ereignisse 3 und 5, also die zwei Elemente aus $A \cap B$. Folglich ist die gesuchte Wahrscheinlichkeit $\frac{2}{3}$ (2 günstige von 3 möglichen Ereignissen). Wir haben dabei nichts anderes getan, als die Wahrscheinlichkeit von $A \cap B$ durch die von B zu dividieren. Die Definition lautet daher:

Definition 5.1.7.1. *Es sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{A}$ mit $\mathbb{W}(B) \neq 0$. Dann ist die **bedingte Wahrscheinlichkeit** $\mathbb{W}(A|B)$ von A unter der Bedingung B definiert durch*

$$\mathbb{W}(A|B) := \frac{\mathbb{W}(A \cap B)}{\mathbb{W}(B)}.$$

Eine besondere Rolle spielen Ereignisse, die einander nicht beeinflussen, wo also die Wahrscheinlichkeit, dass A eintritt, unabhängig davon ist, ob die Bedingung B eingetreten ist:

$$\mathbb{W}(A|B) = \mathbb{W}(A)$$

Nach Definition 5.1.7.1 ist das für $\mathbb{W}(B) \neq 0$ äquivalent zu $\mathbb{W}(A \cap B) = \mathbb{W}(A) \cdot \mathbb{W}(B)$. Entsprechend definiert man:

Definition 5.1.7.2. *Es sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein Wahrscheinlichkeitsraum. Dann heißen zwei Ereignisse $A, B \in \mathcal{A}$ **unabhängig**, wenn*

$$\mathbb{W}(A \cap B) = \mathbb{W}(A) \cdot \mathbb{W}(B)$$

gilt.

Eine **Familie von Ereignissen** $A_i \in \mathcal{A}$, $i \in I$ (Indexmenge), heißt *unabhängig*, wenn für je endlich viele paarweise verschiedene Indizes $i_1, \dots, i_n \in I$

$$\mathbb{W}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_n}) = \mathbb{W}(A_{i_1}) \cdot \mathbb{W}(A_{i_2}) \cdot \dots \cdot \mathbb{W}(A_{i_n})$$

gilt.

Eine Familie von **Zufallsgrößen** $X_i : \Omega \rightarrow \mathbb{R}$, $i \in I$, heißt *unabhängig*, wenn für alle messbaren Teilmengen A_i von \mathbb{R} die Ereignisse $(X_i \in A_i)$, $i \in I$, unabhängig sind.

Es ist klar, dass in einer unabhängigen Familie von Ereignissen auch je zwei verschiedene unabhängig sind ($n = 2$ setzen). Bei der Umkehrung ist aber Vorsicht geboten. Dazu betrachten wir diskrete Gleichverteilung auf der 8-elementigen Menge $\Omega := \{1, 2, 3, 4, 5, 6, 7, 8\}$ und die drei Mengen $A := \{1, 2, 3, 4\}$, $B := \{1, 2, 5, 6\}$ und $C := \{3, 4, 5, 6\}$. Dann ist $A \cap B = \{1, 2\}$, $A \cap C = \{3, 4\}$, $B \cap C = \{5, 6\}$ und $A \cap B \cap C = \emptyset$, also mit $\mathbb{W}(A) = \mathbb{W}(B) = \mathbb{W}(C) = \frac{4}{8} = \frac{1}{2}$ und

$$\begin{aligned}\mathbb{W}(A \cap B) &= \frac{2}{8} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{W}(A) \cdot \mathbb{W}(B) \\ \mathbb{W}(A \cap C) &= \frac{2}{8} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{W}(A) \cdot \mathbb{W}(C) \\ \mathbb{W}(B \cap C) &= \frac{2}{8} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{W}(B) \cdot \mathbb{W}(C).\end{aligned}$$

Folglich sind je zwei dieser Ereignisse unabhängig. Hingegen sind wegen

$$\mathbb{W}(A \cap B \cap C) = \frac{0}{8} = 0 \neq \frac{1}{8} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \mathbb{W}(A) \cdot \mathbb{W}(B) \cdot \mathbb{W}(C),$$

die drei Ereignisse A, B und C als dreielementige Familie nicht unabhängig.

Die typische Situation, wo unabhängige Ereignisse und Zufallsgrößen eine entscheidende Rolle übernehmen, sind Räume von endlichen und unendlichen Folgen, wie wir sie in 5.1.2 behandelt haben. Daran wollen wir anschließen, verallgemeinern und Zufallsgrößen ins Spiel bringen.

Wir gehen aus von Zufallsgrößen X_n , $n \in \mathbb{N}^+$, die auf W-Räumen $(\Omega_n, \mathcal{A}_n, \mathbb{W}_n)$ definiert sind. Wir konstruieren aus diesen Räumen nun den Produktraum $(\Omega, \mathcal{A}, \mathbb{W})$. Und zwar ist Ω das kartesische Produkt der Ω_n , enthält als Elemente also alle Folgen $\omega = (\omega_1, \omega_2, \dots)$ mit $\omega_n \in \Omega_n$. Für eine Menge $A_n \in \mathcal{A}_n$ bezeichne $[A_n]$ die Menge aller Folgen $\omega \in \Omega$, deren n -te Komponente in A_n liegt. Das W-Maß \mathbb{W} ist so definiert, dass stets $\mathbb{W}([A_n]) = \mathbb{W}_n(A_n)$ gilt, außerdem

$$\mathbb{W}([A_{n_1}] \cap [A_{n_2}] \cap \dots \cap [A_{n_k}]) = \mathbb{W}_{n_1}(A_{n_1}) \cdot \mathbb{W}_{n_2}(A_{n_2}) \cdot \dots \cdot \mathbb{W}_{n_k}(A_{n_k})$$

für alle Auswahlen endlich vieler paarweise verschiedener Indizes n_i und $A_{n_i} \in \mathcal{A}_{n_i}$. Es lässt sich zeigen, dass es eine geeignete σ -Algebra \mathcal{A} auf Ω und darauf ein W-Maß \mathbb{W} mit den behaupteten Eigenschaften tatsächlich gibt. Der so definierte W-Raum $(\Omega, \mathcal{A}, \mathbb{W})$ heißt der **Produktraum** der $(\Omega_n, \mathcal{A}_n, \mathbb{W}_n)$ und entspricht im Wesentlichen dem Raum,

den wir schon in 5.1.2 unter „unendliche Versuchswiederholungen“ betrachtet haben. Er dient zur Modellierung von unendlichen Folgen von unabhängigen Zufallsgrößen. Der Bezug ergibt sich daraus: \mathbb{W} ist gerade so definiert, dass alle Familien von Mengen $[A_n]$ mit $A_n \in \mathcal{A}_n$ und $n \in \mathbb{N}^+$ unabhängig sind. Die ursprünglich gegebenen Zufallsgrößen $X_n : \Omega_n \rightarrow \mathbb{R}$ kann man auch als Zufallsgrößen X'_n auf dem gemeinsamen Produktraum interpretieren, indem man

$$X'_n : \Omega \rightarrow \mathbb{R}, \quad \omega = (\omega_1, \omega_2, \dots) \mapsto X_n(\omega_n)$$

setzt. Ist A eine messbare Menge in \mathbb{R} , so gilt

$$X_n'^{(-1)}(A) = [X_n^{(-1)}(A)], \quad n \in \mathbb{N}^+,$$

und solche Familien von Mengen sind unabhängig. Folglich sind auch die Zufallsgrößen X'_n unabhängig. Ist von Folgen unabhängiger Zufallsgrößen die Rede, kann man sich als mathematische Modellierung stets diese Konstruktion eines Produktraumes vorstellen.

Übungsaufgabe 283. (T) Aus einer Urne mit den vier (nummerierten) Kugeln 1, 2, 3 und 4 wird zwei Mal mit Zurücklegen gezogen, d.h. der Ergebnisraum ist $\Omega = \{(a_1, a_2) : 1 \leq a_i \leq 4, i = 1, 2\}$.

- A ist das Ereignis, dass im ersten Zug eine ungerade Zahl gezogen wird.
- B ist das Ereignis, dass im zweiten Zug eine gerade Zahl gezogen wird.
- C ist das Ereignis, dass die Summe der gezogenen Zahlen 2 ist.
- D ist das Ereignis, dass die Summe der gezogenen Zahlen 5 ist.

Berechnen Sie $\mathbb{W}(A)$, $\mathbb{W}(B)$, $\mathbb{W}(C)$, $\mathbb{W}(D)$, $\mathbb{W}(A \cap B)$, $\mathbb{W}(A \cap C)$, $\mathbb{W}(B \cap D)$. Erklären Sie Ihr Ergebnis speziell im Hinblick auf stochastische Unabhängigkeit.

Übungsaufgabe 284. (E) Das in dieser Übungsaufgabe zu behandelnde Beispiel stammt aus einer Fernsehshow vor Jahren und erlangte damals große Bekanntheit. Dabei konfrontierte der Fernsehmoderator einen Kandidaten mit drei verschlossenen Türen. Dem Kandidaten wurde verraten, dass hinter einer dieser Türen ein Auto, hinter den anderen beiden je eine Ziege steht. Unter der stillschweigenden Annahme, dass Teilnehmer an einer Fernsehshow leidenschaftliche Autofahrer sind, aber kein Interesse an Ziegenhaltung haben, wurde dem Kandidaten in Aussicht gestellt, dass das Auto ihm gehöre, wenn er die richtige Tür errät. Der Kandidat zeigte auf eine der drei Türen. Der Moderator öffnete eine der anderen beiden Türen, hinter der man eine Ziege sah, und stellte dem Kandidaten frei, seine Wahl noch zu korrigieren. Was würden Sie dem Kandidaten (unter der Annahme, dass er Autofan und Ziegenhasser ist) empfehlen?

Übungsaufgabe 285. (T) In einem Kartenspiel mit 32 Karten sind 8 Karten der Farbe Pik. Es werden zufällig (d.h. unter der Laplace-Annahme) 4 Karten ohne Zurücklegen gezogen. Berechnen Sie mit Hilfe der Formel für bedingte Wahrscheinlichkeiten die Wahrscheinlichkeit dafür, dass alle 4 Karten Pik sind.

5.1.8 Die Formel von Bayes

Inhalt in Kurzfassung: Bilden die B_k eine Zerlegung eines \mathbb{W} -Raumes und A ein weiteres Ereignis, so liefert die Formel von Bayes die bedingten Wahrscheinlichkeiten $\mathbb{W}(B_k|A)$, sofern man die Wahrscheinlichkeiten $\mathbb{W}(B_k)$ sowie die bedingten Wahrscheinlichkeiten $\mathbb{W}(A|B_k)$ kennt. Die Formel von Bayes liefert in vielen Anwendungssituationen nützliche und gelegentlich auch verblüffende Ergebnisse.

Die **Formel von Bayes** ermöglicht es, beim Vorliegen entsprechender Informationen, bedingte und bedingende Ereignisse zu vertauschen. Er lautet:

Satz 5.1.8.1. *Es sei $(\Omega, \mathcal{A}, \mathbb{W})$ ein Wahrscheinlichkeitsraum und $B_1, B_2, \dots \in \mathcal{A}$ eine Zerlegung von Ω in endlich oder abzählbar unendlich viele paarweise disjunkte messbare Teilmengen. Dann gilt für jedes $A \in \mathcal{A}$ mit $\mathbb{W}(A) > 0$ und alle $k = 1, 2, \dots$*

$$\mathbb{W}(B_k|A) = \frac{\mathbb{W}(B_k)\mathbb{W}(A|B_k)}{\sum_n \mathbb{W}(B_n)\mathbb{W}(A|B_n)}.$$

Beweis. Zunächst gilt

$$A = A \cap \Omega = A \cap \bigcup_n B_n = \bigcup_n (A \cap B_n).$$

Da diese Vereinigung aus paarweise disjunkten Mengen besteht, erhalten wir den sogenannten Satz von der totalen Wahrscheinlichkeit:

$$\mathbb{W}(A) = \sum_n \mathbb{W}(A \cap B_n) = \sum_n \mathbb{W}(B_n)\mathbb{W}(A|B_n).$$

Setzen wir das in

$$\mathbb{W}(B_k|A) = \frac{\mathbb{W}(B_k \cap A)}{\mathbb{W}(A)} = \frac{\mathbb{W}(B_k)\mathbb{W}(A|B_k)}{\mathbb{W}(A)}$$

für den Nenner auf der rechten Seite ein, erhalten wir die Behauptung. \square

Die Nützlichkeit der Formel von Bayes wird bei Anwendungen wie der folgenden typischen deutlich:

Eine seltene Krankheit kann mit einem Test diagnostiziert werden, der allerdings gewisse Fehlerraten aufweist. Angenommen es sei bekannt, dass innerhalb der gesamten Bevölkerung nur eine Person von 100000 diese Krankheit in sich trägt. Bei 99% der Erkrankten liefere der Test ein (korrektes) positives Ergebnis, bei 1% der Kranken ein (falsches) negatives. Außerdem sei bekannt, dass bei Gesunden das Testergebnis mit 90% (korrekt) negativ, jedoch bei 10% (falsch) positiv ausfalle. Wir modellieren mit einem \mathbb{W} -Raum $(\Omega, \mathcal{A}, \mathbb{W})$, der $\mathbb{W}(K) = 10^{-5}$ für die Menge K der Kranken und entsprechend $\mathbb{W}(G) = 99999 \cdot 10^{-5}$ für die Menge G der Gesunden liefert. Außerdem halten wir die bedingten Wahrscheinlichkeiten für P (positiver Test) und N (negativer Test) unter den Bedingungen K und G fest: $\mathbb{W}(P|K) = 0,99$, $\mathbb{W}(N|K) = 0,01$, $\mathbb{W}(N|G) = 0,9$ und $\mathbb{W}(P|G) = 0,1$. Welche Perspektive hat man, wenn man positiv getestet worden ist?

Wenn es sonst keine Indizien für die Krankheit gibt, dann eine bessere als man auf den ersten Blick vielleicht vermuten würde! Für die bedingte Wahrscheinlichkeit $\mathbb{W}(K|P)$, bei positivem Test tatsächlich krank zu sein, liefert die Formel von Bayes (wir haben dort $A = P$, $B_1 = K$ und $B_2 = G$ zu setzen) nämlich nicht mehr als

$$\begin{aligned}\mathbb{W}(K|P) &= \frac{\mathbb{W}(K) \cdot \mathbb{W}(P|K)}{\mathbb{W}(K) \cdot \mathbb{W}(P|K) + \mathbb{W}(G) \cdot \mathbb{W}(P|G)} = \\ &= \frac{10^{-5} \cdot 99 \cdot 10^{-2}}{10^{-5} \cdot 99 \cdot 10^{-2} + 99999 \cdot 10^{-5} \cdot 10^{-1}} = \frac{99}{99 + 999990} = \frac{99}{1000089} \approx 0,0001.\end{aligned}$$

Man macht sich die Situation schnell klar, wenn man an die absoluten Zahlen denkt, z.B. bei einer Gesamtpopulation von 1000000. Dann gibt es nur etwa 10 Kranke (die wahrscheinlich alle positiv getestet würden) und entsprechend 999990 Gesunde, von denen immerhin annähernd 100000 (falsch) positiv getestet würden. Bei einer vollständigen Testung der Gesamtpopulation gäbe es also etwa 10000-mal so viele Gesunde mit positivem Test wie Kranke. Dennoch ist Vorsicht geboten: Wenn man sich zum Beispiel deshalb einem Test unterzogen hat, weil man gewisse Symptome hat, die mit der seltenen Krankheit einhergehen, dann wird die obige Modellierung fragwürdig. Man müsste sich nämlich auf die Grundgesamtheit jener Personen mit diesen Symptomen beschränken, was zu ganz anderen Zahlenwerten führen kann, besonders wenn nur wenige Gesunde diese Symptome aufweisen.

Übungsaufgabe 286. (T) *Es wurde ein neuer Alkohol-Test für das nächste BIZ Fest entwickelt. Man geht davon aus (auf Erfahrungswerten basierend), dass 60% der getesteten Besucher zumindest leicht angetrunken sind.*

Bezüglich der Funktionsweise des Tests wurde ermittelt, dass

- *in 95% der Fälle der Test positiv ist, wenn die getestete Person tatsächlich angetrunken ist und*
- *in 90% der Fälle der Test negativ ist, wenn die Person nicht betrunken ist.*

Wie groß ist die Wahrscheinlichkeit, dass die getestete Person betrunken ist, wenn der Test positiv ist?

5.2 Wichtige Ergebnisse der Wahrscheinlichkeitstheorie

In 5.2.1 werden die wichtigsten Rechenregeln, auf denen die Hauptergebnisse der W-Theorie beruhen, hergeleitet. Vor allem geht es um die Additivität der Varianz unabhängiger Zufallsgrößen. In 5.2.2 und 5.2.3 werden einige wichtige diskrete bzw. stetige reelle Verteilungen mit ihren Erwartungswerten und Varianzen vorgestellt. Das erste Hauptergebnis der W-Theorie, nämlich das Gesetz der großen Zahlen ist Gegenstand von 5.2.4. Eine ziemlich direkte Folgerung ist der Hauptsatz der mathematischen Statistik in 5.2.5. Der Abschnitt wird in 5.2.6 mit dem zweiten Hauptergebnis der W-Theorie, dem Zentralen Grenzwertsatz abgeschlossen.

5.2.1 Rechenregeln für Erwartungswert und Varianz

Inhalt in Kurzfassung: Ziel ist vor allem die Rechenregel $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$ für unabhängige Zufallsgrößen X und Y . Die wichtigste Zwischenstation beim Beweis ist die Multiplikativität $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$ des Erwartungswerts bei Unabhängigkeit.

