



Michael Bildhauer

Höhere Mathematik für Ingenieure

– Ein Viersemestriger Grundkurs –

– [Band 1](#) –

Universität des Saarlandes

F.R. 6.1 Mathematik

Postfach 151150

D-66041 Saarbrücken

Inhaltsverzeichnis

Band 1

| | | |
|----------|---|-----------|
| 0 | Vorwort | xv |
| I | Grundlagen | 1 |
| 1 | Aussagen, Mengen und Funktionen | 3 |
| 1.1 | Aussagen (wahre und falsche; logische Operationen; Quantoren) . . . | 3 |
| 1.2 | Grundbegriffe der Mengenlehre (Mengenalgebra) | 4 |
| 1.3 | Funktionen (injektiv; surjektiv; bijektiv; Umkehrabbildung; Verkettung) | 9 |
| 2 | \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R} | 15 |
| 2.1 | \mathbb{N} , \mathbb{Z} (Gruppen; vollständige Induktion; Kombinatorik; Primzahlen; abzählbar unendliche Mengen) | 15 |
| 2.2 | \mathbb{Q} (Körper; indirekter Beweis) | 24 |
| 2.3 | \mathbb{R} (Axiomatik; Anordnung; Vollständigkeit; Rechnen mit Ungleichungen; Längenmessung; Überabzählbarkeit) | 27 |
| 3 | Reelle Funktionen, Polynominterpolation | 35 |
| 3.1 | Reelle Funktionen (Grundbegriffe; Polynome; rationale Funktionen) | 35 |
| 3.2 | Polynominterpolation (Darstellung nach Lagrange und Newton; Algorithmus von Neville; Splines) | 37 |
| 4 | Folgen, Reihen, Maschinenzahlen | 47 |
| 4.1 | Reelle Zahlenfolgen (Konvergenz; Divergenz; Konvergenzkriterien; Teilfolgen; Satz von Bolzano-Weierstraß) | 48 |
| 4.2 | Reelle Zahlenreihen (Konvergenz; Divergenz; g -adische Zifferndar- stellung; Konvergenzkriterien; absolute Konvergenz) | 60 |

| | | |
|-----------|--|------------|
| 4.3 | Maschinenzahlen (Festpunktdarstellung; Gleitpunktdarstellung; relativer Rundungsfehler; Maschinengenauigkeit; Rundungsabbildung; Maschinenoperationen) | 70 |
| 5 | Funktionenfolgen, Potenzreihen, Exponentialfunktion | 75 |
| 5.1 | Funktionenfolgen, Funktionenreihen (punktweise und gleichmäßige Konvergenz; gleichmäßige absolute Konvergenz) | 75 |
| 5.2 | Potenzreihen (Konvergenzradius; Konvergenzintervall) | 79 |
| 5.3 | Die Exponentialfunktion (Cauchy-Produkt; Logarithmus; allgemeine Potenzfunktion; Umkehrfunktion) | 82 |
| 6 | Der \mathbb{R}^n | 89 |
| 6.1 | Der Vektorraum \mathbb{R}^n (Vektorräume; Funktionenräume; lineare Abhängigkeit; Dimension; Basis; Unterraum) | 89 |
| 6.2 | Die Geometrie des \mathbb{R}^n (Norm; Skalarprodukt; Cauchy-Schwarzsche Ungleichung; Kosinus; orthogonal; orthonormal; Kronecker-Symbol; Gram-Schmidtsches Orthogonalisierungsverfahren; orthogonale Projektion) | 98 |
| 6.3 | Folgen im \mathbb{R}^n (Übertragung des Konvergenzbegriffes) | 110 |
| 6.4 | Die Topologie des \mathbb{R}^n (beschränkte, offene, abgeschlossene und kompakte Mengen im \mathbb{R}^n) | 112 |
| 7 | Die komplexen Zahlen \mathbb{C} | 123 |
| 7.1 | Einführung der komplexen Zahlen (der Körper der komplexen Zahlen; Fundamentalsatz der Algebra) | 123 |
| 7.2 | Potenzreihen im Komplexen (Identitätssatz; Exponentialfunktion; trigonometrische Funktionen; Hyperbelfunktionen; Eulersche Formeln) | 127 |
| 7.3 | Die reellen trigonometrischen Funktionen (Konsistenz der Definitionen; Eigenschaften) | 133 |
| 7.4 | Die Gaußsche Zahlenebene (Polarkoordinaten; Einheitswurzeln) | 136 |
| II | Lineare Algebra (Teil 1) | 139 |
| 8 | Matrizen und lineare Gleichungssysteme | 141 |
| 8.1 | Matrizenkalkül (der Vektorraum $M(n, m, \mathbb{K})$; lineare Abbildungen; lineare Gleichungssysteme; Matrizenprodukt; Transponierte; Konjugierte; Adjungierte; hermitesches Skalarprodukt im \mathbb{C}^n) | 141 |

| | | |
|------------|---|------------|
| 8.2 | Zur Lösbarkeit linearer Gleichungssysteme (homogene und inhomogene Systeme; Kern einer Matrix; allgemeine Lösung; Superposition von Lösungen; erweiterte Matrix; Spalten- und Zeilenrang; Rang einer Matrix; Dimensionsformel) | 152 |
| 8.3 | Das Gaußsche Eliminationsverfahren (elementare Zeilen- und Spaltenumformungen; Trapezschemata zur Rangbestimmung; Gaußsches Eliminationsverfahren zur Lösung linearer Gleichungssysteme; Pivotelement) | 158 |
| 8.4 | Überbestimmte lineare Gleichungssysteme: Eine Anwendung (Residuum; Methode der kleinsten Quadrate; Ausgleich nach Tschebyscheff; lineare Ausgleichsprobleme; Normalgleichung) | 170 |
| 8.5 | Invertierbare Matrizen (inverse, reguläre, singuläre Matrizen; Dreiecksmatrizen; LR -Zerlegung; Cholesky-Zerlegung) | 173 |
| 8.6 | Determinanten (Volumenfunktion; Transposition; Signum; Spat; Parallelotop; Regel von Sarrus; n -Linearform; alternierend; Multiplikationssatz; Streichungsmatrizen; Laplacescher Entwicklungssatz; Cramersche Regel) | 180 |
| 8.7 | Vektorprodukt, Spatprodukt, Ebenen und Geraden im \mathbb{R}^3 (Orientierung; rechte-Hand-Regel; Normalenvektor; Richtungsvektor; Hessesche Normalform; Abstand zwischen zwei Mengen) | 190 |
| 9 | Lineare Abbildungen | 197 |
| 9.1 | Definition und erste Eigenschaften (Kern und Bild einer linearen Abbildung; Rang; Rangformel; Injektivität und Surjektivität einer linearen Abbildung) | 198 |
| 9.2 | Matrixdarstellung (Koordinaten; darstellende Matrix; Basiswechsel; Koordinatentransformation; Transformation der darstellenden Matrix; Dualraum; lineares Funktional; duale Basis; Bidualraum; Tensor; kovariant; kontravariant; Tensorprodukt) | 201 |
| III | Analysis von Funktionen einer Veränderlichen | 217 |
| 10 | Stetige Funktionen | 219 |
| 10.1 | Stetige Funktionen in mehreren Veränderlichen (Lipschitzstetige Funktionen; Grenzwert von Funktionen; Definition der Stetigkeit; Rechenregeln für stetige Funktionen; Stetigkeit und Kompaktheit) | 219 |