Weil der Erwartungswert von Zufallsgrößen über ein Integral definiert ist, vererbt sich dessen Linearität. Für zwei Zufallsgrößen X und Y auf demselben W -Raum und für beliebige Zahlen a und b gilt daher $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$, insbesondere $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$. Die entsprechende Regel für die Multiplikation gilt natürlich nicht, weil das Integral eines Produktes im Allgemeinen nicht das Produkt der Integrale ist. Für die W -Theorie fundamental ist aber die Tatsache, dass der Erwartungswert zweier *unabhängiger* Zufallsgrößen X und Y sehr wohl multiplikativ ist:

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Sind X und Y auf einem endlichen Raum definiert, rechnet man das unmittelbar über die entsprechenden Summen nach. Entscheidend ist dabei das Distributivgesetz. Bei abzählbar unendlichem Raum hat man entsprechend Cauchyprodukte zu bemühen. Für den allgemeinen Fall verwendet man, dass unabhängige Zufallsgrößen mit Hilfe des Produktraums modelliert werden: Für Produkträume gilt die entsprechend verallgemeinerte Variante des Satzes von Fubini. Man kann dann nämlich, ähnlich wie wir das in 5.1.7 getan haben, X und Y als Funktionen in zwei Variablen auffassen, wobei aber X nur von der ersten, Y nur von der zweiten Komponente abhängt. Damit bekommt man in sehr schematischer Schreibweise:

$$\begin{aligned} \mathbb{E}(X \cdot Y) &= \int X(\omega_1, \omega_2) \cdot Y(\omega_1, \omega_2) d\mathbb{W}(\omega_1, \omega_2) = \\ &= \int \int X(\omega_1) \cdot Y(\omega_2) d\mathbb{W}_1(\omega_1) d\mathbb{W}_2(\omega_2) = \\ &= \int X(\omega_1) \left(\int Y(\omega_2) d\mathbb{W}_2(\omega_2) \right) d\mathbb{W}_1(\omega_1) = \\ &= \left(\int X(\omega_1) d\mathbb{W}_1(\omega_1) \right) \cdot \left(\int Y(\omega_2) d\mathbb{W}_2(\omega_2) \right) = \\ &= \mathbb{E}(X) \cdot \mathbb{E}(Y). \end{aligned}$$

Sind X und Y unabhängig, so auch $X - \mathbb{E}(X)$ und $Y - \mathbb{E}(Y)$, folglich gilt

$$\mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(X - \mathbb{E}(X)) \cdot \mathbb{E}(Y - \mathbb{E}(Y)) = 0,$$

weil beide Faktoren 0 sind.

Man nennt den hier auftretenden Ausdruck

$$\text{Cov}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$$

übrigens auch **Kovarianz** von X und Y . Fasst man Zufallsgrößen mit endlicher Varianz als Vektoren eines Vektorraums auf (genauer sind die Elemente dieses Vektorraums

Äquivalenzklassen von fast sicher gleichen Zufallsgrößen), so definiert $\mathbb{E}(X \cdot Y)$ ein Skalarprodukt mit Norm $\|X\| := \sqrt{\mathbb{E}(X^2)}$. Somit gilt die Cauchy-Schwarzsche Ungleichung, die, angewendet auf $X - \mathbb{E}(X)$ und $Y - \mathbb{E}(Y)$ die Ungleichung

$$|\text{Cov}(X, Y)| = |\mathbb{E}((X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y)))| \leq \|X - \mathbb{E}(X)\| \cdot \|Y - \mathbb{E}(Y)\| = \sqrt{\mathbb{V}(X) \cdot \mathbb{V}(Y)}$$

liefert. Folglich ist (für $\mathbb{V}(X), \mathbb{V}(Y) \neq 0$) der sogenannte **Korrelationskoeffizient**

$$\text{Kor}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\mathbb{V}(X) \cdot \mathbb{V}(Y)}}$$

eine reelle Zahl zwischen -1 und 1 . Gemäß dem Kriterium, wann in der Cauchy-Schwarzschen Ungleichung Gleichheit auftritt (siehe Satz 1.4.3.1), ist $\text{Kor}(X, Y) = 1$ genau dann, wenn $X = Y$ fast sicher und $\text{Kor}(X, Y) = -1$ genau dann, wenn $X = -Y$. Im Fall $\text{Cov}(X, Y) = 0$ heißen X und Y **unkorreliert**.

Die obige Rechnung zeigt somit, dass zwei unabhängige Zufallsgrößen X und Y mit endlichen Varianzen $\mathbb{V}(X), \mathbb{V}(Y)$ stets unkorreliert sind. Die Umkehrung gilt im Allgemeinen nicht, wie die folgende Übungsaufgabe zeigt.

Übungsaufgabe 287. (E) Konstruieren Sie zwei unkorrelierte Zufallsgrößen, die nicht unabhängig sind. Hinweis: Wählen Sie als zugrundeliegende Menge Ω von Elementarereignissen die Einheitskreislinie $\{(x, y) : x^2 + y^2 = 1\}$ mit stetiger Gleichverteilung (d.h. gleiche Bogenlängen haben gleiches Maß) und als Zufallsgrößen X und Y die Projektionen auf die x - bzw. auf die y -Achse.

Für unkorrelierte und somit erst recht für unabhängige Zufallsgrößen gilt:

$$\begin{aligned} \mathbb{V}(X + Y) &= \mathbb{E}((X + Y) - \mathbb{E}(X + Y))^2 = \\ &= \mathbb{E}((X - \mathbb{E}(X)) + (Y - \mathbb{E}(Y)))^2 = \\ &= \mathbb{E}\left((X - \mathbb{E}(X))^2 + 2(X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y)) + (Y - \mathbb{E}(Y))^2\right) = \\ &= \mathbb{E}(X - \mathbb{E}(X))^2 + 2\mathbb{E}(X - \mathbb{E}(X))\mathbb{E}(Y - \mathbb{E}(Y)) + \mathbb{E}(Y - \mathbb{E}(Y))^2 = \\ &= \mathbb{V}(X) + \mathbb{V}(Y) \end{aligned}$$

Mittels Induktion lässt sich die Additivität der Varianz von zwei auf n unabhängige Zufallsgrößen ausdehnen. Es gilt also:

Satz 5.2.1.1. Sind X_1, \dots, X_n paarweise unkorrelierte (z.B. also paarweise unabhängige) Zufallsgrößen auf demselben W -Raum, deren Varianzen existieren. Dann gilt:

$$\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i)$$

Übungsaufgabe 288. (T*) Sei X eine Zufallsgröße, welche eine bestimmte Eigenschaft einer Grundmenge Ω misst. μ ist der Erwartungswert von X und σ^2 die Varianz. Man führt nun eine vierfache Messung durch: X_1, X_2, X_3 und X_4 sind die Zufallsgrößen, welche den Wert der jeweiligen Messung angeben. Folglich gilt $\mathbb{E}(X_i) = \mu$ und $\mathbb{V}(X_i) = \sigma^2$ für $i = 1, 2, \dots, 4$. Gegeben sind nun folgende Zufallsgrößen (Schätzer für μ):

- $T_1(X_1, \dots, X_4) = \frac{1}{4}(X_1 + X_2 + X_3 + X_4)$.
- $T_2(X_1, \dots, X_4) = \frac{1}{2}(X_1 + X_4)$.
- $T_3(X_1, \dots, X_4) = \frac{1}{6}(X_1 + X_2) + \frac{2}{3}X_4$.

1. Zeigen Sie, dass T_1 , T_2 und T_3 „erwartungstreu“ für μ sind, d.h., dass $\mathbb{E}(T_i) = \mu$ gilt für $i = 1, 2, 3$.

2. Berechnen Sie $\mathbb{V}(T_i)$, $i = 1, 2, 3$. Interpretieren Sie Ihr Ergebnis.

Hinweis: Begründen Sie, dass die Zufallsvariablen X_i unabhängig sind. Für unabhängige Zufallsvariablen X und Y gilt bekanntlich $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$. Benutzen auch $\mathbb{V}(X_i) = \sigma^2$ für alle $i \geq 1$ bei der Berechnung.

5.2.2 Beispiele diskreter Verteilungen und ihre Erwartungswerte

Inhalt in Kurzfassung: Die behandelten Beispiele diskreter Verteilungen samt Erwartungswerten und Varianzen sind: diskrete Gleich-, Binomial-, geometrische, hypergeometrische und Poissonverteilung.

Folgenden diskrete Verteilungen kommen besonders häufig vor und seien daher samt ihren Erwartungswerten und Varianzen kurz vorgestellt:

- **Diskrete Gleichverteilung** oder **Laplaceverteilung** auf $\{1, 2, \dots, n\}$
Eine Zufallsvariable X heißt gleichverteilt auf $\{1, \dots, n\}$, $n \in \mathbb{N}$, wenn für alle $k = 1, \dots, n$

$$\mathbb{W}(\{X = k\}) = \frac{1}{n}.$$

Der Erwartungswert einer gleichverteilten Zufallsvariable X ist

$$\mathbb{E}(X) = \frac{n+1}{2},$$

ihre Varianz ist

$$\mathbb{V}(X) = \frac{n^2 - 1}{12}.$$

Übungsaufgabe 289. (E) Begründen Sie diese Formeln für Erwartungswert und Varianz nach.

- **Binomialverteilung** zu den Parametern n und p
Eine Zufallsgröße X heißt binomialverteilt zu den Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$, wenn für alle $k = 0, \dots, n$

$$\mathbb{W}(\{X = k\}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Für $n = 1$ wird die Binomialverteilung auch **Bernoulliverteilung** Verteilung zum Parameter p genannt.

Die Bedeutung von Binomial- und Bernoulliverteilung wird durch folgende Überlegung deutlich: Eine beliebige Verteilung auf der zweielementigen Menge $\Omega := \{0, 1\}$ ist eine Bernoulliverteilung, und zwar zum Parameter p , wenn p die Wahrscheinlichkeit für 1 ist. Ihr Erwartungswert ist durch $p \cdot 1 + (1 - p) \cdot 0 = p$ gegeben, ihre Varianz durch $p \cdot (1 - p)^2 + (1 - p) \cdot (0 - p)^2 = p \cdot (1 - p)$. Sind X_1, \dots, X_n unabhängige Bernoulli-verteilte Zufallsgröße zum Parameter p , so ist ihre Summe X binomialverteilt zu den Parametern n und p .

Übungsaufgabe 290. (E) Begründen Sie diesen Zusammenhang.

$\mathbb{W}(\{X = k\})$ berechnet also gerade die Wahrscheinlichkeit, bei n -facher Wiederholung genau k Einsen zu erhalten, sofern eine Eins die Wahrscheinlichkeit p hat.

Wegen der Additivität von Erwartungswert und, für unabhängige Zufallsgrößen, von Varianz, ergibt sich für Erwartungswert und Varianz einer Zufallsgröße X , die binomialverteilt ist zu den Parametern p und n :

$$\mathbb{E}(X) = np, \quad \mathbb{V}(X) = np(1 - p).$$

- **Geometrische Verteilung** zum Parameter p

Eine Zufallsvariable X heißt geometrisch verteilt zum Parameter $p \in [0, 1]$ auf \mathbb{N} , wenn für alle $k \in \mathbb{N}$

$$\mathbb{W}(\{X = k\}) = (1 - p)^{k-1} p.$$

Typischerweise kommen geometrisch verteilte Zufallsgrößen wie folgt zustande: Betrachten wir ein Zufallsexperiment mit zweielementiger Ergebnismenge $\{0, 1\}$, wobei 1 mit Wahrscheinlichkeit p und 0 mit Wahrscheinlichkeit $1 - p$ auftritt. Wir wiederholen dieses Experiment in unabhängiger Weise so lange, bis erstmals 1 auftritt. Die Zufallsvariable X , die angibt, wie oft beim wievielten Versuch das der Fall ist, ist geometrisch verteilt. $\mathbb{W}(\{X = k\})$ gibt also die Wahrscheinlichkeit an, dass die ersten $k - 1$ Wiederholungen nur Nullen liefern und im k -ten Versuch die Eins kommt. Erwartungswert und Varianz ergeben sich nach einigen Rechnungen, auf die wir hier verzichten als

$$\mathbb{E}X = \frac{1}{p} \quad \text{und} \quad \mathbb{V}X = \frac{1 - p}{p^2}.$$

- **Hypergeometrische Verteilung** zu den Parametern r, s und n

Es seien $r, s, n \in \mathbb{N}$ mit $r + s \geq n$. Eine Zufallsvariable X heißt hypergeometrisch verteilt zu den Parametern r, s, n , wenn für alle $k = 0, \dots, n$

$$\mathbb{W}(\{X = k\}) = \frac{\binom{r}{k} \cdot \binom{s}{n-k}}{\binom{r+s}{n}}.$$

Die natürliche Erklärung für hypergeometrische Verteilungen ergibt sich, wenn wir beim Bruch an die Regel „günstige durch mögliche Fälle“ denken. Im Nenner steht die Anzahl der n -elementigen Teilmengen einer $r + s$ -elementigen Menge. Im Zähler steht das Produkt aus den Anzahlen der k -elementigen Teilmengen einer r -elementigen und der $n - k$ -elementigen Teilmengen einer s -elementigen Menge. Dem entspricht offenbar ein Zufallsexperiment, bei dem n Elemente aus einer $(r + s)$ -elementigen Menge ohne Zurücklegen gezogen werden, wobei r Elemente Ausprägung 1 und s Elemente Ausprägung 0 haben. Dann ist die Zufallsvariable X , welche die Anzahl der Elemente mit Ausprägung 1 unter den n ausgewählten Elementen zählt, hypergeometrisch verteilt.

Ohne die aufwendige Berechnung von Erwartungswert und Varianz einer hypergeometrisch verteilten Zufallsvariablen X durchzuführen, seien hier lediglich die resultierenden Formeln angegeben:

$$\mathbb{E}(X) = \frac{nr}{r+s} \quad \text{und} \quad \mathbb{V}(X) = \frac{nrs(r+s-n)}{(r+s)^2(r+s-1)}.$$

- **Poissonverteilung** zum Parameter λ

Eine Zufallsvariable X heißt Poisson verteilt zum Parameter $\lambda > 0$ auf \mathbb{N} , wenn für alle $k \in \mathbb{N}$

$$\mathbb{W}(\{X = k\}) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Erwartungswert und Varianz einer Poisson verteilten Zufallsvariablen X sind (auch hier ohne Beweis) gegeben durch

$$\mathbb{E}(X) = \lambda \quad \text{und} \quad \mathbb{V}(X) = \lambda.$$

(Annähernde) Poissonverteilungen treten typischerweise auf, wenn gezählt wird, wieviele von sehr vielen möglichen unabhängigen, jeweils aber seltenen gedächtnislosen Ereignissen (siehe dazu Exponentialverteilung in 5.2.3) innerhalb eines bestimmten Zeitintervalls eintreten, zum Beispiel: Anzahl der in einem gegebenen Zeitintervall zerfallenden Teilchen einer radioaktiven Substanz oder der in einem Geschäft einlangenden Kunden o.ä. Für jedes Teilchen bzw. für jeden Menschen ist es sehr unwahrscheinlich zu zerfallen bzw. sich gerade in dieses Geschäft zu begeben. Da es aber sehr viele Teilchen bzw. Menschen gibt, kommen insgesamt doch ein paar (nämlich k) zusammen. Je größer der Parameter λ , desto seltener ist ein Einzelereignis.

Übungsaufgabe 291. (T) Betrachten Sie folgendes Spiel: Sie werfen eine faire Münze (Kopf oder Zahl). Dann werfen Sie nochmals: wenn Sie dasselbe Ergebnis wie beim ersten Wurf werfen, dürfen Sie weiterspielen; wenn nicht, ist das Spiel aus. Sie dürfen solange weiter werfen, solange Sie das Ergebnis ihres ersten Wurfs reproduzieren. Wenn Sie das erste mal „falsch“ werfen ist das Spiel aus. Angenommen Sie durften N mal werfen, dann beträgt ihr Gewinn $\frac{2^N}{2^N}$ Euro. (Sie gewinnen also in jedem Fall zumindest $\frac{2^2}{2 \cdot 2} = 1$ Euro.) Es bezeichne X die diskrete Zufallsgröße, welche ihren Gewinn angibt.

1. Finden Sie die Verteilung von X .
2. Berechnen Sie den Erwartungswert von X , d.h. jenen Einsatz den Sie zahlen müssten, damit das Spiel fair wird.

5.2.3 Beispiele stetiger Verteilungen und ihre Erwartungswerte

Inhalt in Kurzfassung: Die behandelten Beispiele stetiger Verteilungen sind: stetige Gleich-, Normal-, Exponential- und Cauchyverteilung.

Wir stellen im Folgenden vier interessante stetige (sogar absolut stetige) Verteilungen und, sofern vorhanden, ihre Erwartungswerte und Varianzen vor:

- **Stetige Gleichverteilung** auf einem Intervall $[a, b]$

Eine Zufallsvariable X heißt gleichverteilt auf $[a, b]$, falls die Verteilung von X die Dichte

$$p(x) = \begin{cases} \frac{1}{b-a} & \text{für } x \in [a, b] \\ 0 & \text{sonst} \end{cases}$$

hat. Wie man mit Hilfe der Dichtefunktion leicht nachrechnet, wie aber auch intuitiv unmittelbar klar ist, ist der Erwartungswert einer auf $[a, b]$ gleichverteilten Zufallsvariable X das arithmetische Mittel von a und b :

$$\mathbb{E}X = \frac{a+b}{2}$$

$$\mathbb{V}X = \frac{(a-b)^2}{12}.$$

Übungsaufgabe 292. (E) Führen Sie die Rechnungen, die zu diesen Formeln führen ausführlich durch.

- **Normalverteilung** zu den Parametern μ und σ

Eine Zufallsvariable X heißt normalverteilt zu den reellen Parametern μ und $\sigma > 0$, wenn die Verteilung von X die Dichte

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

hat. Die Dichte ist die berühmte Gauß'sche Glockenkurve. Sie ist offenbar symmetrisch um μ , weshalb gibt μ der Mittelwert=Erwartungswert der Verteilung ist, an dem die Dichtefunktion $\varphi_{\mu,\sigma}$ auch ihr Maximum annimmt. Weniger offensichtlich ist, dass σ die Streuung, d.h. σ^2 die Varianz dieser Verteilung ist. Für eine nach gemäß der Dichte $\varphi_{\mu,\sigma}$ verteilte Zufallsgröße X gilt also

$$\mathbb{E}(X) = \mu \quad \text{und} \quad \mathbb{V}(X) = \sigma^2.$$

Je größer σ , desto kleiner ist das Maximum und desto flacher und breiter ist die Kurve. Falls $\mu = 0$ und $\sigma = 1$ ist, nennt man X **standardnormalverteilt**. Die überragende Bedeutung der Normalverteilung wird durch den Zentralen Grenzwertsatz in 5.2.6 deutlich werden.

Ist X eine standardnormalverteilte Zufallsvariable, so ist $X^* := \sigma X + \mu$ normalverteilt zu den Parametern μ und σ . Umgekehrt lässt sich eine zu μ und σ normalverteilte Zufallsgröße X^* gemäß $X := \frac{X^* - \mu}{\sigma}$ auf Standardnormalverteilung transformieren. Die Verteilungsfunktion einer standardnormalverteilten Zufallsvariablen ist gegeben durch

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt.$$

Die Funktionswerte von Φ sind z.B. Tabellen wie im Anhang zu entnehmen, heutzutage aber natürlich auch in Softwarepaketen implementiert.