| | | |
|-----------|--|------------|
| 10.2 | Stetige Funktionen einer Variablen (Stetigkeit von Potenzreihen; Umkehrfunktion einer stetigen, streng monotonen Funktion; Zwischenwertsatz) | 236 |
| 11 | Differentialrechnung in einer Veränderlichen | 241 |
| 11.1 | Differenzierbarkeit von Funktionen einer Veränderlichen (Differenzenquotient; Ableitung; Produkt- Quotientenregel; Kettenregel; Ableitung der Umkehrfunktion; gliedweises Differenzieren von Potenzreihen; höhere Ableitungen; die Klasse $C^k(I)$) | 241 |
| 11.2 | Lokale Extrema, Satz von Rolle, Mittelwertsätze (notwendige und hinreichende Bedingungen für lokale Extrema; Satz von Rolle; Ableitung und Monotonie; konvexe und konkave Funktionen; Regeln von l'Hospital) | 253 |
| 11.3 | Numerische Differentiation (Extrapolation zum Limes $h \rightarrow 0$; zentraler Differenzenquotient) | 264 |
| 12 | Eindimensionale Integration | 269 |
| 12.1 | Das bestimmte (Riemannsches) Integral (Zerlegung; Feinheit; Riemannsches Zwischensumme; Untersumme; Obersumme; Verfeinerung; Unterintegral; Oberintegral; Riemann integrierbare Funktionen; Integrabilitätskriterien; Rechenregeln für integrierbare Funktionen; stückweise stetige Funktionen; Integration auf Teilintervallen; Mittelwertsatz der Integralrechnung; orientiertes Riemannsches Integral; Flächeninhalt) | 269 |
| 12.2 | Der Hauptsatz der Differential- und Integralrechnung (Stammfunktion; unbestimmtes Integral) | 278 |
| 12.3 | Integrationstechniken (partielle Integration; Substitution; Partialbruchzerlegung) | 282 |
| 12.4 | Uneigentliche Integrale (lokal integrierbare Funktionen; Konvergenzkriterien) | 290 |
| 12.5 | Parameterabhängige Integrale (Gamma-Funktion; partielle Differenzierbarkeit) | 295 |
| 12.6 | Numerische Integration (Newton-Cotes Formeln; Newton-Cotes Summenformeln; Integrationsverfahren von Romberg) | 297 |
| 13 | Der Satz von Taylor, Fehlerabschätzungen | 305 |
| 13.1 | Taylor-Formel und Taylor-Reihe (Taylor-Polynom; Restglied; Integraldarstellung des Restgliedes; Lagrangesche Restgliedformel; die Klasse C^∞ ; reell analytische Funktionen) | 305 |

| | | |
|------|---|-----|
| 13.2 | Konditionierung einer numerischen Aufgabe (Landausche Symbole; relative Konditionszahlen; Problemfehler) | 314 |
| 13.3 | Stabilität eines Algorithmus (gutartiger Algorithmus) | 319 |
| 13.4 | Restglied bei der Polynominterpolation (Interpolationsfehler) | 325 |

Band 2

| | | |
|-----------|--|------------|
| IV | Gewöhnliche Differentialgleichungen (Teil 1) | 329 |
| 14 | Gewöhnliche lineare Differentialgleichungen | 331 |
| 14.1 | Einführung (explizite, implizite, lineare, homogene, inhomogene, partielle Differentialgleichung; Ordnung einer Differentialgleichung; System von Differentialgleichungen; Anfangswertproblem) | 331 |
| 14.2 | Homogene lineare Differentialgleichungen (Fundamentalsystem; allgemeine Lösung; Wronski-Determinante) | 337 |
| 14.3 | Homogene lineare Differentialgleichungen mit konstanten Koeffizienten (charakteristische Gleichung; Reduktion der Ordnung; charakteristisches Polynom) | 342 |
| 14.4 | Inhomogene lineare Differentialgleichungen (spezielle, allgemeine Lösung der inhomogenen linearen Differentialgleichung; Variation der Konstanten; Resonanz) | 346 |
| V | Lineare Algebra (Teil 2) | 357 |
| 15 | Spektraltheorie quadratischer Matrizen | 359 |
| 15.1 | Eigenwerte und Eigenvektoren (Eigenraum; charakteristisches Polynom einer Matrix; algebraische und geometrische Vielfachheit eines Eigenwertes; ähnliche Matrizen) | 359 |
| 15.2 | Diagonalisierbare Matrizen (Hauptachsentransformation) | 366 |
| 15.3 | Anwendung auf Quadriken (Normalformen von Quadriken) | 373 |
| 15.4 | Die Jordansche Normalform (Hauptvektoren; Kettenbedingung; Jordan-Kästchen) | 379 |
| VI | Gewöhnliche Differentialgleichungen (Teil 2) | 385 |
| 16 | Systeme linearer gewöhnlicher Differentialgleichungen erster Ordnung | 387 |

| | | |
|--|--|------------|
| 16.1 | Systeme linearer gewöhnlicher Differentialgleichungen erster Ordnung (homogenes, inhomogenes System; Fundamentalsystem; allgemeine Lösung; Wronski-Determinante; Reduktion der Ordnung; Variation der Konstanten; Fundamentalmatrix) | 388 |
| 16.2 | Systeme mit konstanten Koeffizienten | 401 |
| VII Analysis von Funktionen mehrerer Veränderlichen | | 405 |
| 17 | Differentialrechnung von Funktionen mehrerer Veränderlicher | 407 |
| 17.1 | Kurven im \mathbb{R}^n (differenzierbare Kurve; Tangentenvektor; reguläre Kurve; singulärer Punkt; Spur; Doppelpunkt; Rektifizierbarkeit; Bogenlänge; Parametertransformation) | 407 |
| 17.2 | Ableitungen (partielle Ableitung; (stetig) partiell differenzierbar; die Klasse $C^1(U)$; Richtungsableitung; Nabla-Operator; Gradient; Tangentialebene; Differential; Jacobi-Matrix; Funktionalmatrix; totale Differenzierbarkeit; Polarkoordinaten; krummlinige Koordinaten; Differentialoperatoren (in krummlinigen Koordinaten); Vektorfeld; Divergenz; Rotation; Laplace-Operator; höhere Ableitungen; Vertauschbarkeit von Ableitungen) | 417 |
| 17.3 | Der Satz von Taylor für Funktionen mehrerer Veränderlicher, lokale Extrema (Multiindex; kritische Punkte; notwendige und hinreichende Bedingungen für die Existenz lokaler Extrema; Definitheit einer Matrix; Hessesche Matrix) | 439 |
| 17.4 | Implizite Funktionen, Umkehrsatz, Extrema mit Nebenbedingungen (Lagrangescher Multiplikator; globale Extrema) | 448 |
| 18 | Kurvenintegrale | 461 |
| 18.1 | Definition und Eigenschaften (Weg; Wegunabhängigkeit des Kurvenintegrals; geschlossener Weg; konservative Vektorfelder; Potential) | 461 |
| 19 | Integralrechnung im \mathbb{R}^n | 471 |
| 19.1 | Das Riemannsches Integral auf quadrierbaren Mengen (Zellen; Jordansches Maß; Zerlegung von Zellen; Untersumme; Obersumme; Unterintegral; Oberintegral; Riemann-Integrierbarkeit; Volumenelement; Mehrfachintegral; charakteristische Funktion; kanonische Fortsetzung; Normalbereich; Cavalierisches Prinzip; Satz von Fubini) | 471 |

| | | |
|-------------|---|------------|
| 19.2 | Der Transformationssatz (Koordinatentransformation; Diffeomorphismus) | 495 |
| 19.3 | Uneigentliche Integrale (reguläre Ausschöpfung; kompakt enthalten; Figur) | 501 |
| VIII | Integralsätze der Vektoranalysis | 505 |
| 20 | Der Gaußsche Integralsatz | 507 |
| 20.1 | Der Gaußsche Integralsatz in der Ebene (positiv orientierter Rand; Greenscher Bereich; äußere Einheitsnormale; Fluss eines Vektorfeldes) | 507 |
| 20.2 | Flächen im \mathbb{R}^3 (Gebiet; zusammenhängend; Parameterdarstellung einer Fläche; Kurven auf Flächen; Tangentialebene; Normaleneinheitsvektor; Oberflächeninhalt; Oberflächenintegral) | 513 |
| 20.3 | Der Gaußsche Integralsatz im \mathbb{R}^3 (äußere Einheitsnormale; Massenfluss; Kontinuitätsgleichung; partielle Integration) | 521 |
| 21 | Der Integralsatz von Stokes | 527 |
| 21.1 | Der Satz (orientierte Fläche; Zirkulation) | 527 |
| IX | Funktionentheorie | 531 |
| 22 | Einführung in die Funktionentheorie | 533 |
| 22.1 | Holomorphe Funktionen (komplexe Differenzierbarkeit; höhere Ableitungen; Rechenregeln) | 534 |
| 22.2 | Die Cauchy- Riemannschen Differentialgleichungen (komplexe Differenzierbarkeit versus reelle Differenzierbarkeit) | 538 |
| 22.3 | Kurvenintegrale (Das komplexe Integral) (Integrationsweg; Wegunabhängigkeit; Stammfunktion) | 541 |
| 23 | Cauchys Integralsatz und Cauchys Integralformel | 551 |
| 23.1 | Der Cauchysche Integralsatz (einfach zusammenhängend; einfache geschlossene Kurven; Fresnelsche Integrale) | 551 |
| 23.2 | Die Cauchysche Integralformel (Existenz beliebiger Ableitungen holomorpher Funktionen; Harmonizität von Real- und Imaginärteil holomorpher Funktionen; Satz von Liouville) | 558 |

| | | |
|-----------|--|------------|
| 24 | Entwicklungen holomorpher Funktionen | 563 |
| 24.1 | Taylor-Reihen (Potenzreihen und holomorphe Funktionen; Differentiation von Potenzreihen) | 563 |
| 24.2 | Laurent-Reihen (holomorphe Funktionen auf Kreisringen; Haupt- und Nebenteil einer Laurent-Reihe; Singularität; gelochte Kreisscheibe; Charakterisierung von isolierten Singularitäten) | 567 |
| 25 | Der Residuensatz | 573 |
| 25.1 | Der Satz (Residuum) | 573 |
| X | Integraltransformationen | 581 |
| 26 | Fourier-Reihen | 583 |
| 26.1 | Einführung (Spektrum; harmonische Analyse; Periode einer Funktion; trigonometrische Reihen; trigonometrische Polynome; gliedweise Integration; Integration und Grenzübergang; Fourier-Koeffizienten) | 583 |
| 26.2 | Der Satz (Sägezahnfunktion; Gibbs-Phänomen; stückweise glatte Funktion; mittlere quadratische Abweichung; Parsevalsche Gleichung) . . | 589 |
| 27 | Fourier-Transformation | 597 |
| 27.1 | Einführung (periodische Fortsetzung; kontinuierliches Spektrum) . . | 597 |
| 27.2 | Fouriers Integralsatz (absolut integrierbar; Fourier-Integral; Diracsche Deltafunktion; weißes Rauschen) | 607 |
| 28 | Bemerkungen zur Laplace-Transformation | 613 |
| 28.1 | Die Transformation (Heaviside-Funktion; konvergenzerzeugender Faktor; exponentielle Ordnung) | 613 |
| XI | Gewöhnliche Differentialgleichungen (Teil 3): Anfangswertprobleme | 619 |
| 29 | Der Banachsche Fixpunktsatz, Nullstellenbestimmung | 621 |
| 29.1 | Der Banachsche Fixpunktsatz (Newton-Verfahren; Fixpunkt; Fixpunktiteration; Verfahrensfunktion; vollständig; Banachraum; Kontraktion) | 621 |
| 29.2 | Nullstellenbestimmung (Konvergenz des Newton-Verfahrens; modifiziertes Newton-Verfahren) | 629 |

| | |
|---|------------|
| 30 Der Satz von Picard-Lindelöf | 635 |
| 30.1 Zwei Beispiele (Nicht-Eindeutigkeit; Nicht-Existenz) | 635 |
| 30.2 Der Satz (Lipschitz-Bedingung; Picard-Iteration) | 637 |
| 31 Einige einfache Differentialgleichungen | 641 |
| 31.1 Geometrische Interpretation (Richtungsfeld; Linienelement; Iso- kline) | 641 |
| 31.2 Spezielle Typen von Differentialgleichungen (Trennung der Variablen) | 642 |
| 32 Numerische Methoden | 647 |
| 32.1 Einführung (stetige Abhängigkeit der Lösung von den Daten; Ein- schrittverfahren; Mehrschrittverfahren) | 647 |
| 32.2 Das Polygonzugverfahren von Euler (konsistentes Verfahren; Ordnung eines Einschrittverfahrens) | 649 |
| 32.3 Das Verfahren von Heun | 652 |
| 32.4 Das modifizierte Euler-Verfahren | 654 |
| 32.5 Runge-Kutta-Verfahren (mehrstufiges Einschrittverfahren) . . . | 654 |
| 32.6 Konvergenz von Einschrittverfahren (globaler Diskretisie- rungsfehler; Verfahrensfehler) | 658 |
| XII Literaturhinweise | 665 |

Kapitel 0

Vorwort

Das vorliegende Manuskript beinhaltet den Stoff einer viersemestrigen ($3 \times (4 + 2) + 1 \times (2 + 1)$) Vorlesung Höhere Mathematik für Ingenieure an der Universität des Saarlandes.

Unabhängig vom Semesterverlauf endet der erste Band mit der Darstellung der Analysis von Funktionen einer Veränderlichen.