- **Exponentialverteilung** zum Parameter λ

Eine Zufallsvariable X heißt exponentialverteilt zum Parameter $\lambda > 0$, wenn die Verteilung von X die Dichte

$$\varrho(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

hat. Erwartungswert einer gemäß dieser Dichte verteilten Zufallsgröße X sind gegeben durch

$$\mathbb{E}X = \frac{1}{\lambda} \quad \text{und} \quad \mathbb{V}X = \frac{1}{\lambda^2}.$$

Exponentialverteilt sind sogenannte gedächtnislose Zufallsgrößen wie z.B. die Dauer X bis zum radioaktiven Zerfall eines Teilchens. Haben wir schon s Zeiteinheiten gewartet, ohne dass das Teilchen bisher zerfallen ist, so ändert das nichts an der Wahrscheinlichkeit, weitere t Zeiteinheiten warten zu müssen. Ist F die Verteilungsfunktion für die Zeitdauer bis zum Zerfall, so bedeutet das in Formelschreibweise:

$$\mathbb{W}(X \geq t + s | X \geq s) = \mathbb{W}(X \geq t) = 1 - F(t)$$

Nach der Definition bedingter Wahrscheinlichkeiten gilt

$$\mathbb{W}(X \geq t + s | X \geq s) = \frac{\mathbb{W}(X \geq t + s, X \geq s)}{\mathbb{W}(X \geq s)} = \frac{\mathbb{W}(X \geq t + s)}{\mathbb{W}(X \geq s)} = \frac{1 - F(t + s)}{1 - F(s)}.$$

Setzt man das oben ein, so erhält man für die Funktion $G(x) := 1 - F(x)$ die Funktionalgleichung $G(t + s) = G(t)G(s)$ der Exponentialfunktion. Weil eine Verteilungsfunktion immer Stetigkeitspunkte hat, muss nach dem Eindeutigkeitsatz für Exponentialfunktionen aus Mathematik 1 also $G = \exp_{-\lambda}$ für ein $\lambda > 0$ gelten. Hieraus folgt $F(x) = 1 - e^{-\lambda x}$ und $\varrho(x) = F'(x) = \lambda e^{-\lambda x}$ für $x \geq 0$, wie behauptet.

- **Cauchyverteilung** zum Parameter a

Diese ist durch die Dichtefunktion

$$\varrho(x) = \frac{a}{\pi(x^2 + a^2)}$$

gegeben. Sie ist von großem theoretischen Interesse und dient vor allem als Beispiel für eine stetige Verteilung ohne Erwartungswert und folglich auch ohne Varianz.

Übungsaufgabe 293. (E) Zeigen Sie, dass es sich bei obiger Funktion ϱ tatsächlich um die Dichte einer W -Verteilung handelt und begründen Sie, warum diese Verteilung keinen Erwartungswert hat.

5.2.4 Das Gesetz der großen Zahlen

Inhalt in Kurzfassung: Das Gesetz der großen Zahlen besagt, grob gesprochen, dass die Mittelwerte von unabhängigen identisch verteilten Zufallsgrößen gegen den Erwartungswert konvergieren. Etwas genauer besehen gibt es mehrere Varianten. Für eine relativ speziell wird hier auch ein Beweis gegeben, in dem als wesentliches Hilfsmittel die Ungleichung von Tschebyschev verwendet wird.

Beträchtliche Teile der Stochastik handeln in der einen oder anderen Form vom Phänomen, dass sich bei langen Messreihen von Daten, die derselben Verteilung unterliegen, häufig sehr stabile arithmetische Mittel einstellen. Man spricht dabei von einem **Gesetz der großen Zahlen** (GdgZ). Etwas präziser gesprochen: Sind X_i identisch verteilte Zufallsgrößen, so geht es darum, ob die Mittelwerte $\frac{1}{n} \sum_{i=1}^n X_i$ konvergieren. Unterschiedliche Aussagen in diese Richtung sind von Interesse und unter geeigneten Voraussetzungen auch beweisbar.

Als Grenzwert der Mittel erweist sich typischer Weise der Erwartungswert $\mathbb{E}(X_i)$, dessen Existenz wir voraussetzen wollen. Es geht also darum, ob sich die Abweichungen $X_i - \mathbb{E}(X_i)$ nach Summation über $i = 1, 2, \dots, n$ und Division durch n weitgehend aufheben. Damit ist eher dann zu rechnen ist, wenn die X_i unabhängig oder wenigstens unkorreliert sind, denn andernfalls könnten wir auch für alle i dasselbe X_i nehmen, wo der Mittelungseffekt sicher nicht auftreten wird. Keinesfalls können wir erwarten, dass die gewünschte Konvergenz immer stattfindet. Folgen die X_i beispielsweise einer Bernoulli-Verteilung zum Parameter $p = \frac{1}{2}$ ($\mathbb{W}(X_i = 0) = \mathbb{W}(X_i = 1) = \frac{1}{2}$), so kommen im Produktraum aller unendlichen 0-1-Folgen sehr wohl auch die konstanten Folgen $0, 0, 0, \dots$ und $1, 1, 1, \dots$ vor, deren Mittel 0 bzw. 1 sind, die also nicht gegen $\mathbb{E}(X_i) = \frac{1}{2}$ konvergieren. Wir können aber hoffen, dass diese Folgen eher die Ausnahme darstellen und dass, wenn man alle denkbaren Folgen mit ihren Wahrscheinlichkeiten in die Waagschale legt, die gewünschte Konvergenz immerhin mit hoher Wahrscheinlichkeit eintritt. Das wird in der folgenden Definition eingefangen.

Definition 5.2.4.1. Seien die X_n , $n = 1, 2, \dots$, Zufallsgrößen, die auf demselben W -Raum definiert sind. (Typisch als Menge Ω der Elementarereignisse ist die Menge aller

möglichen unendlichen Folgen von Versuchsausgängen und als X_n jene Funktion, die das n -te Folgenglied herausgreift.) Außerdem sei vorausgesetzt, dass die Erwartungswerte $\mathbb{E}(X_n) \in \mathbb{R}$ existieren.

Wir sagen, dass die X_i ein **starkes Gesetz der großen Zahlen** erfüllen, wenn

$$\mathbb{W} \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}(X_i)) = 0 \right) = 1$$

gilt. Wir sagen, dass die X_i ein **schwaches Gesetz der großen Zahlen** erfüllen, wenn für jedes $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{W} \left(\left| \frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}(X_i)) \right| \geq \varepsilon \right) = 0$$

gilt.

Klarerweise folgt aus dem starken das schwache GdgZ. Die Umkehrung gilt aber nicht. Zum Beispiel wäre es a priori nicht auszuschließen, dass die Ausnahmemengen A_n , wo $\left| \frac{1}{n} \sum_{i=1}^n (X_i - \mathbb{E}(X_i)) \right| \geq \varepsilon$ gilt, etwa $\mathbb{W}(A_n) = \frac{1}{n}$ erfüllen (was dem schwachen GdgZ entspräche), dass aber eine Menge A von positivem Maß $\mathbb{W}(A) > 0$ existiert derart, dass jedes $\omega \in A$ in unendlich vielen A_n liegt, was dem starken GdgZ widerspräche.

Man beachte weiters, dass im Fall identisch verteilter X_n mit gemeinsamem Erwartungswert μ das starke GdgZ umgeschrieben werden kann zur Konvergenz

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mu$$

fast sicher (d.h. mit Wahrscheinlichkeit 1). Das schwache GdgZ bedeutet in diesem Fall

$$\lim_{n \rightarrow \infty} \mathbb{W} \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| \geq \varepsilon \right) = 0$$

für alle $\varepsilon > 0$.

Die Frage an die W-Theorie lautet nun: Unter welchen Voraussetzungen kann ein Gesetz der großen Zahlen (stark oder wenigstens schwach) garantiert werden? Wir werden in Satz 5.2.4.3 zwei Bedingungen angeben, von denen jede hinreichend ist. Einen Beweis werden wir nur für ein schwaches GdgZ geben unter der relativ starken, restriktiven Voraussetzung, dass die X_n identisch verteilt, paarweise unkorreliert und beschränkt sind. Bei diesem vergleichsweise einfachen Beweis wird aber immerhin transparent, auf welche Weise die Rechenregeln aus 5.2.1 einfließen, insbesondere die Additivität der Varianzen unabhängiger Zufallsgrößen. Die Beweise der stärkeren Aussagen in Satz 5.2.4.3 erfordern etwas mehr technischen Aufwand, den wir hier nicht treiben werden.

Basis für den Beweis ist die **Tschebyschevsche Ungleichung**:

Satz 5.2.4.2. Sei X eine Zufallsvariable über $(\Omega, \mathcal{A}, \mathbb{W})$. Dann gilt für jedes $\varepsilon > 0$

$$\mathbb{W}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}.$$

Beweis. Es sei Y die Indikatorfunktion der Menge $A := \{|X - \mathbb{E}(X)| \geq \varepsilon\}$, also $Y(\omega) = 1$ für $\omega \in A$ und $Y(\omega) = 0$ sonst. Dann gilt offenbar

$$|X - \mathbb{E}(X)| \geq \varepsilon \cdot Y.$$

Daher erhalten wir wegen $Y^2 = Y$

$$\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \geq \mathbb{E}(\varepsilon^2 \cdot Y) = \varepsilon^2 \mathbb{E}(Y) = \varepsilon^2 \mathbb{W}(|X - \mathbb{E}(X)| \geq \varepsilon).$$

Division durch ε^2 liefert die Behauptung. \square

Wir gehen jetzt von einer Folge unkorrelierter und identisch verteilter Zufallsgrößen X_n aus, die der Einfachheit halber überdies als beschränkt angenommen seien. Insbesondere existiert dann auch ein gemeinsamer Erwartungswert $\mu = \mathbb{E}(X_n)$ und eine gemeinsame Varianz $\sigma^2 = \mathbb{V}(X_n)$, jeweils für alle n . Halten wir ein n fest, so hat die Zufallsgröße $X := \frac{1}{n} \sum_{i=1}^n X_i$ wegen der Rechenregeln aus 5.2.1 ebenfalls den Erwartungswert $\mathbb{E}(X) = \mu$ und die Varianz

$$\mathbb{V}(X) = \mathbb{V}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \mathbb{V}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{\sigma^2}{n}.$$

Man beachte, dass in der mittleren Gleichheit die Unkorreliertheit der X_i verwendet wurde.

Wir geben uns nun ein beliebiges $\varepsilon > 0$ vor, setzen in der Tschebyschevschen Ungleichung ein und erhalten

$$0 \leq \mathbb{W}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| > \varepsilon\right) \leq \frac{\mathbb{V}(X)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}.$$

Weil der Ausdruck rechts für $n \rightarrow \infty$ gegen 0 konvergiert, ist damit das schwache GdgZ bewiesen. Mit etwas mehr Aufwand kann man sogar auf das starke GdgZ schließen. Ein einfacher, aber sehr typischer Spezialfall ist ein Zufallsexperiment, das immer wieder (unabhängig) durchgeführt wird. X_n sei die Indikatorfunktion für ein Ereignis A im n -ten Versuch. Dann zählt $\frac{1}{n} \sum_{i=1}^n X_i$ die relative Häufigkeit des Ereignisses A unter den ersten n Versuchen. Nach dem starken GdgZ konvergiert sie für $n \rightarrow \infty$ fast sicher gegen $\mathbb{W}(A)$:

$$\mathbb{W}\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mathbb{W}(A)\right) = 1$$

Im folgenden Satz sind sehr allgemeine Bedingungen gegeben, unter denen ein starkes GdgZ bewiesen werden kann (was wir, wie bereits angekündigt, hier aber nicht tun werden):

Satz 5.2.4.3. *Gegeben sei eine Folge von paarweise unkorrelierten (z.B. unabhängigen) Zufallsgrößen X_n , deren Erwartungswerte $\mathbb{E}(X_n) \in \mathbb{R}$ existieren. Gilt eine der folgenden beiden Bedingungen, dann erfüllen die X_n ein starkes Gesetz der großen Zahlen:*

1. Es existieren sämtliche Varianzen $\mathbb{V}(X_n)$ und sind durch eine reelle Zahl $v \geq \mathbb{V}(X_n)$ beschränkt.
2. Die X_n sind paarweise unabhängig und identisch verteilt.

Wir schließen unsere Behandlung des GdgZ mit einem **Beispiel eines vorteilhaften Spiel, bei dem man auf Dauer verliert:**

Ein Spiel heißt *vorteilhaft*, wenn in jeder Runde der erwartete Gewinn größer als der erwartete Verlust ist. Es sei nun $X_0 = 1$ unser Startkapital. Wir werfen in jeder Runde eine faire Münze. Unser Kapital halbiert sich, wenn eine Null fällt. Bei einer Eins hingegen gewinnen wir zwei Drittel unseres Kapitalstands hinzu. Um das zu formalisieren sei

$$Y_n = \begin{cases} \frac{1}{2} & \text{bei einer Null im } n\text{-ten Wurf} \\ \frac{5}{3} & \text{bei einer Eins im } n\text{-ten Wurf} \end{cases}$$

Dann sind die Zufallsvariablen Y_1, \dots, Y_n unabhängig und identisch verteilt mit Erwartungswert

$$\mathbb{E}Y_n = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{5}{3} = \frac{13}{12} > 1$$

für alle $n \in \mathbb{N}$. Unser Kapital X_n nach der n -ten Runde ist gegeben durch $X_n = Y_1 \cdot Y_2 \cdot \dots \cdot Y_n$. Dieses Spiel ist vorteilhaft, denn bei einem Kapital x nach n Runden ist das erwartete Kapital eine Runde später gleich

$$\mathbb{E}(x \cdot Y_{n+1}) = \frac{13}{12}x.$$

Wir erwarten also einen Gewinn von $x/12$. Aus der Unabhängigkeit der Y_i folgt

$$\mathbb{E}(X_n) = \mathbb{E}(Y_1) \cdot \dots \cdot \mathbb{E}(Y_n) = \left(\frac{13}{12}\right)^n.$$

Daher wächst der Erwartungswert von X_n mit größer werdendem n exponentiell. Dennoch folgt aus dem Gesetz der großen Zahlen, dass man bei diesem Spiel auf lange Sicht verliert. Das sieht man wie folgt.

Es bezeichne μ den Erwartungswert von $\ln Y_i$. Dann ist μ negativ, denn

$$\mu = \mathbb{E}(\ln Y_i) = \frac{1}{2} \ln \frac{1}{2} + \frac{1}{2} \ln \frac{5}{3} = \frac{1}{2} \ln \frac{5}{6} < \frac{1}{2} \ln 1 = 0.$$

Aus dem Gesetz der großen Zahlen folgt, dass für alle $\varepsilon > 0$ gilt:

$$\mathbb{W} \left(\left| \frac{1}{n} \ln X_n - \mu \right| > \varepsilon \right) \leq \mathbb{W} \left(\left| \frac{1}{n} \sum_{i=1}^n \ln Y_i - \mu \right| > \varepsilon \right) \rightarrow 0.$$

Setzen wir nun $\varepsilon = -\mu/2$, so folgt

$$\mathbb{W}(X_n > e^{\frac{\mu}{2}n}) \rightarrow 0.$$

Es wird also immer wahrscheinlicher (die Wahrscheinlichkeit dafür konvergiert gegen 1), dass für unser Kapital gilt

$$X_n \leq e^{\frac{\mu}{2}n} = \left(\frac{5}{6}\right)^{\frac{n}{4}}.$$

Höchstwahrscheinlich haben wir also auf lange Sicht kaum noch Kapital. Dieses Beispiel zeigt wieder, dass der Erwartungswert allein nicht jede Situation adäquat beschreibt.

5.2.5 Der Hauptsatz der mathematischen Statistik

Inhalt in Kurzfassung: Ziemlich schnell folgt aus dem starken Gesetz der großen Zahlen der Hauptsatz der mathematischen Statistik: Empirische Verteilungsfunktionen unabhängiger identisch verteilter Zufallsgrößen konvergieren fast sicher gegen die gemeinsame Verteilungsfunktion. Dieser Sachverhalt ist die Grundlage dafür, dass man unbekannte Verteilungen beliebig gut bestimmen kann, sofern man Zugriff auf beliebig große unabhängige Stichproben hat. Das ist ein wichtiger theoretischer Hintergrund für die Tauglichkeit so vieler statistischer Verfahren.

Die Zufallsgrößen X_1, X_2, \dots seien identisch verteilt mit Verteilungsfunktion F . Für jedes $x \in \mathbb{R}$ definieren wir $F_n(x) := \frac{k}{n}$, wobei k die Anzahl der $i = 1, \dots, n$ mit $X_i \leq x$ ist. F_n ist die Verteilungsfunktion der diskreten Gleichverteilung auf der Menge $\{X_1, \dots, X_n\}$, eventuell unter Berücksichtigung der Vielfachheiten, sofern gewisse Werte mehrmals auftreten. Man nennt F_n auch die **empirische Verteilungsfunktion** der X_n . Mit diesen Bezeichnungen gilt der **Hauptsatz der mathematischen Statistik**:

Satz 5.2.5.1. *Sind X_1, X_2, \dots unabhängige Zufallsgrößen mit derselben Verteilungsfunktion F , so konvergieren ihre empirische Verteilungsfunktionen F_n fast sicher gegen F .*

Beweis. Für festes x definieren wir die Zufallsgröße $Y_n := \mathbf{1}_{(-\infty, x]}(X_n)$. Sie nimmt den Wert 1 an, wenn $X_n \leq x$, sonst 0. Weil die X_n unabhängig sind, sind es auch die mit ihrer Hilfe definierten Y_n . Außerdem sind die Y_n identisch verteilt auf der zweielementigen Menge $\{0, 1\}$ mit Erwartungswert $\mathbb{E}(Y_n) = \mathbb{W}(X_n \leq x) = F(x)$. Nach dem starken Gesetz der großen Zahlen konvergieren die arithmetischen Mittel $\frac{1}{n} \sum_{i=1}^n Y_n$ der Y_n mit Wahrscheinlichkeit 1 gegen $F(x)$. Diese Mittel stimmen aber mit $F_n(x)$ überein. Also gilt fast sicher (d.h. mit Wahrscheinlichkeit 1)

$$\lim_{n \rightarrow \infty} F_n(x) = F(x).$$

Anders formuliert: Für die Ausnahmemenge A_x , wo $F_n(x)$ nicht gegen $F(x)$ konvergiert, gilt $\mathbb{W}(A_x) = 0$. Sei A die Vereinigung A aller A_x mit $x \in D := \mathbb{Q} \cup U$. Dabei bezeichne \mathbb{Q} wie üblich die Menge der rationalen Zahlen und U die Menge der Unstetigkeitsstellen von F . Wegen der Monotonie von F ist U höchstens abzählbar unendlich, wegen der Abzählbarkeit von \mathbb{Q} ist also auch D abzählbar. Die σ -Additivität von \mathbb{W} impliziert daher

$$0 \leq \mathbb{W}(A) = \mathbb{W}\left(\bigcup_{x \in \mathbb{Q}} A_x\right) \leq \sum_{x \in \mathbb{Q}} \mathbb{W}(A_x) = \sum_{x \in \mathbb{Q}} 0 = 0,$$

also $\mathbb{W}(A) = 0$. Mit anderen Worten: Mit Wahrscheinlichkeit 1 stimmt der Grenzwert der F_n auf D mit F überein. Man überlegt sich schnell, dass wegen der Rechtsstetigkeit und der Monotonie der Verteilungsfunktion F und der F_n deren Grenzwert folglich sogar auf ganz \mathbb{R} mit F übereinstimmt. Damit ist der Satz bewiesen. \square

5.2.6 Der zentrale Grenzwertsatz

Inhalt in Kurzfassung: Neben dem Gesetz der großen Zahlen ist der (deutlich schwieriger zu beweisende) Zentrale Grenzwertsatz das vielleicht wichtigste Ergebnis der W-Theorie überhaupt. Er besagt, grob gesprochen, dass die Summe sehr vieler kleiner und unabhängiger Summanden annähernd normalverteilt ist. Die präzise Formulierung ist komplizierter, insbesondere wenn einem an möglichst schwachen, aber trotzdem hinreichenden Bedingungen gelegen ist. Diese präzisieren, was unter *kleinen* Summanden zu verstehen ist. Eine stärkere Bedingung, die hinreicht, ist auch identische Verteilung der Summanden und Existenz von deren Erwartungswert und Varianz. Der zentrale Grenzwertsatz ist der Grund für die überragende Bedeutung der Normalverteilung in der Stochastik.