Die Vorlesung wurde in dieser Konzeption zum ersten Mal beginnend mit dem Wintersemester 2004/2005 gelesen, in dem der Studiengang Mechatronik eingeführt wurde. Der wesentliche Unterschied zu früheren Semestern ist die Kürzung auf eine zweistündige Vorlesung im vierten Semester, wodurch ein eigenständiger Teil zur Numerik entfallen musste. Numerische Aspekte sind in entsprechend reduzierter Form in den Verlauf der Vorlesung integriert.

Zum Aufbau der Vorlesung:

In Teil I werden Grundlagen diskutiert, die mathematische Sprache wird eingeführt und es werden die Objekte, Räume und Methoden vorgestellt, die das “Handwerkszeug” für die weiteren Betrachtungen sind.

Dazu sind zunächst einige Bemerkungen über Aussagen, Mengen und Funktionen notwendig.

Desweiteren sind zu Beginn die verschiedenen Zahlenbereiche vorzustellen. Begleitend dazu können bereits einige algebraische Strukturen

wie Gruppe, Körper etc. diskutiert werden. Ebenso werden mathematische Beweismethoden beispielhaft erläutert. Die Idee dabei ist, generelle Strukturen und Gesetzmäßigkeiten dort einzuführen, wo sie in natürlichen Beispielen zum ersten Mal in Erscheinung treten.

Der wesentliche Begriff zur Beschreibung von Vorgängen in der Natur ist der einer Funktion. Einfache Funktionen wie Polynome und rationale Funktionen werden im dritten Kapitel behandelt. Damit kann auch die Interpolation von Messdaten durch Polynome schon verstanden werden. Weitere aus der Schule bekannte Funktionen wie insbesondere die Exponentialfunktion werden bewusst noch nicht angesprochen, da die Grundlagen für eine "saubere" Definition etwa von $e^{\sqrt{2}}$ noch nicht zur Verfügung stehen.

Diese werden in den Kapiteln 4 und 5 gelegt. Als sicherlich wichtigster Begriff des ersten Semesters wird der Konvergenzbegriff für Folgen und Reihen ausführlich im vierten Kapitel diskutiert. Die Erweiterung auf Funktionenfolgen und Potenzreihen im fünften Kapitel ermöglicht schließlich eine Definition der Exponentialfunktion, die sich nicht auf nachzulieferndes Wissen beruft.

Waren die bisherigen Betrachtungen eindimensionaler Natur, ist der \mathbb{R}^n der Gegenstand des Kapitels 6. Dieser wird von verschiedenen Gesichtspunkten aus beleuchtet: Die Vektorraumstruktur wird ebenso thematisiert wie die Geometrie, die Konvergenz von Folgen und die Topologie. Insgesamt soll Kapitel 6 zu einem frühen Zeitpunkt das räumliche Denken fördern.

Die komplexen Zahlen werden im siebten Kapitel relativ spät definiert. Dafür gibt es zwei Hauptgründe. Zum Einen wird wie üblich die Gaußsche Zahlenebene mit dem \mathbb{R}^2 identifiziert. Dies ist natürlich besonders einfach, wenn zuvor der \mathbb{R}^n diskutiert wurde, es sind keine Zusatzbemerkungen etwa zur Topologie notwendig. Zum Anderen steht die Exponentialfunktion jetzt als Hilfsmittel zur Verfügung, um die Darstellung in Polarkoordinaten verstehen zu können. Die Verallgemeinerung vieler bisher erzielter Resultate auf den komplexen Fall kann als Übungsaufgabe gestellt werden und ist eine gute Möglichkeit

zur Wiederholung.

Es schließt sich Teil II zur Linearen Algebra (Teil 1) an. Hier wird in Kapitel 8 der Umgang mit Matrizen, Determinanten und linearen Gleichungssystemen eingeübt, lineare Ausgleichsprobleme dienen als Beispiel für überbestimmte Systeme. Wesentlich für den späteren Verlauf der Vorlesung sind die Abschnitte zur Invertierbarkeit von Matrizen und zur Determinante (motiviert als orientiertes Volumen), während sich Paragraph 8.7 auf das Notwendigste beschränkt.

In direkter Folge wird die Brücke vom Matrizenkalkül zu linearen Abbildungen geschlagen. Nach der Definition und einigen vorbereitenden Betrachtungen wird in Kapitel 9.2 besonderer Wert auf die Matrixdarstellung linearer Abbildungen sowie deren Transformationsverhalten gelegt.

Der dritte Teil der Vorlesung beschäftigt sich mit der Analysis von Funktionen einer Veränderlichen. Lediglich beim Studium stetiger Funktionen wird in Kapitel 10.1 der Fall mehrerer Veränderlichen in die Betrachtungen eingeschlossen, da dies ohne großen Mehraufwand möglich ist und da wieder anhand von Beispielen die geometrische Intuition geschult wird.

In den Kapiteln 11 und 12 wird ausführlich die Infinitesimalrechnung in einer Variablen vorgestellt, wobei auch numerische Konzepte angesprochen werden. Hier ist im Wesentlichen der Aufbau vorgegeben. Es gibt lediglich eine Ausnahme:

Der Satz von Taylor wird im eigenständigen Kapitel 13 behandelt, um dort mit seiner Hilfe auch exemplarisch Fehlerabschätzungen zu diskutieren.

Die Infinitesimalrechnung einer Veränderlichen ermöglicht im nächsten Schritt einführende Betrachtungen zum Thema lineare gewöhnliche Differentialgleichungen.

Um auch lineare Systeme behandeln zu können, ist in natürlicher

Weise die Spektraltheorie quadratischer Matrizen im fünften Teil der Vorlesung anzusprechen. Deren Anwendung auf Systeme mit konstanten Koeffizienten schließt sich in Kapitel 16.2 an.

Im zentralen Teil VII, der Analysis von Funktionen mehrerer Veränderlichen, wird die Differenzierbarkeit einer Funktion möglichst anschaulich motiviert. Unterschiede zwischen den Differenzierbarkeitsbegriffen werden zwar angesprochen, in der Regel werden aber alle Funktionen als “hinreichend gut” angenommen. Ansonsten ist der Aufbau wie in den Kapiteln 11, 12, 18 und 19 kanonisch.

Kurvenintegrale und die Integralrechnung im \mathbb{R}^n werden in den Kapiteln 18 und 19 vorgestellt, wobei auch hier technische Feinheiten deutlich reduziert werden. So kann der Integralbegriff als Riemannsches Integral auf quadrierbaren Mengen sehr schnell ohne maßtheoretische Betrachtungen eingeführt werden. Betont werden dagegen der Satz von Fubini und der Transformationssatz als konkrete Hilfsmittel zur Berechnung von Integralen.

Wird in Teil VIII der Vorlesung der Gaußsche Integralsatz zusammen mit dem Flächenbegriff im \mathbb{R}^3 noch relativ ausführlich behandelt, so wird der Integralsatz von Stokes lediglich skizziert.

Die elementaren Betrachtungen zur Funktionentheorie bilden einerseits als Teil IX eine unabhängige Einheit, die für sich genommen als eigenständiges Element der Analysis anzusehen sind. Andererseits sind sie genau dieser Stelle zum Verständnis des Folgenden einzufügen.

Integraltransformationen werden im zehnten Teil diskutiert. Dabei wurde das Kapitel Fourier-Reihen bewusst so weit hinten platziert, um die Fourier-Transformation unmittelbar auf den diskreten Fall zurückführen zu können. Besonderer Wert ist dabei darauf gelegt, unzulässige Vertauschungen von Grenzwerten zu vermeiden, die leider in der Literatur zur Ingenieurmathematik oft zu finden sind. Gleichzeitig bietet Kapitel 27 die Gelegenheit, zahlreiche analytische Konzepte anzuwenden und zu vertiefen. Aus Zeitgründen wurde dahingegen das Kapitel 28 zur Laplace-Transformation relativ knapp gehalten.

Den Abschluss bildet die systematische Diskussion von Anfangswertproblemen. Hier ist der Banachsche Fixpunktsatz als Einstieg gewählt, um in Kapitel 30 als Anwendung den Satz von Picard-Lindelöf in seiner einfachsten Form präsentieren zu können. Einige einfache Differentialgleichungen und die benötigten Rechentechniken sind schließlich noch ebenso vorzustellen wie die Grundlagen der numerischen Aspekte in Kapitel 32.

Die Verteilung des Stoffes auf vier Semester sieht wie folgt aus: Im ersten Semester können die Kapitel 1–8 behandelt werden, das zweite Semester schließt mit Kapitel 15. Der Einschub zur Funktionentheorie ist als letzter Teil des dritten Semesters gedacht, die restlichen Kapitel werden im (verkürzten) vierten Semester behandelt.

Zur äußeren Form sei festgehalten: Die farblichen Hinterlegungen im Text sollen die Orientierung erleichtern: **Neu auftauchende** Begriffe im laufenden Text sind **blau** gekennzeichnet, besonders **betont** sind die **roten** Bemerkungen. Definitionen, Sätze etc. sind komplett hervorgehoben. Die Darstellung ist streng strukturiert, damit wichtige Fakten nicht im laufenden Text untergehen. Dies erleichtert gerade Studierenden in den ersten Semestern den Überblick, zumal auch das Nachschlagen wesentlich erleichtert wird. Zum gleichen Zweck sind die Hinweise in den Kapitelüberschriften als Alternative zu einem Index gedacht.

Die historischen Daten sind überwiegend den Lehrbüchern von Ansgar und Oberle entnommen, die als Begleitmaterial ebenso empfohlen werden wie die von Bärwolff, Burg/Haf/Wille und Hoffmann/Marx/Vogt.

Das vorliegende Manuskript befindet sich zur Zeit in der Überarbeitung. Hinweise auf Fehler oder Unstimmigkeiten sowie weitere Anregungen jeglicher Art sind ausdrücklich erwünscht.

Teil I

Grundlagen

Kapitel 1

Aussagen, Mengen und Funktionen

Zu Beginn müssen die **mathematische Sprache** und die mathematischen Grundregeln festgelegt werden.

1.1 Aussagen (wahre und falsche; logische Operationen; Quantoren)

Als **Aussagen** werden **Sätze** bezeichnet, denen man einen der beiden **Wahrheitswerte** **wahr** oder **falsch** zuordnen kann. Bezeichnet weiter $w(A)$ den Wahrheitswert einer Aussage A , so gilt **per definitionem**

$w(A) = 0$ genau dann, wenn gilt: Aussage A ist falsch ;

$w(A) = 1$ genau dann, wenn gilt: Aussage A ist richtig .

Aussagen können durch **logische Operationen** verknüpft werden. Sind A und B zwei Aussagen, so schreibt man:

$\neg A$ “nicht A ” (**Negation**);

$A \wedge B$ “ A und B ” (**Konjunktion**);

$A \vee B$ “ A oder B ” (**Disjunktion**);

$A \Rightarrow B$ “aus A folgt B ” (**Implikation**);

$A \Leftrightarrow B$ “ A äquivalent zu B ” (**Äquivalenz**).

Es entstehen neue Aussagen, deren Wahrheitswert durch die folgende Tabelle **definiert** ist:

| $w(A)$ | $w(B)$ | $w(\neg A)$ | $w(A \wedge B)$ | $w(A \vee B)$ | $w(A \Rightarrow B)$ | $w(A \Leftrightarrow B)$ |
|--------|--------|-------------|-----------------|---------------|----------------------|--------------------------|
| 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| 0 | 0 | 1 | 0 | 0 | 1 | 1 |

Bemerkungen.

- i) $A \vee B$ ist also als “nicht-ausschließendes oder” zu verstehen (und nicht als “entweder ... oder”).
- ii) Ist A falsch, so ist die Aussage “aus A folgt B ” immer richtig.
- iii) Mit obiger Tabelle verifiziert man (als Übung)

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A) .$$

- iv) Ein **mathematischer Satz** ist eine Implikation $A \Rightarrow B$. Hierbei ist A die **Voraussetzung** und B die **Behauptung**. Ein **direkter Beweis** wird in Form eines **Kettenschlusses** geführt:

$$A =: A_0 \Rightarrow A_1 \Rightarrow A_2 \Rightarrow \cdots \Rightarrow A_n = B .$$

Der Doppelpunkt hinter dem ersten Gleichheitszeichen besagt dabei, dass A_0 **per definitionem** gleich A ist.

Schließlich sind **Aussageformen** (oft ebenfalls als Aussagen bezeichnet) Aussagen, die von **Variablen** abhängen, eine Aussageform in den Variablen x und y ist beispielsweise

$$A(x, y) : \Leftrightarrow x^2 + y^2 = 1 .$$

Dann ist $A(0, 1)$ wahr, $A(1, 1)$ ist dagegen falsch.

Man benutzt die **Quantoren**:

- $\forall x : A(x)$ “für alle x gilt $A(x)$ ”;
- $\exists x : A(x)$ “es gibt (wenigstens) ein x , sodass $A(x)$ wahr ist”;
- $\exists! x : A(x)$ “es gibt genau ein x , sodass $A(x)$ wahr ist”.

1.2 Grundbegriffe der Mengenlehre (Mengenalgebra)

Die “naive”² Vorstellung des **Mengenbegriffes** geht auf Georg Cantor³ zurück:

¹Das Zeichen $:\Leftrightarrow$ bedeutet: Die Aussage auf der linken Seite ist **per definitionem äquivalent** zur Aussage auf der rechten Seite.

²Auf eine Diskussion der **Axiomatik** sei hier verzichtet.

³G. Cantor, 1845-1918; Berlin, Halle.

Definition 1.2.1

Eine *Menge* A ist eine Zusammenfassung von wohl bestimmten und wohl unterschiedenen Objekten (den *Elementen* der Menge) zu einem Ganzen.

Schreibweisen.

- i) $x \in A$: x ist Element der Menge A ;
- ii) $x \notin B$: x ist kein Element der Menge B ;
- iii) $A = \{x, y, z, \dots\} = \{z, y, x, \dots\}$: A ist die Menge, die aus den Elementen x, y, z, \dots besteht (explizites Auflisten) – dabei **spielt die Reihenfolge keine Rolle**.
- iv) $A = \{x : x \text{ hat die Eigenschaft } E\}$
 $(A = \{y : y \text{ hat die Eigenschaft } E\})$:
 A ist die Menge aller Elemente, die die Eigenschaft E haben.