Im Zentrum der Stochastik steht neben dem Gesetz der großen Zahlen (GdgZ), wonach arithmetische Mittel unter sehr schwachen Voraussetzungen gegen den Erwartungswert konvergieren, der Zentrale Grenzwertsatz (ZGWS) als zweites, noch tiefer liegendes und ebenso ubiquitäres Phänomen. Aus ihm ergibt sich die universelle Bedeutung der Normalverteilung (Glockenkurve). Worum es dabei geht, lässt sich am einfachsten anhand der Binomialverteilung illustrieren, das Phänomen ist aber wesentlich universeller.

Betrachten wir der Einfachheit halber wiederholten Münzwurf („Kopf oder Adler“) mit einer fairen Münze. Zählen wir mit der Zufallsgröße X , wie oft beispielsweise unter 100 Würfeln „Kopf“ aufscheint, so legt das GdgZ nahe, dass für X Werte nahe bei 50 wesentlich wahrscheinlicher sind als nahe bei 0 oder 100. Der ZGWS macht eine noch wesentlich genauere Aussage. Er fängt ein, dass das durch das GdgZ suggerierte Ergebnis $X = 50$ nicht als exakter Wert wahrscheinlich ist, und macht quantitative Angaben, wie sehr z.B. das Ereignis $40 \leq X \leq 60$ wahrscheinlicher ist als die Ereignisse $0 \leq X \leq 20$ oder $80 \leq X \leq 100$. Die Wahrscheinlichkeiten der Binomialverteilung entsprechen nämlich verblüffend genau einer „Gauß’schen Glockenkurve“, deren Maximum (Mittelwert μ) in diesem Fall bei 50 liegt und deren Varianz σ^2 jener der Binomialverteilung zu den Parametern $p = \frac{1}{2}$ und $n = 100$ entspricht, also $\sigma^2 = 100 \cdot p \cdot (1-p) = 25$, d.h. $\sigma = \sqrt{25} = 5$. Im standardisierten Fall ist $\mu = 0$ und $\sigma = 1$. Die Hauptrolle spielen dabei das Integral $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ aus 4.1.7 und seine Transformierten $\int_{\mathbb{R}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$.

Übungsaufgabe 294. (E) Verwenden Sie die bereits bekannte Formel $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ aus 4.1.7, um für beliebige $\mu \in \mathbb{R}$ und $\sigma > 0$ die Formel $\int_{\mathbb{R}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$ zu bestätigen.

Aus dieser Formel folgt, dass, so wie bereits in 5.2.3 behauptet, die Funktionen

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R},$$

für alle $\mu \in \mathbb{R}$ und $\sigma > 0$ Dichtefunktionen von W-Verteilungen sind, den Normalverteilungen.

Definition 5.2.6.1. Die W-Verteilung mit der Dichtefunktion

$$\varphi_{\mu,\sigma}(x) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}$$

mit $\mu \in \mathbb{R}$ und $\sigma > 0$ heißt **Normalverteilung** zu den Parametern μ und σ . Im Fall $\mu = 0$ und $\sigma = 1$ spricht man von **Standardnormalverteilung**. Jede der Dichtefunktionen $\varphi_{\mu,\sigma}$ heißt auch **Gauß'sche Glockenkurve**. Die Verteilungsfunktion der Standardnormalverteilung wird auch mit Φ bezeichnet:

$$\Phi(x) := \int_{-\infty}^x \varphi_{0,1}(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Am Ende des Kapitels findet sich eine Tabelle mit Werten von Φ , der Verteilungsfunktion der Standardnormalverteilung. Sie zeichnet sich durch Erwartungswert 0 und Varianz 1 aus. Ganz allgemeinen erweisen sich die Parameter μ und σ nach elementarer Rechnung als Erwartungswert und Streuung:

Proposition 5.2.6.2. Die Normalverteilung zu den Parametern $\mu \in \mathbb{R}$ und $\sigma > 0$ hat Erwartungswert μ und Varianz σ^2 .

Übungsaufgabe 295. (E) Bestätigen Sie durch Rechnung Proposition 5.2.6.2. Hinweis: Führen Sie zunächst mit Hilfe einer geeigneten Substitution den allgemeinen auf den standardisierten Fall zurück. Argumentieren Sie sodann für den Mittelwert $\mu = 0$ mit Hilfe eines Symmetriearguments. Für die Berechnung der Varianz bietet sich partielle Integration an.

Wir haben nun alles beisammen, um genauer zu formulieren, was wir unter dem ZGWS verstehen wollen. Wie beim GdgZ geht es um die Partialsummen, die aus einer Folge von Zufallsgrößen X_i gebildet werden. Wir wollen also die Verteilung der Summen $S_n := \sum_{i=1}^n X_i$ mit einer Normalverteilung zu geeigneten Parametern vergleichen. Die Summe S_n hat den Erwartungswert $\mu_n := \sum_{i=1}^n \mathbb{E}(X_i)$ und, wenn wir Unabhängigkeit der X_i voraussetzen, Varianz $\sigma_n^2 := \sum_{i=1}^n \mathbb{V}(X_i)$. In jedem Fall wollen wir demnach die Existenz der Erwartungswerte $\mathbb{E}(X_i)$ und Varianzen $\mathbb{V}(X_i)$ der einzelnen X_i voraussetzen. Es geht also darum, ob S_n annähernd normalverteilt ist zu den Parametern μ_n und σ_n . Das ist äquivalent damit, dass die transformierten Zufallsgrößen

$$Y_n := \frac{S_n - \mu_n}{\sigma} = \frac{\sum_{i=1}^n (X_i - \mathbb{E}(X_i))}{\sqrt{\mathbb{V}(X_1 + \dots + X_n)}}$$

annähernd standardnormalverteilt ist. Die einzige Unschärfe liegt noch in der Frage, was hier „annähernd“ bedeutet. Natürlich geht es um den Grenzwert $n \rightarrow \infty$. Und sollen die Verteilungsfunktion der Y_n (punktweise, was in diesem Fall sogar gleichmäßig impliziert) gegen die Verteilungsfunktion Φ der Standardnormalverteilung konvergieren. Wir definieren daher:

Definition 5.2.6.3. Seien X_1, X_2, \dots unabhängige Zufallsgrößen, für die sowohl Erwartungswert $\mathbb{E}(X_i)$ als auch Varianz $\mathbb{V}(X_i) > 0$ als endliche reelle Zahlen existieren. Man sagt, dass für diese Folge von Zufallsgrößen der **Zentrale Grenzwertsatz (ZGWS)** gilt, wenn für die Verteilungsfunktionen F_n der Zufallsgrößen

$$Y_n := \frac{\sum_{i=1}^n (X_i - \mathbb{E}(X_i))}{\sqrt{\mathbb{V}(\sum_{i=1}^n X_i)}}$$

für alle $x \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

gilt.

Sind die X_i unkorreliertheit oder, was im Zusammenhang mit dem ZGWS meist vorausgesetzt wird, unabhängig, so kann man den Ausdruck unter der Wurzel im Nenner der Definition der Y_n zu $\mathbb{V}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathbb{V}(X_i)$ umformen kann. Sind die X_i identisch verteilt mit $\mathbb{E}(X_i) = \mu$ und $\mathbb{V}(X_i) = \sigma$, so lässt sich die Gültigkeit des ZGWS daher umformulieren zu:

Für alle $x \in \mathbb{R}$ gilt

$$\lim_{n \rightarrow \infty} \mathbb{W} \left(\frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n}\sigma} \leq x \right) = \Phi(x).$$

Noch ist die Frage offen, unter welchen Voraussetzungen sich ein ZGWS beweisen lässt. Greifen wir das einleitende Beispiel mit dem Münzwurf nochmals auf, so ist ein relativ elementarer Zugang möglich. Denn für die auftretenden binomialverteilten Zufallsgrößen X zu den Parametern p und n gilt bekanntlich

$$\mathbb{W}(\{X = k\}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n.$$

Auch Erwartungswert $\mathbb{E}(X) = np$ und Varianz $\mathbb{V}(X) = np(1-p)$ der Binomialverteilung sind aus 5.2.2 bereits bekannt. Es geht also darum, für die entsprechend transformierte Zufallsgröße $\frac{X-np}{\sqrt{np(1-p)}}$ annähernde Normalverteilung nachzuweisen. Dabei hat man die obigen Binomialkoeffizienten entsprechend zu transformieren. Die in den Binomialkoeffizienten $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ auftretenden Faktoriellen lassen sich mit Hilfe der Stirlingschen Formel $n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$ aus Mathematik 1 beherrschen. Mit Hilfe einiger Rechnungen (wo z.B. auch Taylorapproximation einfließt, was hier aber nicht ausgeführt wird) lässt sich tatsächlich zeigen, dass die relevanten Größen aus der Binomialverteilung gegen die Werte von $\varphi_{0,1}$ an den entsprechenden Stellen konvergieren. Damit wäre der Beweis des ZGWS in diesem sehr speziellen Fall erbracht.

Die überragende Bedeutung des ZGWS ergibt sich daraus, dass er aber nicht nur für identisch verteilte X_n gilt. Entscheidend sind erstens Linie Unabhängigkeit und zweitens, dass nicht einzelne Summanden X_n gegenüber den anderen dominieren. Das könnte nämlich zur Folge haben, dass die Verteilung dominanter X_n auch in der Summe durchschlagen und Normalverteilung verhindern würde. Zur Präzisierung, worum es geht, können mehrere Bedingungen herangezogen werden:

Definition 5.2.6.4. Für eine Folge von unabhängigen Zufallsgrößen X_1, X_2, \dots , deren Erwartungswerte $\mathbb{E}(X_n)$ und Varianzen $\mathbb{V}(X_n) < \infty$ existieren, sei

$$s_n := \sqrt{\mathbb{V}(X_1) + \dots + \mathbb{V}(X_n)}.$$

Wir formulieren folgende drei Bedingungen:

- **Asymptotische Vernachlässigbarkeit:** Für die Zufallsgrößen $X_{n,i} := \frac{X_i - \mathbb{E}(X_i)}{s_n}$, $i = 1, \dots, n$, $n = 1, 2, \dots$, und alle $\varepsilon > 0$ gilt

$$\lim_{n \rightarrow \infty} \left(\max_{i: 1 \leq i \leq n} \mathbb{W}(|X_{n,i}| \geq \varepsilon) \right) = 0.$$

- **Feller-Bedingung:**

$$\lim_{n \rightarrow \infty} \left(\max_{i: 1 \leq i \leq n} \frac{\sqrt{\mathbb{V}(X_i)}}{s_n} \right) = 0$$

- **Lindeberg-Bedingung:** Für alle $\varepsilon > 0$ gilt

$$\lim_{n \rightarrow \infty} \left(\frac{1}{s_n^2} \sum_{i=1}^n \int_{\{|x - \mathbb{E}(X_i)| \geq \varepsilon s_n\}} (x - \mathbb{E}(X_i))^2 d\mathbb{W}_{X_i}(x) \right) = 0.$$

Man mache sich klar, dass jede dieser drei Bedingungen aus Definition 5.2.6.4 als Präzisierung der weiter oben etwas vage formulierten Forderung verstanden werden kann, nämlich, dass keiner der Summanden X_n gegenüber den anderen dominieren dürfe.

Übungsaufgabe 296. (E) Formulieren Sie Ihre Überlegungen dazu ausführlich.

Studiert man die drei Bedingungen etwas genauer, so erkennt man, dass sie immer stärker werden und z.B. identisch verteilte X_n alle drei erfüllen:

Proposition 5.2.6.5. Unter den Voraussetzungen von Definition 5.2.6.4 gilt:

1. Sind die X_n identisch verteilt, so ist die Lindeberg-Bedingung erfüllt.
2. Aus der Lindeberg-Bedingung folgt die Feller-Bedingung.
3. Aus der Feller-Bedingung folgt asymptotische Vernachlässigbarkeit.

Übungsaufgabe 297. (E) Beweisen Sie Proposition 5.2.6.5.

Hat man sich diese Zusammenhänge vor Augen geführt, erkennt man die Tragweite des nun folgenden **Satzes von Lindeberg-Feller**, einer sehr umfassenden Version des **Zentralen Grenzwertsatzes (ZGWS)**:

Satz 5.2.6.6. Für eine Folge unabhängiger Zufallsgrößen mit endlichen Erwartungswerten $\mathbb{E}(X_n)$ und positiven endlichen Varianzen $\mathbb{V}(X_n)$ sind die folgenden Aussagen äquivalent:

5 Stochastik

1. Es gilt die Lindeberg-Bedingung.
2. Es gilt die Feller-Bedingung und der ZGWS.
3. Es gilt asymptotische Vernachlässigbarkeit und der ZGWS.

Insbesondere gilt im Falle identisch verteilter X_n mit $\mathbb{E}(X_n) = \mu$ und $\mathbb{V}(X_n) = \sigma^2$ der ZGWS, explizit:

$$\lim_{n \rightarrow \infty} \mathbb{W} \left(a \leq \frac{\sum_{i=1}^n (X_i - \mu)}{\sqrt{n\sigma^2}} \leq b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx = \Phi(b) - \Phi(a).$$

Für spätere Zwecke notieren wir noch den Spezialfall Bernoulli-verteilter unabhängiger Zufallsgrößen X_n zum Parameter p . In diesem Fall gilt $\mathbb{E}(X_n) = p$ und $\mathbb{V}(X_n) = p(1-p)$, nach dem ZGWS also für alle $a < b \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} \mathbb{W} \left(a \leq \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq b \right) = \Phi(b) - \Phi(a).$$

Beispiel: Um die Wahrscheinlichkeit zu berechnen, bei $n = 6000$ Würfeln eines fairen Würfels mindestens 1100 Sechsen zu erhalten, ist die Summe

$$\sum_{i=1100}^{6000} \binom{6000}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{6000-k}$$

auszurechnen. Dies ist selbst mit Hilfe eines Computers eine schwierige Aufgabe. Hier kann aber der Zentrale Grenzwertsatz helfen. Das Würfeln einer Sechsen ist ein Bernoulli-Experiment mit $p = 1/6$. Definieren wir die Zufallsvariablen

$$X_i = \begin{cases} 1 & \text{bei einer Sechsen im } i\text{-ten Wurf} \\ 0 & \text{sonst} \end{cases}$$

so ist die gesuchte Wahrscheinlichkeit gerade $\mathbb{W}(\{S_{6000} \geq 1100\})$, wobei wieder

$$S_n = \sum_{i=1}^n X_i.$$

Nach dem Zentralen Grenzwertsatz gilt

$$\mathbb{W}(S_{6000} \geq 1100) = \mathbb{W} \left(\frac{S_{6000} - np}{\sqrt{np(1-p)}} \geq \frac{1100 - 6000 \cdot \frac{1}{6}}{\sqrt{1000 \cdot \frac{1}{6} \cdot \frac{5}{6}}} \right) \approx 1 - \Phi(2\sqrt{3}) \approx 0,00028.$$

Interessiert man sich für die Wahrscheinlichkeit, dass bei 600 Würfeln mindestens 110 Sechsen fallen, so wäre man ohne stochastisches Wissen eventuell versucht zu vermuten, dass dies derselbe Wert wäre wie für 1100 bei 6000 Würfeln ist, denn schließlich geht es in beiden Fällen um denselben relativen Anteil an Sechsen unter sämtlichen Würfeln. Das wäre allerdings ein Irrtum. Die entsprechende Rechnung wie oben liefert nämlich

$$\mathbb{W}(\{S_{600} \geq 110\}) \approx 1 - \Phi \left(\sqrt{\frac{5}{6}} \right) \approx 0,1367.$$

Übungsaufgabe 298. (E) Führen Sie die hier ausgesparten Rechnungen für $\mathbb{W}(S_{600} \geq 110)$ durch. Vergleichen Sie mit dem obigen Ergebnis für $\mathbb{W}(S_{6000} \geq 1100)$ und versuchen Sie, daraus unter Bezugnahme auf den ZGWS möglichst allgemeine Schlüsse zu ziehen.

Übungsaufgabe 299. (T) Aus Erfahrung weiß man, dass 15% aller Wiener gerne auf die Alte Donau rudern gehen. Wie groß ist die Wahrscheinlichkeit unter 5000 Wienern

1. weniger als 500

2. mehr als 300

Ruderer zu finden? Berechnen Sie die Wahrscheinlichkeiten mit dem zentralen Grenzwertsatz.

5.3 Statistik

In der Statistik geht es darum, aus empirischen Daten (Stichproben) Rückschlüsse auf eine Grundgesamtheit, d.h. auf deren Verteilung zu machen. Man interpretiert dabei die empirischen Daten als Werte von Zufallsgrößen mit der gesuchten Verteilung und versucht unter Zuhilfenahme der W-Theorie, aus den Daten möglichst viel Information herauszuholen. Zunächst geht es in 5.3.1 um elementare Unterscheidungen verschiedener Arten von Daten sowie in 5.3.2 um einfache Größen wie empirische Mittelwerte und empirische Varianzen, die aus den Daten gebildet werden können. Sodann wird in 5.3.3 das grundsätzliche Anliegen der Statistik ausführlicher diskutiert. Schließlich werden drei wichtige Methoden der beurteilenden Statistik kurz in ihren Grundideen vorgestellt: Punktschätzer in 5.3.4, Konfidenzintervalle in 5.3.5 und das Testen von Hypothesen in 5.3.6.