Beispiele.

- i) $A = \{1, 2, 3\} = \{n : n \text{ ist eine natürliche Zahl}^4 \text{ kleiner als } 3.\}$.
- ii) $A = \mathbb{N} = \{1, 2, 3, \dots\} = \{n : n \text{ ist eine natürliche Zahl.}\}$.
- iii) Die **leere Menge** \emptyset ist die Menge, die **kein** Element enthält⁵.

Beziehungen zwischen Mengen?
Definition 1.2.2

Es seien A und B zwei Mengen.

- i) $A = B :\Leftrightarrow A$ und B haben dieselben Elemente.
- ii) $A \subset B :\Leftrightarrow$ Jedes Element von A gehört zu B
(A ist Teilmenge von B).

⁴Was ist das eigentlich?

⁵Die **Existenz der leeren Menge** gehört zur Axiomatik ebenso wie die Diskussion der **Menge aller Mengen**.

Beispiel. Gleichheit von A und B überprüft man in der Regel mittels der beiden Teilmengenrelationen $A \subset B$ und $B \subset A$. Ist etwa

$A := \{\text{Dreiecke } D : D \text{ ist rechtwinklig und beide Katheten haben die Länge } 1\}$,

$B := \{\text{Dreiecke } D : D \text{ hat zwei } 45 \text{ Grad Winkel und die längste Seite hat die Länge } \sqrt{2}\}$,

so schließt man: Ist $D \in A$, so folgt (da beide Katheten gleiche Länge haben), dass D zwei 45 Grad Winkel haben muss. Aus dem Satz des Pythagoras⁶ folgt weiter, dass die längste Seite die Länge $\sqrt{2}$ hat, also gilt $A \subset B$.

Ist umgekehrt $D \in B$, so folgt (Winkelsumme im Dreieck ist 180 Grad), dass D rechtwinklig ist. Die längste Seite ist die Hypotenuse (nach Annahme von der Länge $\sqrt{2}$), die Katheten haben gleiche Länge (zwei 45 Grad Winkel), nach dem Satz des Pythagoras sind sie von der Länge 1. Also $B \subset A$ und somit auch $A = B$.

Wie kann man mit Mengen operieren? (**Mengenalgebra**)

Definition 1.2.3

*Es seien A und B zwei Mengen. Dann sind i) die **Ver- einigung**, ii) der **Durchschnitt**, iii) die **Differenz**, iv) das **kartesische Produkt** und v) die **Potenzmenge** definiert als*

$$i) A \cup B := \{x : x \in A \vee x \in B\};$$

$$ii) A \cap B := \{x : x \in A \wedge x \in B\};$$

$$iii) A - B := \{x : x \in A \wedge x \notin B\};$$

$$iv) A \times B := \{(a, b) : a \in A \wedge b \in B\};$$

$$v) \mathcal{P}(A) := \{X : X \subset A\}.$$

⁶Pythagoras, um 570 v. Chr.; Samos.

Bemerkungen.

- i) Veranschaulicht werden diese Beziehungen und Operationen in so genannten **Venn-Diagrammen**.

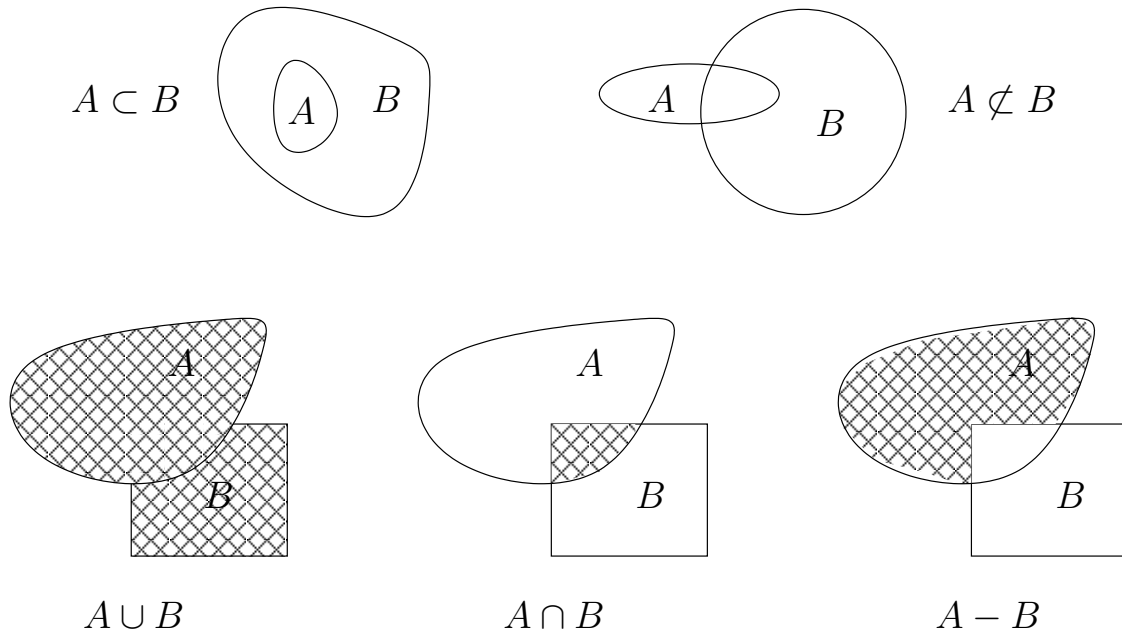


Abbildung 1.1: Venn-Diagramme.

- ii) Das kartesische Produkt ist intuitiv als **geordnetes Paar** geschrieben. Es gilt $(a, b) = (c, d) :\Leftrightarrow (a = c \text{ und } b = d)$.

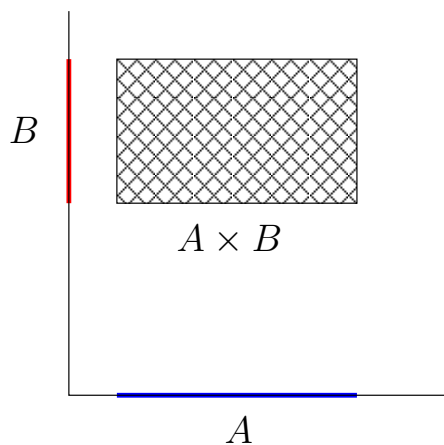


Abbildung 1.2: Veranschaulichung des kartesischen Produktes.

- iii) Die Potenzmenge ist eine **Menge von Mengen**.

- iv) Die Differenz $A - B$ heißt auch das **Komplement** von B in A .
Andere Schreibweisen: $C_A B$, $A \setminus B$.
- v) Zwei Mengen heißen **disjunkt**, falls $A \cap B = \emptyset$.
- vi) Vereinigung, Durchschnitt und kartesisches Produkt lassen sich unmittelbar auf mehrere Mengen verallgemeinern, Notation:

$$\bigcup_{k=1}^n A_k := A_1 \cup A_2 \cup \cdots \cup A_n ;$$

$$\bigcap_{k=1}^n A_k := A_1 \cap A_2 \cap \cdots \cap A_n ;$$

$$\prod_{k=1}^n A_k := A_1 \times A_2 \times \cdots \times A_n .$$

- vii) Beispielsweise werden im dreidimensionalen **Euklidischen**⁷ **Raum**⁸ $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ die **Koordinaten** eines Punktes $\underline{\mathbf{x}}^{(0)}$ als **geordnetes Tupel** $(x^{(0)}, y^{(0)}, z^{(0)})$ oder $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)})$ bezeichnet, die **Achsen** werden entsprechend mit x , y , z oder x_1 , x_2 , x_3 beschriftet. Die Achsen sind dabei in der Regel wie Daumen, Zeige- und Mittelfinger der rechten Hand angeordnet (vgl. Abbildung 1.3).