5.3.1 Beschreibende Statistik – Merkmale und Skalen

Inhalt in Kurzfassung: Wir beginnen den Abschnitt mit deskriptiven Aspekten von Daten (Unterscheidung zwischen Nominal-, Ordinal- und metrischen Skalen) .

Bei einer Statistik im Sinne von Datenerhebung werden (zufällig) *Untersuchungseinheiten* ω aus einer *Grundgesamtheit* Ω ausgewählt, um gewisse Daten oder Merkmale dieser ω zu erfassen. Die Funktion, die ω eine Merkmalausprägung x zuordnet, bezeichnen wir in Anlehnung an unsere Notation für Zufallsgrößen mit X . Es ist also

$$X : \Omega \rightarrow S, \quad \omega \mapsto X(\omega),$$

wobei S die Menge aller möglichen Merkmalausprägungen bezeichnet. Natürlich gibt es auch viele Untersuchungen, in denen d verschiedene Merkmale gleichzeitig erhoben werden, dann ist $X(\omega) = (X_1(\omega), \dots, X_d(\omega))$ ein d -Tupel. Besteht die Menge S nur aus endlich bzw. abzählbar vielen Merkmalausprägungen, so spricht man von diskreten Merkmalen. Kann jeder Wert eines Intervalls als Merkmalausprägung auftreten, so spricht man von stetigen Merkmalen.

Beispiele:

- (a) Sind wir an der Altersverteilung der Bevölkerung von Wien interessiert, so besteht unsere Grundgesamtheit Ω aus allen Einwohnern Wiens, eine Untersuchungseinheit $\omega \in \Omega$ ist dann ein einzelner Einwohner und $S = \mathbb{N}$. Die Funktion $X : \Omega \rightarrow \mathbb{N}$ ordnet jedem Einwohner ω sein Alter in Lebensjahren, abgerundet auf ganze Jahre zu.
- (b) Um die Augenfarben von Studierenden des Bauingenieurwesens zu untersuchen, wählen wir als Grundgesamtheit Ω alle solche Studierenden, eine Untersuchungseinheit ω ist ein einzelner Student und die Menge S besteht aus einer Teilmenge des Farbspektrums. Die Funktion $X : \Omega \rightarrow S$ weist ω seine Augenfarbe zu.
- (c) Wir möchten Daten der Bevölkerung von Wien erfassen, um zu untersuchen, ob es einen Zusammenhang zwischen dem Rauchverhalten, dem Alter und dem Geschlecht gibt. Dazu sei Ω also wieder die Menge aller Einwohner von Wien und

$$S = \{r, n\} \times \mathbb{N} \times \{w, m\}$$

die Menge der möglichen Merkmalausprägungen, wobei r und n für Raucher bzw. und Nichtraucher stehen und w und m für weiblich bzw. männlich.

Wie die obigen Beispiele zeigen, gibt es verschiedene Typen von Daten. Dementsprechend ist zu unterscheiden, auf welcher Art von Skala die Merkmale liegen:

- *Nominalskala*

Die Ausprägungen nominal skaliert Merkmale können nicht geordnet werden. Der einzig mögliche Vergleich ist der Test auf Gleichheit. Das Merkmal „Farbe“ oder „Wohnort“ ist z.B. nominal skaliert.

- *Ordinalskala*

Die Merkmalausprägungen können geordnet werden. Während eine Interpretation der Rangordnung möglich ist, bleibt eine Abstandsinterpretation fragwürdig oder gar sinnlos. Das Merkmal „Schulnote“ ist etwa ordinal skaliert.

- *Metrische Skala*

Unter den Merkmalausprägungen gibt es eine Rangordnung. Zusätzlich können auch die Abstände zwischen den Merkmalausprägungen gemessen und interpretiert werden. Metrisch skalierte Merkmale können noch weiter unterteilt werden:

- *Intervallskala*

Zwischen den Merkmalausprägungen sind Differenzbildungen sinnvoll; man kann so Abstände bestimmen. Ein Beispiel ist das Merkmal „Temperatur“.

- *Verhältnisskala*

Zu den Eigenschaften der Intervallskala (wobei negative Werte allerdings meist ausgeschlossen sind) kommt noch ein natürlicher Nullpunkt hinzu. Eine Quotientenbildung ist zulässig und Verhältnisse sind sinnvoll interpretierbar. Börsenkurse sind ein Beispiel. Durch Logarithmieren kann eine Verhältnisskala in eine Intervallskala transformiert werden.

Wir werden uns vorwiegend mit diskreten und eindimensionalen metrisch skalierten (d.h. durch Zahlen darstellbaren) Merkmale befassen.

5.3.2 Kennzahlen der beschreibenden Statistik

Inhalt in Kurzfassung: Ein empirischer Datensatz besteht aus einer endlichen Anzahl von Daten x_1, \dots, x_n und induziert, indem man jedem x_i dasselbe Gewicht $\frac{1}{n}$ gibt, einen diskreten W -Raum. Somit lassen sich allgemeine Begriffe über W -Verteilungen auf diesen Fall übertragen, und man erhält Begriffe wie empirische Verteilungsfunktion, empirischer Mittelwert etc.

Jeder endliche Datensatz (Vektor von Daten) $\mathbf{x} = (x_1, \dots, x_n)$ mit Elementen aus einer Grundmenge Ω induziert in natürlicher Weise einen W -Raum $(\Omega, \mathcal{A}, \mathbb{W}_{\mathbf{x}})$. Entsprechend lässt sich jeder Begriff, der sich auf beliebige W -Verteilungen bezieht auch auf empirische Datensätze anwenden. Sind die x_i reelle Zahlen, so lässt sich $\mathbb{W}_{\mathbf{x}}$ auch als reelle W -Verteilung auffassen, weshalb sich auch Begriffe wie Erwartungswert etc. auf empirische Daten übertragen lassen. Die wichtigsten Beispiele solcher begrifflicher Übertragungen fassen wir in der folgenden Definition zusammen.

Definition 5.3.2.1. Für eine empirische Datenerhebung sei Ω die Menge aller möglichen Daten und ein konkreter Datensatz $\mathbf{x} = (x_1, \dots, x_n)$ mit $x_i \in \Omega$ gegeben. Wir führen die folgenden Sprech- und Bezeichnungsweisen ein. Wann immer erforderlich, sei dabei $x_i \in \mathbb{R}$ vorausgesetzt.

1. Sei \mathcal{A} die Potenzmenge (Menge aller Teilmengen) von Ω . Das W -Maß $\mathbb{W}_{\mathbf{x}}$, definiert durch

$$\mathbb{W}_{\mathbf{x}}(A) := \frac{1}{n} |\{i : x_i \in A\}|,$$

heißt auch die **empirische Häufigkeitsverteilung** von \mathbf{x} , $(\Omega, \mathcal{A}, \mathbb{W}_{\mathbf{x}})$ der von Ω und \mathbf{x} induzierte **empirische Wahrscheinlichkeitsraum**. Die Zahl $\mathbb{W}_{\mathbf{x}}(A)$ heißt die **empirische Häufigkeit** von A in \mathbf{x} .

2. Ist $A \subseteq \Omega$ die Menge jener $\omega \in \Omega$, die ein gewisses Merkmal aufweisen, so heißt

$$\mathbb{W}_{\mathbf{x}}(A) = \frac{|\{i : x_i \in A\}|}{n}$$

die **relative Häufigkeit** des Merkmals in \mathbf{x} . (Diese ist zu unterscheiden von der Zahl $|\{i : x_i \in A\}| \in \mathbb{N}$, der **absoluten Häufigkeit** des Merkmals in \mathbf{x} .)

3. Die **empirische Verteilungsfunktion** $F_{\mathbf{x}}$ eines reellen Datensatzes \mathbf{x} ist die Verteilungsfunktion von $\mathbb{W}_{\mathbf{x}}$.
4. Der **empirische Mittelwert** $\mu_{\mathbf{x}}$ von \mathbf{x} (bezeichnet oft auch mit $\bar{\mathbf{x}}$ oder \bar{x}) ist definiert als

$$\mu_{\mathbf{x}} := \mathbb{E}(\mathbb{W}_{\mathbf{x}}) = \frac{1}{n} \sum_{i=1}^n x_i.$$

5. Für $n \geq 2$ ist die **empirische Varianz** $\mathbb{V}_{\mathbf{x}}$ von \mathbf{x} definiert als

$$\mathbb{V}_{\mathbf{x}} := \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_{\mathbf{x}})^2$$

(Der Nenner $n-1$ statt n wird deshalb gewählt, weil nur so ein erwartungstreuer Schätzer für die Varianz entsteht, siehe 5.3.4.)

6. Die Quadratwurzel $\sqrt{\mathbb{V}_{\mathbf{x}}}$ der empirischen Varianz heißt **empirische Standardabweichung** von \mathbf{x} .

7. Die **empirischen α -Quantile** bezüglich reeller Datensätze sind definiert als die entsprechenden α -Quantile von $\mathbb{W}_{\mathbf{x}}$ und werden oft mit \tilde{x}_{α} oder \tilde{x}_{α} bezeichnet. Insbesondere heißen Median ($\alpha = \frac{1}{2}$) und Quartile ($\alpha \in \{\frac{1}{4}, \frac{3}{4}\}$) von $\mathbb{W}_{\mathbf{x}}$ auch **empirischer Median** bzw. **empirische Quartile** von \mathbf{x} . Darüber hinaus heißt die Differenz zwischen maximalem und minimalem x_i auch **Spannweite**, die Differenz zwischen α - und β -Quantilen auch **α - β -Quantilabstand**, der zwischen den beiden Quantilen **Quantilabstand** von \mathbf{x} .

Einige Bemerkungen zu diesen Definitionen:

- Die Mittelbildungen ist nicht immer sinnvoll, auch von reellen Daten. Zum Beispiel sind Mittel von Schulnoten problematisch, von Postleitzahlen sinnlos.
- So wie generell bei Verteilungsfunktionen sind α -Quantile für rationales α , insbesondere also Median und Quartile ($\alpha \in \{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}\}$) nicht immer eindeutig bestimmt, entsprechend auch Quantilabstände. Ist nämlich $\alpha = \frac{m}{n}$ mit $m \in \mathbb{N}$, so ist folgende Situation möglich: Genau für m verschiedene Indizes i gilt $x_i \leq a$, für wenigstens einen davon $x_i = a$; für die restlichen $n-m$ Indizes i gilt $x_i \geq b > a$, für wenigstens einen von ihnen $x_i = b$. Dann sind die α -Quantile genau alle $x \in [a, b]$. Oft wird deshalb auch das ganze Intervall als das (als Menge dann eindeutiges) α -Quantil bezeichnet.
- Der Mittelwert muss für einen Datensatz nicht typisch sein. Ist beispielsweise 2,3 die mittlere Anzahl der Kinder einer Familie in einer Personengruppe, so kann man sicher sein, dass es kein einziges x_i gibt, das genau diesen Wert annimmt.
- Median und Quartile sind, im Gegensatz zum Mittelwert, stabil bezüglich „Ausreißern“, d.h. gegenüber einzelnen Daten, die viel kleiner oder viel größer sind als die Mehrzahl der Daten. Zum Beispiel bei Messungen, wo erratische, große Messfehler möglich sind, hat deshalb der Median oft größere Aussagekraft als der Mittelwert.

Empirischer Mittelwert und Median haben interessante Eigenschaften, die wir noch extra hervorheben wollen. Die Beweise sind nicht schwer und Gegenstand einer Übungsaufgabe.

Proposition 5.3.2.2. *Mit den Bezeichnungen aus Definition 5.3.2.1 gilt:*

1. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(z) := \sum_{i=1}^n (z - x_i)^2$, hat genau eine Extremstelle, und zwar ein Minimum bei $z = \mu_{\mathbf{x}}$, dem empirischen Mittelwert von \mathbf{x} .
2. Die Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, $f(z) := \sum_{i=1}^n |z - x_i|$ hat als Extremstellen ausschließlich Minima und zwar genau an jenen Stellen, die ein empirischer Median von \mathbf{x} sind.

Übungsaufgabe 300. (E) Beweisen Sie Proposition 5.3.2.2

Zum Abschluss noch ein **Beispiel**:

Mäuse mit einem bestimmten Gendefekt überleben nicht die ersten zwei Wochen nach der Geburt:

Tag	1	2	3	4	5	6	7	8	9
Anzahl Mäuse	13	23	136	114	45	10	2	1	1

Die Spannweite dieses Datensatzes ist $9 - 1 = 8$, die Quartile sind $\tilde{x}_{0,25} = 3$ und $\tilde{x}_{0,75} = 4$. Mindestens ein Viertel aller Mäuse überlebt also nicht die ersten drei Tage und mindestens drei Viertel aller Mäuse wird nicht älter als vier Tage. Der Quartilabstand ist damit 1. Er gibt nach Definition einen Bereich an, in dem mindestens die Hälfte aller Daten liegt.

Ebenso ergibt sich wegen $\tilde{x}_{0,05} = 2$ und $\tilde{x}_{0,95} = 5$, dass mehr als 90% aller Mäuse zwischen dem zweiten und fünften Tag sterben. Wir sehen damit, dass Quantilsabstände sowie der Median viel weniger anfällig gegenüber Ausreißern sind als die Spannweite.

Übungsaufgabe 301. (T) Zwanzig Studierende werden befragt, an wie vielen Partys Sie im vergangenen Semester teilgenommen haben:

3, 5.5, 4, 8.5, 6, 7, 4.5, 2, 3, 5, 6, 4.5, 2, 3.5, 7, 3.5, 5, 2, 4, 3.

1. Berechnen Sie den Median und den empirischen Mittelwert.
2. Angenommen ein 21. Studierender, welcher jeden Tag ein Fest besucht hat, liefert ein weiteres Messergebnis, nämlich 100. Wie lautet der neue Mittelwert, wie der neue Median?

Übungsaufgabe 302. (T) Teilen Sie die Ergebnisse der Befragung aus Beispiel 301 in die Klassen $[0, 5)$, $[5, 10)$, $[10, 15)$ und $[15, 20)$. Bestimmen und skizzieren Sie die empirische Häufigkeitsverteilung bezüglich dieser Klasseneinteilung. Berechnen Sie außerdem die empirische Varianz, die mittlere absolute Abweichung sowie den Quartilsabstand.

Übungsaufgabe 303. (E) Machen Sie bei 3 Messwerten x_1, x_2, x_3 plausibel, dass der Median \tilde{x} tatsächlich den Ausdruck

$$\delta(\tilde{x}) = \sum_{i=1}^3 |x_i - \tilde{x}|$$

minimiert.

Hinweis: Wie lautet der Median \tilde{x} bei drei Messwerten konkret? Angenommen $c < \tilde{x}$ bzw. $c > \tilde{x}$. Überlegen Sie sich, dass in jedem der beiden Fälle stets $\sum_{i=1}^3 |x_i - \tilde{x}| < \sum_{i=1}^3 |x_i - c|$.

5.3.3 Allgemeine Vorbemerkungen zur Beurteilenden Statistik

Inhalt in Kurzfassung: Die W -Theorie beschäftigt sich mit der Verteilung von Zufallsgrößen und versucht Prognosen über das Verhalten zufälliger Prozesse zu machen. In der (beurteilenden) Statistik versucht man umgekehrt aus empirischen Daten auf zugrunde liegende Verteilungen zu schließen. Das wird hier zunächst unter grundsätzlichen Gesichtspunkten erörtert und mit einer kurzen Vorschau auf die nachfolgenden Unterabschnitte über Punktschätzungen, Konfidenzintervalle und Hypothesentests verbunden.

Während es in der beschreibenden Statistik darum ging, die Daten einer Erhebung aufzubereiten und Kenngrößen zu berechnen, wollen wir nun Schlüsse aus unseren Daten ziehen. Hierbei versuchen wir, eine Aufgabe zu lösen, die der Sichtweise der Wahrscheinlichkeitstheorie genau entgegengesetzt ist. Wir versuchen, von den Daten einer Stichprobe auf die der Stichprobe zugrunde liegende Verteilung zu schließen oder genauer: Angenommen wir kennen eine Stichprobe, die gemäß einer unbekanntem Wahrscheinlichkeit \mathbb{W} verteilt ist; was können wir aufgrund dieser Stichprobe über \mathbb{W} aussagen?

Wir beginnen mit einer sehr allgemeinen Aufgabe dieser Art, aus einem reellen Datensatz $\mathbf{x}_n = (x_1, \dots, x_n)$ auf die zugrundeliegende W -Verteilung \mathbb{W} zu schließen, ohne irgendein Vorwissen einbringen zu können. Dann bietet sich die empirische Verteilungsfunktion $F_{\mathbf{x}}$ als Approximation der Verteilungsfunktion $F_{\mathbb{W}}$ an. Laut Hauptsatz der mathematischen Statistik 5.2.5.1 herrscht für $n \rightarrow \infty$ mit Wahrscheinlichkeit 1 Konvergenz von $F_{\mathbf{x}_n}$ gegen $F_{\mathbb{W}}$. Wenn n hinreichend groß wird werden wir also fast sicher eine zufriedenstellende Approximation erhalten. Es also darum, möglichst viele unabhängige Daten erheben zu können, was in der Praxis schwierig, aufwendig und teuer sein kann. Wir werden in 5.3.6 den Kolmogorov-Smirnov-Anpassungstests kennen lernen, wo es um die Approximation einer gesuchten Verteilungsfunktion durch eine empirische geht.

Sehr häufig erscheinen zusätzliche Annahmen über die gesuchte Verteilungsfunktion $F_{\mathbb{W}}$ sinnvoll, beispielsweise Stetigkeit. Dann kann man sich verschiedene Methoden überlegen, um $F_{\mathbf{x}}$ durch eine stetige Verteilungsfunktion möglichst nahe bei $F_{\mathbf{x}}$ zu ersetzen. Das wäre aber immer noch ein Beispiel für **nicht-parametrische Statistik**.