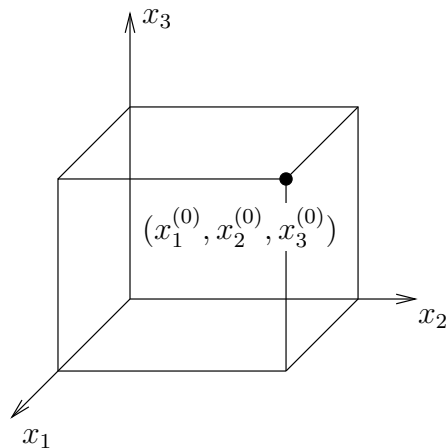


Abbildung 1.3: Der dreidimensionale Euklidische Raum.

In der **Raum-Zeit-Darstellung** erhält man ($n = 4$) das Tupel $(x^{(0)}, y^{(0)}, z^{(0)}, t^{(0)})$.

⁷Euklid, um 300 v. Chr.; Alexandria.

⁸ \mathbb{R} wird auch noch definiert.

viii) In den Übungen werden die **Kommutativgesetze**, das **Assoziativgesetz**, das **Distributivgesetz** und die **Regeln von DeMorgan** diskutiert.

ix) Man vereinbart $A \times \emptyset = \emptyset \times B = \emptyset$.

1.3 Funktionen (injektiv; surjektiv; bijektiv; Umkehrabbildung; Verkettung)

Die Beschreibung von Vorgängen in der realen Welt erfolgt mittels **Funktionen** (auch **Abbildungen** genannt).

Definition 1.3.1

Eine **Abbildung** f von einer Menge A in eine Menge B ordnet jedem Element x aus A **genau ein** Element y aus B zu. Notation:

$$f : A \rightarrow B, \quad y = f(x),$$

oder auch

$$f : A \ni x \mapsto f(x) \in B.$$

A heißt der **Definitionsbereich**, B der **Bildbereich** und $y = f(x)$ der **Bildpunkt** des **Urbildes** x unter f .

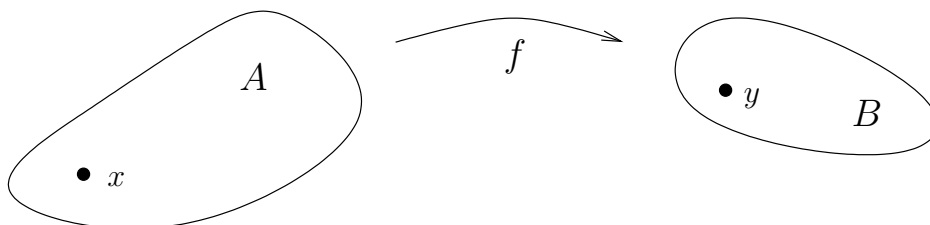


Abbildung 1.4: Eine Abbildung zwischen zwei Mengen.

Bemerkungen.

- i) Zwei verschiedenen Punkten x_1 und x_2 aus A kann durchaus derselbe Bildpunkt zugeordnet werden. Dann gilt $f(x_1) = f(x_2)$. An zwei Raumpunkten kann etwa die gleiche Temperatur herrschen. Am selben Ort können aber nicht unterschiedliche Temperaturen herrschen (deshalb “genau ein” in der Definition).
- ii) Es müssen nicht alle Punkte aus B Bildpunkte sein, Beispiel: $f: \mathbb{N} \rightarrow \mathbb{N}$, $f(n) \equiv 1^9$. Notation:

$$f(A) = \text{Bild}(f) = \{f(x) : x \in A\} \subset B.$$

- iii) $x \in A$ bezeichnet man auch als **Argument** oder **unabhängige Variable** von f , die Menge der geordneten Paare

$$\{(x, f(x)) : x \in A\}$$

heißt der **Graph** von f .

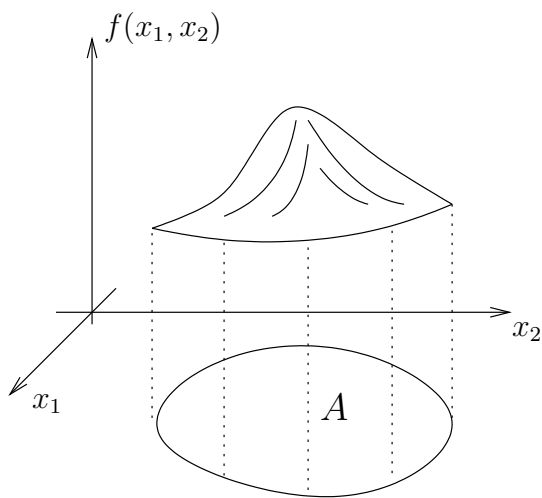


Abbildung 1.5: Der Graph einer Funktion.

Beispiele.

- i) $f: A \rightarrow A$, $f(x) := x \forall x \in A$ (die **Identität** $\text{Id}(A)$ von A).
- ii) $f: A \rightarrow B$, $f(x) := b \forall x \in A$ mit fixiertem $b \in B$ (**konstante Abbildung**).

⁹Das bedeutet: $f(n) = 1 \forall n \in \mathbb{N}$.

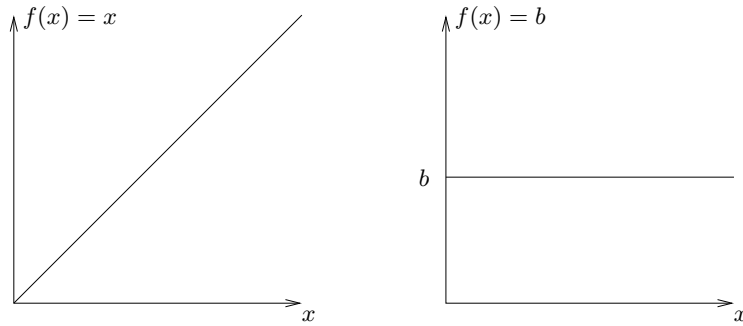


Abbildung 1.6: Identität und konstante Abbildung.

iii) Kein Graph ist aber die Menge G_1 aus Abbildung 1.7, da es Punkte $x \in A$ gibt, über denen mehrere Punkte liegen. Die Menge G_2 aus Abbildung 1.8 ist ebenfalls kein Graph, da es Punkte $x \in A$ gibt, über denen gar kein Punkt aus G_2 liegt.