Im Gegensatz dazu macht sich die parametrische Statistik zunutze, dass in vielen Anwendungen qualitative Überlegungen nahelegen, dass eine gesuchte W -Verteilung \mathbb{W} aus einer speziellen Klasse stammt oder wenigstens sehr gut durch eine Verteilung aus dieser Klasse approximiert werden kann. So war fast jede der Klassen von Verteilungen aus 5.2.2 und 5.2.3 typisch für eine bestimmte, in der Stochastik häufig auftretende Situation. Beispielsweise ist ziemlich klar, unter welchen Bedingungen man eine Binomialverteilung annehmen darf (wenn es um Summen unabhängiger Bernoulli-verteilter Zufallsgrößen ging), wann eine Normalverteilung (wenn hinreichende Voraussetzungen für die Gültigkeit eines ZGWS erfüllt sind), wann eine geometrische Verteilung (bei einer gedächtnislosen stetigen Verteilung) etc. Die meisten dieser Klassen bestehen aus W -Verteilungen, von denen jede (innerhalb dieser Klasse) durch wenige oder gar nur einen reellen Parameter vollständig charakterisiert ist. In solchen Fällen ist es deshalb meist

effizienter, nicht die Verteilungsfunktion, sondern nur direkt diese(n) charakterisierenden Parameter (etwa Mittelwert und Varianz einer Normalverteilung) zu schätzen oder zu approximieren. Das ist Gegenstand der **parametrischen Statistik**. Die angewandte Statistik ist reich unzähligen Methoden, die jeweils speziellen Situationen angepasst sind.

Die letzten drei verbleibenden Abschnitte werden den folgenden, vorwiegend parametrischen Ansätzen gewidmet sein:

- *Punktschätzungen*
Parameter oder Werte, die vom Parameter der Verteilung abhängen, werden direkt geschätzt.
- *Konfidenzintervalle*
Ein Intervall, in dem der Parameter mit großer Sicherheit liegt, wird angegeben (geschätzt).
- *Hypothesentests*
Vorgegebene Annahmen über den gesuchten Parameter werden getestet und dann, je nach Testergebnis, verworfen oder beibehalten.

Auf den folgenden, diesen drei Themenkreisen gemeinsamen Aspekt sei vorab hingewiesen, weil er ein beständiger Quell für unzulässige Vereinfachungen aufgrund von Missverständnissen ist. Und zwar geht es jeweils um die (annähernde) Bestimmung von W-Verteilungen bzw. deren Parameter. Diese Verteilungen/Parameter werden in der hier präsentierten klassischen, auch **frequentistisch** genannten Statistik als objektiv gegeben (wenn auch unbekannt) betrachtet. (Anders wäre das beispielsweise in der **Bayes-Statistik**, der ein **subjektivistischer** Wahrscheinlichkeitsbegriff zugrunde liegt.) Unter der frequentistischen Sichtweise hat es daher keinen Sinn, zu sagen, der Parameter hat „mit Wahrscheinlichkeit α “ diesen oder jenen Wert. Korrekter müsste es heißen:

Unter der Annahme, dass der Parameter diesen oder jenen Wert hat (bzw. Element einer bestimmten Menge von Parametern ist), liegt bzw. läge eine bestimmte (empirisch, d.h. aus den Daten bestimmbar) statistische Größe mit Wahrscheinlichkeit $\geq \alpha$ in jenem anzugebenden Bereich.

5.3.4 Punktschätzungen

Inhalt in Kurzfassung: Punktschätzungen beziehen sich auf Parameter von Verteilungen wie Erwartungswert, Varianz etc. und verwenden empirische Stichproben. Zum Beispiel kann man einer unbekanntem Verteilung n unabhängige Messwerte entnehmen und deren arithmetisches Mittel als Schätzung für den Erwartungswert nehmen. Wegen des Gesetzes der großen Zahlen hat man bei großem n gute Chancen, nahe am wahren Wert zu liegen. Qualitätsmerkmale von Schätzern sind: Erwartungstreue oder wenigstens asymptotische Erwartungstreue, Konsistenz etc.

Wir gehen aus von der Aufgabe, einen Parameter θ (z.B. Mittelwert oder Varianz einer Verteilung) einer W -Verteilung \mathbb{W}_θ zu schätzen, eventuell auch eine Funktion $\tau(\theta)$ dieses Parameters. Der einfacheren Sprechweise halber, wollen wir uns aber auf den Fall $\tau(\theta) = \theta$ konzentrieren. Die wichtigsten Aspekte werden bereits daran deutlich. Oft ist es wichtig, zwischen dem Parameter und den möglichen Werten zu unterscheiden (ähnlich wie in der Informatik). Um dem gerecht zu werden, bezeichnen wir mit $\Theta \subseteq \mathbb{R}$ die Menge jener reellen Zahlen r , die der Parameter θ als Wert annehmen kann.

Einzigste Grundlage unserer Schätzung für den korrekten Wert von θ soll ein Datensatz $\mathbf{x} = (x_1, \dots, x_n)$ mit $x_i \in \mathbb{R}$ sein. Wir bringen die W -Theorie ins Spiel, indem wir die x_i als Werte $x_i = X_i(\omega)$ unabhängiger Zufallsgrößen $X_i : \Omega \rightarrow \mathbb{R}$ interpretieren, die nach \mathbb{W}_θ mit gesuchtem Parameterwert θ verteilt sind. Jedes x_i entspricht einer einzelnen Messung. Das Elementarereignis ω dürfen wir z.B. als unsere gesamte konkrete Datenerhebung interpretieren (viele andere Datenerhebungen gleichen Umfangs, die derselben W -Verteilung unterliegen, sind ebenso denkbar). Da ω in unserem Kontext konstant ist, werden wir es bestenfalls in Ausnahmefällen anschreiben.

Weil unsere Schätzung soll nur von $\mathbf{x} = (x_1, \dots, x_n)$ abhängt, handelt es sich um eine Funktion, die auf der Menge der möglichen n -Tupel von reellen Messwerten, am besten also gleich auf ganz \mathbb{R}^n definiert ist. Da wir an möglichst guten Schätzungen interessiert sind (über sinnvolle Qualitätsmerkmale werden wir gleich nachdenken) und der Hauptsatz der mathematischen Statistik 5.2.5.1 die Einbeziehung großer Datenmengen nahelegt, wollen wir n variabel halten. Dem werden wir gerecht, indem wir eigentlich Familien $T = (T_n)_{n \in \mathbb{N}^+}$ von messbaren³ Funktionen

$$T_n : \mathbb{R}^n \rightarrow \mathbb{R}, \quad n \in \mathbb{N}^+,$$

betrachten, von denen wir erwarten, dass $T_n(X_1, \dots, X_n)$ für große n mit hoher Wahrscheinlichkeit nahe an θ liegt. In diesem Zusammenhang nennen wir T einen **Schätzer für θ** .

Bisher ist es nur unsere subjektive Absicht, θ zu schätzen. Die T_n an sich haben noch überhaupt nichts mit θ zu tun, weil wir ja zunächst jede Folge von (messbaren) Funktionen $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$ zugelassen haben. Um einen Zusammenhang zwischen den T_n und θ herzustellen, werden wir nun wünschenswerte Eigenschaften formulieren, die rechtfertigen, bei T von einem Schätzer „für θ “ zu sprechen:

Definition 5.3.4.1. Die X_i , $i \in \mathbb{N}^+$, seien unabhängige Zufallsgrößen, die alle nach einer gesuchten Verteilung \mathbb{W}_θ verteilt sind. Als mögliche Werte r für θ seien alle $r \in \Theta$ mit einem festen $\Theta \subseteq \mathbb{R}$ zugelassen. Gegeben seien nun die messbaren Funktionen $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$ und somit Zufallsgrößen $T_n(X_1, \dots, X_n)$ (als Funktionen $\omega \mapsto T_n(X_1(\omega), \dots, X_n(\omega))$), die eine Familie $T = (T_n)_{n \in \mathbb{N}^+}$ bilden.

- Ein einzelnes T_n heißt **erwartungstreu** für θ , wenn für alle möglichen reellen

³Messbarkeit bedeutet hier: Urbilder von reellen Intervallen unter T_n sind n -dimensionale Borelmengen, d.h. in der kleinsten σ -Algebra auf \mathbb{R}^n enthalten, die alle offenen Mengen enthält. Wem diese Begriffe unsympathisch sind, kann diese Voraussetzung hier gefahrlos ignorieren. Denn alle Funktionen, die bei uns auftreten werden, haben diese Eigenschaft.

Werte $r \in \Theta$

$$\mathbb{E}_{\theta=r}(T_n(X_1, \dots, X_n)) = r$$

gilt. (Dabei steht $\mathbb{E}_{\theta=r}$ für den Erwartungswert bezüglich der Verteilung mit dem konkreten Parameterwert $\theta = r$, analog in den folgenden Teilen dieser Definition.) T heißt als Schätzer für θ erwartungstreu, wenn alle T_n erwartungstreu für θ sind.

- T heißt als Schätzer für θ **asymptotisch erwartungstreu**, wenn für alle $r \in \Theta$

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\theta=r}(T_n(X_1, \dots, X_n)) = r$$

gilt.

- T heißt als Schätzer für θ **konsistent**, wenn für jedes $\varepsilon > 0$ und alle $r \in \Theta$

$$\lim_{n \rightarrow \infty} \mathbb{W}_{\theta=r}(|T_n(X_1, \dots, X_n) - r| > \varepsilon) = 0$$

gilt.

- Ist ein zweiter Schätzer $T^* = (T_n^*)_{n \in \mathbb{N}^+}$ für θ gegeben und sind sowohl T als auch T^* erwartungstreu, so heißt T_n **effizienter** als T_n^* , wenn für alle $r \in \Theta$

$$\mathbb{V}_{\theta=r}(T_n(X_1, \dots, X_n)) < \mathbb{V}_{\theta=r}(T_n^*(X_1, \dots, X_n))$$

gilt. T heißt effizienter als T^ , wenn T_n für alle $n \in \mathbb{N}^+$ effizienter als T_n^* ist. (Auch zu diesem Begriff sind Varianten wie „asymptotisch effizienter“ u.ä. von Interesse.)*

- Sind die in Frage kommenden Verteilungen \mathbb{W}_θ diskret, so heißt T_n **Maximum Likelihood-Schätzer**, wenn für jeden Datensatz (X_1, \dots, X_n) die Wahrscheinlichkeit $\mathbb{W}_\theta(T_n(X_1, \dots, X_n) = r)$ unter allen möglichen Werten $r \in \Theta$ für den Wert $r = T_n(x_1, \dots, x_n)$ maximal ist. (Ähnlich kann man Maximum-Likelihood-Schätzer für stetige Verteilungen mit Dichtefunktionen definieren.)

Zur Illustration die wichtigsten Beispiele:

Der empirische Mittelwert $T_n(x_1, \dots, x_n) := \mu_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i$ ist stets ein erwartungstreuere Schätzer für den Erwartungswert $\mathbb{E}(\mathbb{W})$ einer Verteilung \mathbb{W} , sofern dieser Erwartungswert nur existiert. Sind nämlich X_i unabhängige, nach \mathbb{W} verteilte Zufallsgrößen, so gilt nämlich:

$$\mathbb{E}(T_n(X_1, \dots, X_n)) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\mathbb{W}) = \mathbb{E}(\mathbb{W})$$

Der empirische Mittelwert ist sogar ein besonders effizienter Schätzer. Das wird durch die folgende Übungsaufgabe beleuchtet.

Übungsaufgabe 304. (E) Der Erwartungswert $\mathbb{E}(\mathbb{W})$ einer Verteilung \mathbb{W} soll durch lineare Schätzer T_n der Form $T_n(x_1, \dots, x_n) := \sum_{i=1}^n p_i x_i$ geschätzt werden. Zeigen Sie:

1. T_n ist genau dann erwartungstreu, wenn $\sum_{i=1}^n p_i = 1$ gilt.
2. Unter allen erwartungstreuen linearen Schätzern für $\mathbb{E}(\mathbb{W})$ ist der empirische Mittelwert (= das arithmetische Mittel) der effizienteste, d.h. jener mit minimaler Varianz.

Achtung! Aus Übungsaufgabe 304 darf nicht geschlossen werden, dass das empirische Mittelwert immer der beste erwartungstreue Schätzer ist. In der Klasse der stetigen Gleichverteilungen $\mathbb{W}_{[a,b]}$, $a < b \in \mathbb{R}$, auf Intervallen beispielsweise ist es meist effizienter, aus einem Datensatz $\mathbf{x} = (x_1, \dots, x_n)$ statt dem arithmetischen Mittel aller x_i nur jenes des kleinsten, x_{i_0} , und des größten, x_{i_1} , zu nehmen. Man kann sich das leicht so plausibel machen: Diese beiden Messdaten tragen die gesamte Information über das Intervall $[a, b]$, die wir besitzen, nämlich $a \leq x_{i_0} \leq x_{i_1} \leq b$. Darüber hinaus sagen uns die übrigen x_i gar nichts.

Übungsaufgabe 305. (E) Vergleichen Sie diese beiden Schätzer für $n = 3$ hinsichtlich Effizienz.

Übungsaufgabe 306. (E) Zeigen Sie, dass der empirische Mittelwert als Schätzer für den Erwartungswert (sofern dieser existiert) konsistent ist. Hinweis: Gesetz der großen Zahlen verwenden.

Übungsaufgabe 307. (E) Untersuchen Sie, ob der empirische Mittelwert stets (d.h. für alle Verteilungen mit Erwartungswert) ein Maximum Likelihood-Schätzer ist.

Zum Abschluss behandeln wir noch einen Schätzer für die Varianz einer Zufallsgröße, wenn ein Datensatz $\mathbf{x} = (x_1, \dots, x_n)$ von $n \geq 2$ unabhängigen Messdaten x_i vorliegt. (Von einer einzigen Messung wird man sich schwerlich eine sinnvolle Aussage über ein Streumaß erwarten.) Als erwartungstreuer Schätzer erweist sich die schon aus Definition 5.3.2.1 bekannte empirische Varianz:

$$\sigma_{\mathbf{x}}^2 := \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_{\mathbf{x}})^2, \quad \mu_{\mathbf{x}} := \frac{1}{n} \sum_{i=1}^n x_i.$$

Um das zu überprüfen, betrachten wir die unabhängigen Zufallsgrößen, die nach einer unbekanntem Verteilung \mathbb{W} mit Erwartungswert μ und Varianz σ^2 verteilt sind. Bei fixem $n \geq 2$ müssen wir zeigen, dass jede der Zufallsgrößen

$$Y_i := \left(X_i - \frac{1}{n} \sum_{j=1}^n X_j \right)^2, \quad i = 1, \dots, n,$$

den Erwartungswert $\mathbb{E}(Y_i) = \frac{n-1}{n} \sigma^2$ hat. Denn daraus folgt

$$\mathbb{E} \left(\sigma_{X_1, \dots, X_n}^2 \right) = \frac{1}{n-1} \mathbb{E} \left(\sum_{i=1}^n Y_i \right) = \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}(Y_i) = \frac{n}{n-1} \cdot \frac{n-1}{n} \sigma^2 = \sigma^2,$$

also die behauptete Erwartungstreue für $\sigma_{\mathbf{x}}^2$. Zunächst berechnen wir für $i = 1, \dots, n$

$$\begin{aligned} Y_i &= \left(X_i - \frac{1}{n} \sum_{j=1}^n X_j \right)^2 = \left((X_i - \mu) - \frac{1}{n} \sum_{j=1}^n (X_j - \mu) \right)^2 = \\ &= (X_i - \mu)^2 - \frac{2}{n} \sum_{j=1, j \neq i}^n (X_i - \mu)(X_j - \mu) - \frac{2}{n} (X_i - \mu)^2 + \frac{1}{n^2} \left(\sum_{j=1}^n (X_j - \mu) \right)^2. \end{aligned}$$

Von diesem Ausdruck ist der Erwartungswert zu bilden. Wir behandeln jeden Summanden einzeln. Für den ersten gilt

$$\mathbb{E}(X_i - \mu)^2 = \mathbb{V}(X_i) = \sigma^2.$$

Der zweite Summand ist selbst eine Summe, deren Summanden wegen der Unabhängigkeit der X_i den Erwartungswert

$$\mathbb{E}((X_i - \mu)(X_j - \mu)) = \mathbb{E}(X_i - \mu)\mathbb{E}(X_j - \mu) = 0 \cdot 0 = 0$$

haben. Für den dritten Summanden gilt

$$\mathbb{E}\left(\frac{2}{n}(X_i - \mu)^2\right) = \frac{2}{n}\mathbb{V}(X_i) = \frac{2}{n}\sigma^2$$

und für den vierten wegen der Additivität der Varianz unabhängiger Zufallsgrößen

$$\mathbb{E}\left(\frac{1}{n^2} \left(\sum_{j=1}^n (X_j - \mu) \right)^2\right) = \frac{1}{n^2} \mathbb{V}\left(\sum_{j=1}^n X_j\right) = \frac{1}{n^2} \sum_{j=1}^n \mathbb{V}(X_j) = \frac{1}{n^2} n \sigma^2 = \frac{1}{n} \sigma^2.$$

Insgesamt erhalten wir

$$\mathbb{E}(Y_i) = \sigma^2 - 0 - \frac{2}{n}\sigma^2 + \frac{1}{n}\sigma^2 = \left(1 - \frac{2}{n} + \frac{1}{n}\right)\sigma^2 = \frac{n-1}{n}\sigma^2,$$

was zu zeigen war. Wir fassen zusammen:

Satz 5.3.4.2. *Existiert für eine unbekannte W -Verteilung \mathbb{W} der Erwartungswert $\mathbb{E}(\mathbb{W})$, so ist für einen Datensatz $\mathbf{x} = (x_1, \dots, x_n)$ das empirische Mittel*

$$\mu_{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i$$

ein erwartungstreuer Schätzer für $\mu := \mathbb{E}(\mathbb{W})$. Existiert auch die Varianz $\mathbb{V}(\mathbb{W})$ und ist $n \geq 2$, so ist die empirische Varianz

$$\sigma_{\mathbf{x}}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_{\mathbf{x}})^2$$

ein erwartungstreuer Schätzer für $\sigma^2 := \mathbb{V}(\mathbb{W})$.

Noch eine abschließende Bemerkung zum auf den ersten Blick vielleicht überraschenden Faktor $\frac{1}{n-1}$ in der empirischen Varianz. Der einfachere Vorfaktor $\frac{1}{n}$ wäre dann angebracht, wenn der wahre Erwartungswert $\mu = \mathbb{E}(W)$ bekannt wäre und in der Formel für den Schätzer statt $\mu_{\mathbf{x}}$ eingesetzt werden könnte. Ist das aber nicht der Fall, so wird durch Verwendung des empirischen Mittelwertes $\mu_{\mathbf{x}}$ die Varianz systematisch unterschätzt. Man erinnere sich aus 5.3.2 daran, dass $\mu(\mathbf{x})$ die Funktion $t \mapsto \sum_{i=1}^n (x_i - t)^2$ minimiert.

5.3.5 Konfidenzintervalle

Inhalt in Kurzfassung: Bei Konfidenzintervallen sucht man nicht nach einem einzigen Wert als Schätzung für einen unbekanntem Parameter einer Verteilung, sondern nach einem Intervall, in dem der Wert mit hoher Sicherheit liegt. Die Grenzen des Intervalls ergeben sich wieder aus empirischen Daten. Zum Beispiel kann man von einem Schätzer für den Parameter ausgehen und die linke Intervallgrenze kleiner als diesen Schätzer wählen, die rechte Intervallgrenze größer. Um wie viel kleiner bzw. größer die Grenzen sein sollen, ist dann die interessante Frage. Einerseits möchte man das Intervall nicht zu klein wählen, weil dann die Gefahr eines Irrtums steigt, andererseits wäre ein sehr großes Intervall nicht sehr aussagekräftig. Zur Beschreibung dieses Dilemmas steht der Begriff des Irrtumsniveaus α bzw. äquivalent dazu, des Sicherheitsniveaus $1 - \alpha$ im Zentrum.

Mit dem Schlagwort **Konfidenzintervall** ist das Anliegen verbunden, einen Parameter einer Verteilung (Erwartungswert, Varianz, Parameter λ bei einer Exponentialverteilung o.ä.) zu schätzen und ein Intervall anzugeben, in dem der Wert gemäß der Stichprobe „mit hoher Wahrscheinlichkeit“ liegt. Doch Achtung, die logische Struktur der gewünschten Aussage ist komplizierter als diese Formulierung suggeriert! Denn sie täuscht vor, dass der Parameter ein zufälliges Ereignis ist, was aber nicht der Fall ist. Im Sinne der klassischen Statistik (anders wäre es in der sogenannten Bayes-Statistik) gibt es nämlich nur einen wahren Parameter, der allerdings unbekannt ist. Zufällig ist die Stichprobe, aufgrund derer wir unsere Schlüsse ziehen. Um Missverständnisse zu vermeiden, könnte man statt von hoher „Wahrscheinlichkeit“ von hoher „Sicherheit“ sprechen. Unsere Aufgabe lautet also folgendermaßen:

Wir haben die Möglichkeit, Stichproben aus einer unbekanntem Verteilung \mathbb{W} zu erhalten. Immerhin wissen wir, dass \mathbb{W} ein Element einer parametrisierten Klasse von Verteilungen ist, d.h. $\mathbb{W} = \mathbb{W}_\theta$ für ein $\theta \in \Theta$ mit einer Menge $\Theta \subseteq \mathbb{R}$ reeller Parameter.

Wir entnehmen n Messwerte, die wir als unabhängige und nach der unbekanntem Verteilung \mathbb{W} verteilte Zufallsgrößen X_1, X_2, \dots, X_n auffassen. Gesucht sind ein unterer Schätzer t^- und ein oberer Schätzer t^+ für θ . Beide setzen wir als Funktionen in n reellen Variablen x_i an mit

$$t^-(x_1, \dots, x_n) < t^+(x_1, \dots, x_n)$$

für alle möglichen Werte der x_i . Von einem Konfidenzintervall vom Irrtumsniveau α (nahe 0) bzw. vom Sicherheitsniveau $1 - \alpha$ (nahe 1) sprechen wir dann, wenn *für alle*

$\theta \in \Theta$ gilt:

$$\mathbb{W}_\theta \left(t^-(X_1, \dots, X_n) < \theta < t^+(X_1, \dots, X_n) \right) \geq 1 - \alpha$$

Man könnte die Logik dahinter mit einem indirekten Beweis vergleichen: Für jeden Parameter $\theta \in \Theta$ wäre es äußerst unwahrscheinlich, dass aufgrund einer zufälligen Stichprobe der untere Schätzer t^- einen zu großen oder der obere Schätzer t^+ einen zu kleinen Wert für θ liefert.

Ist beispielsweise $\theta = \mu$ der unbekannte Erwartungswert der Verteilung \mathbb{W} . Ginge es nur um eine Schätzung und nicht um ein Konfidenzintervall, so böte sich als Schätzer das empirische (arithmetische) Mittel

$$\mu_{\mathbf{x}} := \frac{1}{n} \sum_{i=1}^n x_i$$

an. Um ein brauchbares Konfidenzintervall zu bekommen, wird man als unteren und oberen Schätzer einen kleineren bzw. einen größeren Wert als $\mu_{\mathbf{x}}$ nehmen; um wie viel kleiner, wird aber von der Varianz σ^2 abhängen. Ist $\sigma > 0$ bekannt, kann man für ein angestrebtes Irrtumsniveau von maximal α beispielsweise $t^- := m - c\sigma$ und $t^+ := m + c\sigma$ mit einem an α angepassten $c > 0$ nehmen.

Konkrete Formeln für Konfidenzintervalle von Parametern zu den unterschiedlichsten Verteilungen finden sich in der Fachliteratur.

Übungsaufgabe 308. (T) Sie haben die verantwortungsvolle Aufgabe, für Recht und Ordnung bei der Befüllung von Biergläsern zu sorgen: Im Spezialitätenrestaurant „Zum Rohbau“ messen Sie die Menge an tatsächlich ausgeschenktem Bier pro Glas. Sie erhalten als Zufallsvektor $(X_i)_{i=1}^{24}$ folgende Ergebnisse in Liter:

0.498, 0.502, 0.490, 0.495, 0.507, 0.492, 0.513, 0.515, 0.492, 0.485, 0.496, 0.504, 0.497, 0.507, 0.505, 0.509, 0.491, 0.503, 0.504, 0.499, 0.494, 0.502, 0.509, 0.489.

1. Wählen Sie eine geeignete Klasseneinteilung in 5 Klassen und skizzieren Sie die empirische Dichte- und Verteilungsfunktion bezüglich Ihrer Klasseneinteilung.
2. Angenommen diese Füllmengen sind normalverteilt. Schätzen Sie μ und σ und skizzieren Sie die Dichtefunktion der μ - σ -Normalverteilung.
3. Berechnen Sie das Konfidenzintervall für den Erwartungswert μ , ausgehend vom Wert des Schätzers $T_n = \frac{1}{n}(X_1 + \dots + X_n)$. Das Konfidenzniveau sei $1 - \alpha$, einmal mit $\alpha = 0.05$ und einmal mit $\alpha = 0.10$.

Hinweis: Rechnen Sie zuerst nach, dass $\mathbb{V}(T_n) = \frac{\sigma^2}{n}$. Nehmen Sie außerdem an, dass $\sqrt{\mathbb{V}(X)}$ gleich dem geschätzten σ -Wert ist.

Übungsaufgabe 309. (T) Die Fahrzeit eines bestimmten Zuges (die aufgrund von Umwelteinflüssen variieren kann) zwischen Ottakring und Simmering wurde an insgesamt

58 aufeinanderfolgenden Tagen gemessen. Es ergibt sich folgende Tabelle der absoluten Häufigkeiten der Fahrzeit in Minuten

Fahrzeit	50	51	52	53	54	55	56	57	58	59	60
absolute Häufigkeit	2	4	4	5	8	9	10	10	3	2	1

Sei X die Zufallsvariable im entsprechenden wahrscheinlichkeitstheoretischen Modell, welche die Fahrzeit misst. Angenommen X ist μ - σ -normalverteilt. Schätzen Sie aus den Messergebnissen μ und σ .

5.3.6 Hypothesentests

Inhalt in Kurzfassung: Wie bei Punktschätzern und Konfidenzintervallen versucht man auch beim Testen von Hypothesen, mit Hilfe empirischer Daten Informationen über unbekannte Parameter von Verteilungen zu gewinnen. Die Logik ist ähnlich wie bei den Konfidenzintervallen, nur lautet die Frage wieder etwas anders, nämlich: Ist die Annahme, dass der gesuchte Parameter in einer gegebenen Menge H_0 liegt (Nullhypothese), plausibel, oder sollen wir die Nullhypothese verwerfen? Man definiert dazu eine Testgröße, die man aus n Messwerten zu ermitteln gedenkt und definiert *a priori* einen Annahmebereich und einen kritischen Bereich, die Komplementärmenge des Annahmebereichs. Fällt die Testgröße in den Annahmebereich, hält man an der Nullhypothese fest, sonst verwirft man sie. Offenbar geht es dabei darum, zwei mögliche Arten von Fehlern zu kontrollieren: eine wahre Nullhypothese irrtümlich zu verwerfen oder eine falsche Nullhypothese irrtümlich beizubehalten.

In 5.3.4 haben wir unbekannte Verteilungsparameter θ (oder davon abhängige Größen $\tau(\theta)$) geschätzt. Nun wollen wir Aussagen über θ formulieren, die wir mit (in einem zu präzisierenden Sinne) großer Sicherheit als richtig oder falsch nachweisen können. Dabei stellt man also eine Hypothese auf, dass der unbekannte Parameter θ der Verteilung in einer bestimmten Menge von Verteilungen liegt, und versucht, diese mathematisch aufgrund der Beobachtungen zu untermauern oder zu widerlegen. Man spricht in diesem Zusammenhang von der **Nullhypothese** H_0 und ihrer **Gegenhypothese** (Alternative) H_1 . Entsprechend ist ein statistischer Test ein Verfahren zur Überprüfung von Annahmen der Gestalt „der Parameter θ der unbekanntenen Verteilung liegt in H_0 “. Ausgehend von Messergebnissen x_1, \dots, x_n soll dabei entschieden werden, ob diese Hypothese plausibel ist, oder ob sie als widerlegt zu betrachten ist und verworfen wird.

Ein statistischer Test kann dabei auf zwei verschiedene Arten zu einem falschen Ergebnis führen:

- *Fehler 1. Art*
Die Nullhypothese H_0 ist wahr, wird aber verworfen.
- *Fehler 2. Art*
Die Nullhypothese H_0 ist falsch, wird aber beibehalten.

Offensichtlich ist es in der Regel unmöglich, beide Fehler gleichzeitig beliebig klein zu halten. Versucht man beispielsweise den Fehler erster Art zu minimieren, indem man das Verwerfen von H_0 erschwert, so vergrößert dies unausweichlich die Chancen auf einen Fehler zweiter Art. In der Statistik hat sich daher eingebürgert, nach einem Test zu suchen, bei dem die Wahrscheinlichkeit für einen Fehler 1. Art kleiner ist als ein vorgegebenes *Signifikanzniveau* α . Die Wahl des Werts α hängt dabei natürlich von der jeweiligen Situation ab: je gravierendere Folgen ein Fehler 1. Art hätte, desto kleiner sollte man α wählen.

Die allgemeine Strategie zur Entwicklung statistischer Tests beruht auf folgenden Punkten:

1. Formulierung einer Nullhypothese H_0 , die gegen ihre Alternative H_1 getestet wird.
2. Wahl einer Testgröße $T_n(X_1, \dots, X_n)$ (beispielsweise eine Schätzfunktion) und Bestimmung der (möglichen) Verteilung(en) von T_n unter H_0 .
3. Festlegung eines Signifikanzniveaus α und Bestimmung des *kritischen Bereichs* B , wo H_0 verworfen werden soll. Der kritische Bereich kann unter Verwendung der Verteilung von T_n (welche in Schritt 2 ermittelt wurde) aus der Bedingung

$$\mathbb{W}_\theta(T_n \in B) \leq \alpha$$

für alle $\theta \in H_0$ berechnet werden. Die Nullhypothese H_0 soll also verworfen werden, wenn $T_n \in B$. Der kritische Bereich hängt von α ab und legt fest, wann genau H_0 abgelehnt wird. Je kleiner α gewählt wird, umso kleiner wird der kritische Bereich und umso schwerer kann H_0 verworfen werden.

4. Einsetzen der Stichprobe x_1, \dots, x_n in T_n . Liegt $T_n(x_1, \dots, x_n)$ in B , so ist nach Wahl des kritischen Bereichs B ein Ereignis eingetreten, gälte die Nullhypothese, sehr unwahrscheinlich wäre. Als Konsequenz wird die Nullhypothese H_0 verworfen und die Alternativhypothese H_1 angenommen. Liegt $T_n(x_1, \dots, x_n)$ nicht in B , so wird H_0 nicht verworfen.

Als sehr typisches Beispiel wollen wir einen ein- und zweiseitigen **Binomialtest** besprechen. Sei dazu X eine Bernoulli-verteilte Zufallsvariable zum unbekanntem Parameter p , die ein Zufallsexperiment mit den beiden möglichen Ergebnissen $\{0, 1\}$ modellieren soll. Zu einer festen (vorläufig angenommenen) Vergleichswahrscheinlichkeit p_0 mit $0 \leq p_0 \leq 1$ testet man beim einseitigen Test die Hypothese

$$H_0 : p \leq p_0 \quad \text{gegen} \quad H_1 : p > p_0,$$

beim zweiseitigen Test

$$H_0 : p = p_0 \quad \text{gegen} \quad H_1 : p \neq p_0.$$

In beiden Fällen bietet sich als Testgröße die Häufigkeit $S_n(X_1, \dots, X_n)$ der Ergebnisse bei n -facher Durchführung des Experiments an.

Unter Verwendung der aus 5.2.2 bekannten Formeln für die Binomialverteilung

$$\mathbb{W}_{p_0}(\{S_n = k\}) = \binom{n}{k} p_0^k (1-p_0)^{n-k},$$

erhält man bei einem Signifikanzniveau α für den einseitigen Test die Bedingung

$$\mathbb{W}_{p_0}(\{S_n \geq K_\alpha\}) = \sum_{i=K_\alpha}^n \binom{n}{i} p_0^i (1-p_0)^{n-i}$$

für ein Verwerfen von H_0 .

Gilt bei gegebener Stichprobe X_1, \dots, X_n die Ungleichung $S_n(X_1, \dots, X_n) < K_\alpha$, so kann H_0 nicht verworfen werden. Beim zweiseitigen Test geht man analog von der Bedingung

$$\mathbb{W}_{p_0}(\{|S_n - np_0| \geq K_\alpha\}) = 1 - \sum_{i: np_0 - K_\alpha < i < np_0 + K_\alpha} \binom{n}{i} p_0^i (1-p_0)^{n-i},$$

aus. Hier liegen Annahme- und Verwerfungsbereich symmetrisch um den Erwartungswert np_0 , was aber keineswegs zwingend ist, sondern hier nur der leichten Fasslichkeit halber so gewählt wurde. Gilt bei gegebener Stichprobe X_1, \dots, X_n die Ungleichung $|S_n(x_1, \dots, x_n) - np_0| < K_\alpha$, so kann H_0 nicht verworfen werden.

Ist die Berechnung der exakten Wahrscheinlichkeiten mit den vielen Binomialkoeffizienten zu aufwendig, so macht man sich bei großem n üblicherweise den ZGWS zunutze und verwendet statt der Binomialverteilung eine Normalverteilung mit demselben Erwartungswert und derselben Varianz („Normalapproximation der Binomialverteilung“).

Beispiel: Wir werfen eine Münze 1000 Mal, um herauszufinden ob sie fair ist. Dazu verwenden wir den zweiseitigen Test mit $p_0 = \frac{1}{2}$, d.h. wir testen den Verteilungsparameter p der Bernoulli-Verteilung auf

$$H_0 : p = \frac{1}{2} \quad \text{gegen} \quad H_1 : p \neq \frac{1}{2}.$$

Das Testniveau betrage 5%. Der kritische Bereich ergibt sich unter Verwendung der Normalapproximation. Unter H_0 hat S_n den Erwartungswert $\mathbb{E}(S_n) = np_0$ und Varianz $\mathbb{V}(S_n) = np_0(1-p_0) = \frac{n}{4}$, für unsere Werte $p_0 = \frac{1}{2}$ und $n = 1000$ also $\mathbb{E}(S_{1000}) = 500$ und $\mathbb{V}(S_{1000}) = 250$. Wir approximieren also mit der Normalverteilung mit $\mu = 500$ und $\sigma^2 = 250$, was zur Bedingung

$$\mathbb{W}_{\frac{1}{2}}(\{|S_{1000} - 500| \geq K_\alpha\}) \approx 2 \cdot \Phi\left(-\frac{K_\alpha}{\sqrt{250}}\right) \leq 0,05,$$

führt. Daraus erhält man numerisch den Wert $31 \leq K_\alpha < 32$. Wir können H_0 also nicht verwerfen, wenn wir zwischen 469 und 531 Einsen sehen. Vergrößert man die Stichprobe, beispielsweise von 1000 auf 10000 Münzwürfe, so wird sich der Annahmebereich der Nullhypothese um den Erwartungswert 5000 herum zwar in absoluten Zahlen vergrößern, in Prozenten jedoch enger werden. Eine genauere Diskussion ist Gegenstand der folgenden Übungsaufgabe.

Übungsaufgabe 310. (E) Rechnen Sie obiges Beispiel mit dem Zahlenwert 10000 statt 1000 durch und versuchen Sie daraus eine möglichst allgemeine Aussage über den Zusammenhang zwischen der Größe der Stichprobe (Anzahl der unabhängigen Messwerte) und der Größe des Annahmebereichs eines Tests bei gegebenem Testniveau abzuleiten.

Ähnliche Überlegungen lassen sich zum Beispiel auf Wahlhochrechnungen anwenden. Sie zeigen, warum Hochrechnungen schon bei sehr geringem Auszählungsgrad beeindruckende Genauigkeit haben – ganz im Gegensatz zu Wahlprognosen, die sich nicht auf echte Stimmabgaben, sondern auf oft sehr fragwürdige Umfragen stützen. Typischerweise werden bei solchen Hochrechnungen aufgrund einiger weniger früh ausgezählter Wahlsprengel die Wählerströme zwischen den wahlwerbenden Parteien (welcher Anteil der Wähler von Partei A beim letzten Mal hat diesmal Partei B gewählt?) geschätzt. Unter Verwendung aller Ergebnisse der letzten Wahl wird davon auf das Gesamtergebnis geschlossen. Es ist ein bemerkenswertes soziologisches Phänomen, dass die geschätzten Wählerströme (als relative Anteile) ziemlich unabhängig von den unterschiedlichen soziologischen Gegebenheiten in unterschiedlichen Wahlkreisen sind.

Um den rein stochastischen Aspekt deutlich zu machen, wollen wir auf den einfachsten Fall eingehen, wo es nur zwei Parteien/Kandidaten A und B gibt, auf die insgesamt N_A bzw. N_B Stimmen entfallen sind und es keine ungültigen Stimmen gibt. Wir nehmen an, dass bisher $n \leq N := N_A + N_B$ Stimmen ausgezählt worden sind, von denen n_A Stimmen auf A und n_B Stimmen auf B entfallen sind, also $n_A + n_B = n$ gilt. Weil wir den einfachsten Fall studieren wollen, sei auch angenommen, dass die ausgezählten Stimmen völlig zufällig ausgewählt worden sind, was in diesem Fall bedeuten soll: nach fiktiver Durchmischung (d.h. unter den ausgezählten Stimmen sind keine Wahlkreise, in denen das Ergebnis vermutlich vom Gesamtergebnis abweicht, systematisch überproportional vertreten) und unabhängig voneinander, was bedeutet: wird eine Stimme zur Auszählung ausgewählt, so hat das keinen Einfluss auf die Wahrscheinlichkeit für eine bestimmte andere Stimme, zur Auszählung ausgewählt zu werden. Alles was uns interessiert ist, ob die Nullhypothese (Kandidat A hat gewonnen) oder die Gegenhypothese H_1 (Kandidat A hat nicht gewonnen, d.h. Kandidat B hat gewonnen oder es herrscht, bei großem N sehr unwahrscheinlich, Stimmengleichstand). Noch *vor der Auszählung* haben wir einen Test zu definieren. Und zwar nehmen wir als Testgröße T_n die Anzahl n_A der Stimmen für den Kandidaten A unter den ersten n ausgezählten Stimmen. Der Annahmebereich für unseren Test soll durch einen Schwellenwert K_α so definiert werden, dass unter der Annahme $N_A \leq N_B$ (bei bekanntem $N = N_A + N_B$) die Wahrscheinlichkeit, dass $\frac{n_A}{n_B} > K_\alpha$ kleiner als α ist. Die Wahrscheinlichkeiten sind der hypergeometrischen Verteilung aus 5.2.2 zu entnehmen. Übertragen wir die Notation von dort auf die aktuelle, so lautet die entscheidende Größe

$$\mathbb{W}_{N_A, N_B, n}(\{X = n_A\}) = \frac{\binom{N_A}{n_A} \cdot \binom{N_B}{n_B}}{\binom{N}{n}}.$$

Die Grenze zwischen den beiden möglichen entgegengesetzten Wahlausgängen (wie wollen die theoretische Möglichkeit eines Stimmengleichstandes vernachlässigen) liegt bei

$N_A = N_B$, also müsste ein möglichst kleines K_α gefunden werden, so dass gerade noch

$$\mathbb{W}(n_A \geq K_\alpha) = \sum_{k: K_\alpha \leq k \leq n} \frac{\binom{N_A}{k} \cdot \binom{N_A}{n_B = n - n_A}}{\binom{N = N_A + N_A}{n}} \leq \alpha$$

gilt. Die hier aus der hypergeometrischen Verteilung resultierenden Summen sind extrem aufwendig zu berechnen. Deshalb argumentieren wir wie folgt: Ist n klein gegenüber N , dann verändert sich die Verteilung der Grundgesamtheit durch Entnahme einer auszählenden Stimme nur unmerklich. Wir können also annähernd von einer Bernoulli-verteiltern Zufallsgröße sprechen zum unbekanntem Parameter $p := \frac{N_A}{N}$. Damit kann ähnlich wie beim vorangegangenen Beispiel mit der fairen/unfairen Münzen mit der Normalverteilung gearbeitet werden. Die Nullhypothese lautet diesmal jedoch $p > p_0 := \frac{1}{2}$, was einen einseitigen Test sinnvoll macht. Die Rechnung mit interessanten Zahlenwerten sei der folgenden Übungsaufgabe vorbehalten.

Übungsaufgabe 311. (E) Sei \mathbb{W}_p eine Bernoulli-Verteilung mit unbekanntem Parameter p . Getestet werden soll die Nullhypothese $p \geq p_0 := \frac{1}{2}$. Formulieren Sie für verschiedene Werte von n (Größe der Stichprobe (x_1, \dots, x_n)) und α (Niveau des Testes) einen einseitigen Binomialtest, indem Sie den Schwellenwert zwischen Annahme und Verwerfen der Nullhypothese angeben. Verwenden Sie dabei Normalapproximation.

Wendet man die Ergebnisse aus Übungsaufgabe 311 auf das vorangegangene Beispiel mit der Wahl an, so stellt sich die Frage, in welche Richtung der Fehler wirkt, der dadurch entsteht, dass eine hypergeometrische Verteilung durch eine Binomialverteilung ersetzt worden ist. Eine kurze Überlegung zeigt deutlich, was in der folgenden Übungsaufgabe genauer zu argumentieren ist:

Übungsaufgabe 312. (E) Die beschriebene Vereinfachung der hypergeometrischen durch eine Binomialverteilung erhöht tendenziell die Sicherheit des resultierenden Tests, d.h. die Wahrscheinlichkeit für Fehler 1. Art wird tendenziell kleiner, die für Fehler 2. Art tendenziell größer.

Zum Abschluss kommen wir noch kurz auf den in 5.3.3 erwähnten (nicht parametrischen) Kolmogorov-Smirnov-Anpassungstest zu sprechen. Dabei lautet die Nullhypothese H_0 , dass die zu testende Verteilung mit einer bestimmten Verteilung übereinstimmt, deren Verteilungsfunktion F_0 bekannt ist. Man definiert die Testgröße

$$T_n(x_1, \dots, x_n) := \sup_{x \in \mathbb{R}} |F_{\mathbf{x}}(x) - F_0(x)|.$$

Die den Test definierenden Schwellenwerte K_α sind dann möglichst klein zu wählen, dass aber immer noch

$$\mathbb{W}(T_n > K_\alpha) \leq \alpha$$

gilt. Konkrete Werte für K_α sind Tabellenwerken bzw. Softwarepaketen zu entnehmen.

Anhang, die Verteilungsfunktion $\Phi(x)$ der Standard-Normalverteilung:

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84850	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99897	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976

Index

, 67

Ableitung, 116

absolut stetig, 302

absolute Häufigkeit, 333

Abstand, 104

affine Funktion, 115

ähnliche Matrizen, 97

algebraische Vielfachheit, 94

Anfangsbedingung, 171

Anfangspunkt, 108

Anfangswert, 167

Anfangswertproblem, 167, 171

anziehender stationärer Punkt, 180

Äquivalenzrelation, 97

asymptotisch erwartungstreuer Schätzer,
339

asymptotische Vernachlässigbarkeit, 329

aufgespannter Unterraum, 13

Ausgleichsproblem, 163

Ausgleichsrechnung, 82

Austauschsatz von Steinitz, 22

Banachscher Fixpunktsatz, 175

Basis, 18

Bayes-Statistik, 337

bedingte Wahrscheinlichkeit, 309

Bernoullische Differentialgleichung, 229

Bernoulliverteilung, 317

Bestimmung von Potentialen, 258

Bestimmung von Stammfunktionen, 225

Bewegungen, 63

bilineare Abbildung, 55

Bilinearform, 51, 55

Binomialtest, 345

Binomialverteilung, 316

Cauchy-Schwarzsche Ungleichung, 245

Cauchyverteilung, 321

charakteristische Gleichung, 93, 186

charakteristisches Polynom, 93, 186

Cramersche Regel, 90

Defekt, 29

Defekt einer Matrix, 41

Definition des Doppelintegrals, 244

Determinante einer linearen Transforma-
tion, 70

Determinante einer Matrix, 69

Determinantenrang, 89

DGL, 165

Dichte, 302

Dichtefunktion, 302

Differential, 116, 221

Differentialform, 278

Differentialgleichung, 165, 167

Differentialoperator, 182

Differentiation von Parameterintegralen,
249

differenzierbar, 116

diskrete Gleichverteilung, 295, 316

diskrete Verteilung, 302

diskreter Wahrscheinlichkeitsraum, 300

Divergenzsatz, 286

Dreiecksungleichung, 245

duale Basis, 55

Durchmesser, 243

effizienter Schätzer, 339

Eigenraum, 91

Eigenvektor, 91

Index

- Eigenwert, 91
einfach zusammenhängend, 133
Einheitsmatrix, 44
Einteilung, 243
Eintrag einer Matrix, 37
Element einer Matrix, 37
elementarer Bereich, 243
Eliminationsverfahren von Gauß, 76, 83
empirische Häufigkeit, 333
empirische Häufigkeitsverteilung, 333
empirische Varianz, 334, 340
empirische Verteilungsfunktion, 325, 333
empirischer Median, 334
empirischer Mittelwert, 333
empirischer Wahrscheinlichkeitsraum, 333
empirisches Quantil, 334
empirisches Quartil, 334
endlichdimensional, 22
Endpunkt, 108
erwartungstreuer Schätzer, 339
Erwartungswert, 305
Erzeugendensystem, 13
Erzeugnis, 13
erzeugter Unterraum, 13
euklidische Länge, 103
euklidische Norm, 103
euklidischer Raum, 57
euklidischer Standardvektorraum, 57
Eulerscher Multiplikator, 227
exakte Differentialgleichung, 223, 224
Exaktheitskriterium, 225
Existenz- und Eindeutigkeitsatz für Differentialgleichungen, 175
explizite Differentialgleichung, 174
Exponentialmatrix, 197
Exponentialverteilung, 320
Extrema unter Nebenbedingungen, 156
Extremstelle, 145
Extremum, 145

fast sicheres Ereignis, 292
Feller-Bedingung, 329
finite Elemente, 232
Fläche, 108

Folgenstetigkeit, 104
Formel von Bayes, 312
Fortsetzungssatz, 30
Fourierkoeffizient, 171
Fourierreihe, 171
frequentistische Statistik, 337
Fundamentallösungen, 189
Funktionaldeterminante, 120, 239
Funktionalgleichung, 166
Funktionalmatrix, 120
Funktionsgebirge, 106

Gauß'sche Glockenkurve, 327
Gaußsches Fehlerintegral, 249
Gebiet, 111, 133
Gegenhypothese, 344
gekoppeltes System, 172
geometrische Verteilung, 317
geometrische Vielfachheit, 94
gerade Permutation, 68
Gesetz von Faraday, 269
Gesetz von Gauß, 269
gewöhnliche Differentialgleichung, 169
Grad, 150
Gradient, 114
Gradientenfeld, 128, 271
Gravitationsfeld, 134
Grenzwert, 104

Halbdiagonalform, 76
Hauptachsentransformation, 100, 154
Hauptdiagonale, 45
Hauptminorenkriterium, 149
Hauptsatz der mathematischen Statistik, 325
Hauptsatz über Umkehrfunktionen, 140
Hauptsatz über implizite Funktionen, 142
Hermiteische Form, 60
Hesse-Matrix, 126, 149
homogene (nichtlineare) Differentialgleichung, 229
homogene Differentialgleichung, 229
homogene lineare Differentialgleichung, 182

- homogenes Gleichungssystem, 83
- hypergeometrische Verteilung, 317
- indefinit, 151
- infinitesimale Größe, 168
- Inhalt, 243
- inhomogene lineare Differentialgleichung, 183
- inhomogenes Gleichungssystem, 83
- instabiler stationärer Punkt, 180
- Integrabilitätsbedingung, 225
- Integrabilitätsbedingungen, 132
- integrierender Faktor, 194, 227
- inverse Matrix, 45
- Isometrien, 63
- isomorph, 32
- Isomorphismus, 32
- Jordan'sche Normalform, 100
- kanonische Basis, 18
- kanonischer Einheitsvektor, 14
- kartesische Koordinaten, 110
- Kern einer linearen Abbildung, 28
- Kern einer Matrix, 41
- Kettenregel, 122
- Koeffizient, 13, 150
- Koeffizientenmatrix eines linearen Gleichungssystems, 73
- kollinear, 17
- Komponente, 10
- Komponente einer Matrix, 37
- Komponente eines Tensors, 55
- Komponenten, 20
- Komponenten einer Funktion, 103
- Konfidenzintervall, 342
- konservatives Vektorfeld, 128
- konsistenter Schätzer, 339
- Kontraktionsprinzip, 175
- Koordinate, 10
- Koordinaten, 20
- Korrelationskoeffizient, 315
- Kovarianz, 314
- Kronecker-Delta, 45
- Kurve, 108
- Kurvenintegral, 255
- Kurvenintegrale auf Potentialfeldern, 257
- Lagrange-Multiplikatoren, 156
- Lagrangesche Multiplikatoren, 157
- Länge, 56
- Laplace-Experiment, 295
- Laplacescher Entwicklungssatz, 81
- Laplaceverteilung, 316
- Lindeberg-Bedingung, 329
- linear abhängig, 15
- linear unabhängig, 15
- lineare Abbildung, 25
- lineare Differentialgleichung, 182
- lineare Hülle, 13
- lineare Transformation, 25
- lineares Gleichungssystem, 82
- Linearform, 34
- Linearkombination, 13
- Lipschitz-Bedingung, 174
- Lipschitz-Stetigkeit, 174
- LK (Linearkombination), 13
- lokal umkehrbar, 139
- lokale Extremstelle, 145
- lokale Maximumsstelle, 145
- lokale Minimumsstelle, 145
- lokale Umkehrfunktion, 139
- lokales Extremum, 145
- lokales Maximum, 145
- lokales Minimum, 145
- Lösung einer Differentialgleichung, 170
- Lösungsmenge, 83
- Lösung eines linearen Gleichungssystems, 73
- Lösungsmenge eines linearen Gleichungssystems, 73
- Masse, 250
- Massenträgheitsmomente, 250
- Matrix einer linearen Abbildung, 39
- Matrixdarstellung einer linearen Abbildung, 39
- Matrizenmultiplikation, 44
- Matrizenpotenzen, 99

Index

- Matrizenprodukt, 43
Maximum, 145
Maximum Likelihood-Schätzer, 339
Maximumsstelle, 145
mehrfach differenzierbar, 125
Methode der kleinsten Quadrate, 163
Metrik, 57, 104
Minimum, 145
Minimumsstelle, 145
Mittelwert, 305
Mittelwertsatz für Doppelintegrale, 245
mittlere absolute Abweichung, 308
Multilinearform, 51
Multiplikation linearer Abbildungen, 44
Multiplikationssatz für Determinanten, 69

negativ definit, 151
negativ semidefinit, 152
nicht-parametrische Statistik, 336
Norm, 56
normal, 58
Normalverteilung, 319, 327
normiert, 58
 n -stellige multilineare Abbildung, 51
Nullabbildung, 41
Nullhypothese, 344
Nullmenge, 236
Nullraum, 13
Nullvektor, 10

OB, 58
Oberfläche, 260
Oberfläche der Kugel, 261
Oberflächenintegral, 263
Oberflächenintegrale über Vektorfelder, 267
Oberflächenintegral, 261, 265
ONB, 58
ONS, 58
orthogonal, 58
Orthogonalbasis, 58
orthogonale Abbildung, 61
orthogonale Matrix, 61
orthogonale Transformation, 61
Orthogonalisierungsverfahren von Schmidt, 59
Orthogonalprojektion, 64
Orthogonalsystem, 58
Orthonormalbasis, 58
Orthonormalsystem, 58
OS, 58

Parameterintegrale, 248
Parameterschätzung, 338
parametrische Statistik, 337
partielle Ableitung der Ordnung n , 115
partielle Differentialgleichung, 170
partikuläre Lösung, 183
partikuläre Lösung, 74
Phasendiagramm, 179
Poissonverteilung, 318
Polarkoordinaten, 110
Polynom in n Variablen, 150
positiv definit, 151
positiv semidefinit, 152
Potentialfeld, 128
Potentialfläche, 271
Potentialflächen, 133
Potentiallinien, 133
Potenzreihenansatz, 220
Produktraum, 310
Projektion, 36

quadratische Form, 55, 152
quadratische Matrix, 37
Quantilabstand, 334
Quartil, 301
Quartilabstand, 334

Randwertproblem, 171
Rang, 29
Rang einer Matrix, 41, 89
Rangformel, 29
reeller Wahrscheinlichkeitsraum, 301
Regressionsgerade, 163
regulär, 29
reguläre Matrix, 41
Rekursionsatz, 166

- relative Häufigkeit, 333
- Rotation, 63
- Rückführung auf Doppel- und Einfachintegrale, 251
- Rückführung auf Einfachintegrale, 246
- Sarrus, 67
- Sattelpunkt, 148, 151, 153
- Satz vom integrierenden Faktor, 227
- Satz von Bolzano-Weierstraß, 105
- Satz von Fubini, 236, 237
- Satz von Gauß, 286
- Satz von Green, Divergenzversion, 282
- Satz von Green, Rotationsversion, 279
- Satz von Picard-Lindelöf, 175
- Satz von Stokes, 284
- Satz von Stokes, allgemein, 278
- Satz von Taylor, 151
- Säulenintegration, 251
- Schätzer, 338
- Schätzer für einen Parameter, 338
- Schichtenintegration, 251
- schiefsymmetrische Multilinearform, 53
- Schrödingergleichung, 167
- Schwarzscher Vertauschbarkeitssatz, 124
- Schwerpunkt, 250
- σ -Algebra, 293
- Signum einer Permutation, 68
- singuläre Matrix, 41
- singuläres Maß, 303
- Singularität, 128
- Skalarfeld, 107
- Spalte einer Matrix, 37
- Spaltenrang, 89
- Spaltenvektor, 37
- Spannweite, 334
- Spiegelung, 63
- stabiler stationärer Punkt, 180
- Stammfunktion, 224
- Standardabweichung, 308
- Standardnormalverteilung, 320, 327
- Standardskalarprodukt, 58
- Standardvektorraum, 10
- stationärer Punkt, 146, 180
- stationäres Feld, 112
- statische Massenmomente, 250
- Stelligkeit, 51
- stetig, 104
- stetige Gleichverteilung, 297, 303, 319
- subjektivistische Wahrscheinlichkeit, 337
- symmetrische Gruppe, 68
- symmetrische Matrix, 100
- symmetrische Multilinearform, 53
- Systemmatrix, 73
- Taylorpolynom, 151
- Tensor, 55
- Transformation auf Polarkoordinaten, 248
- Transformation von Doppelintegralen, 247
- Transformation von Dreifachintegralen, 251
- transponierte Matrix, 38
- Trennung der Variablen, 222
- trivialer Unterraum, 13
- Tschebyschevsche Ungleichung, 322
- überbestimmte Gleichungssysteme, 162
- unabhängige Ereignisse, 296, 309
- unabhängige Zufallsgrößen, 310
- unendlich oft differenzierbar, 125
- unendlichdimensional, 22
- ungerade Permutation, 68
- unitärer Raum, 60
- unkorreliert, 315
- Unterraum, 12
- Varianz, 308
- Vektor, 10
- Vektoraddition, 10
- Vektorfeld, 180, 265
- Vektorprodukt, 72
- Vektorraum, 10
- Verteilung, 301
- Verteilung einer Zufallsgröße, 305
- Volumen der Kugel, 261
- Wahrscheinlichkeitsmaß, 294
- Wahrscheinlichkeitsraum, 293
- Wahrscheinlichkeitsverteilung, 301

Index

Weg, 108
wegzusammenhängend, 108
Winkel, 58
Wärmeleitungsgleichung, 171

Zeile einer Matrix, 37
Zeilenrang, 89
Zeilenvektor, 37
Zentraler Grenzwertsatz, 329
zentrales Kraftfeld, 133
Zerlegung, 243
Zufallsgröße, 304
Zufallsvariable, 304
zusammenhängend, 111, 133