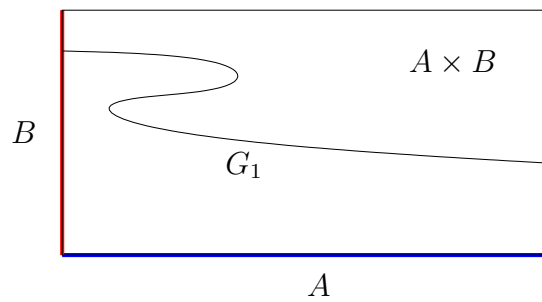


Abbildung 1.7: Kein Graph.

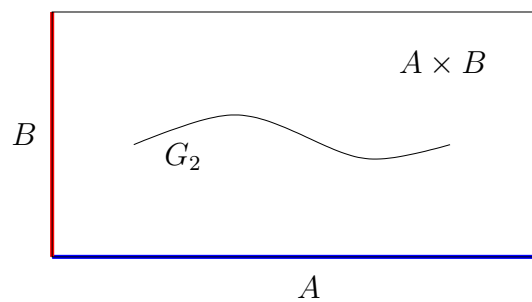


Abbildung 1.8: Kein Graph.

Definitions- und Bildbereich sind wesentliche Bestandteile einer Abbildung, beispielsweise hängt die folgende Definition stark davon ab:

Definition 1.3.2

Es sei f eine Abbildung von A nach B . Dann heißt f

- i) **injektiv**, falls aus $x_1, x_2 \in A$ mit $x_1 \neq x_2$ folgt $f(x_1) \neq f(x_2)$;
- ii) **surjektiv**, falls jeder Punkt von B ein Bildpunkt ist, d.h. $\text{Bild } f = B$;
- iii) **bijektiv**, falls f sowohl injektiv als auch surjektiv ist.

Ist f bijektiv, so existiert die **Umkehrabbildung** f^{-1} , definiert via $f^{-1}: B \rightarrow A$,

$$f^{-1}(y) := x \text{ für } y = f(x) .$$

Beispiele.

- i) Die Umkehrabbildung ist in Abbildung 1.9 veranschaulicht.

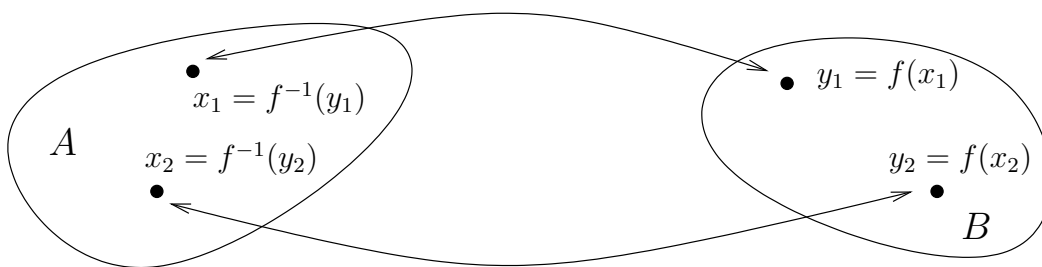


Abbildung 1.9: Zur Umkehrabbildung.

Nach Voraussetzung (bijektiv) existiert zu $y_1 \in B$ genau ein $x_1 \in A$ mit $y_1 = f(x_1)$ und man setzt $f^{-1}(y_1) := x_1$.

- ii) $f: \mathbb{R} \rightarrow \{y \in \mathbb{R} : y \geq 0\}$, $x \mapsto x^2$, ist surjektiv aber nicht injektiv.
- iii) $f: \{x \in \mathbb{R} : x \geq 0\} \rightarrow \{y \in \mathbb{R} : y \geq 0\}$, $x \mapsto x^2$, ist bijektiv.

iv) $f: \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$, ist weder injektiv noch surjektiv.

Schließlich können Funktionen in natürlicher Weise **verkettet** werden:

Definition 1.3.3

Sind drei Mengen A, B, C und sind zwei Funktionen $f: A \rightarrow B$ und $g: B \rightarrow C$ gegeben, so heißt die Abbildung

$$g \circ f : A \rightarrow C, \quad (g \circ f)(x) := g(f(x))$$

die **Hintereinanderschaltung** (**Verkettung**) von f und g .

Beispiele.

i) Die Verkettung wird in Abbildung 1.10 verdeutlicht.

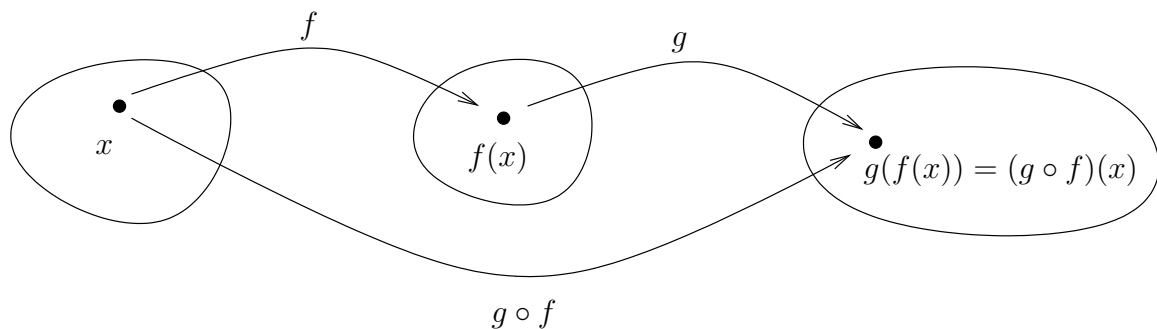


Abbildung 1.10: Zur Verkettung von Abbildungen.

ii) Ist $f: A \rightarrow B$ bijektiv, so gilt $(f^{-1} \circ f)(x) = x$ für alle $x \in A$ (siehe Definition 1.3.2).

Kapitel 2

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$

2.1 \mathbb{N}, \mathbb{Z} (Gruppen; vollständige Induktion; Kombinatorik; Primzahlen; abzählbar unendliche Mengen)

Das Begriff der **natürlichen Zahlen** \mathbb{N} ist intuitiv bekannt, sie sind im Wesentlichen¹ charakterisiert durch die Existenz einer “**Nachfolgeabbildung**”:

- i) 1 ist eine natürliche Zahl;
- ii) zu jeder natürlichen Zahl n gibt es eine nachfolgende natürliche Zahl $(n + 1)$.

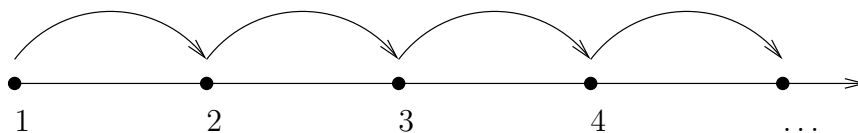


Abbildung 2.1: Zur Nachfolgeabbildung.

Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ sind also per definitionem mit einer **Addition** $+$ versehen

Problem. Zu gegebenen $n, m \in \mathbb{N}$ ist die Gleichung

$$m + z = n$$

¹Genauer: Die natürlichen Zahlen sind durch die **Peano-Axiome** gekennzeichnet (G. Peano, 1858-1932; Turin). (1) $1 \in \mathbb{N}$; (2) $n \in \mathbb{N} \Rightarrow (n + 1) \in \mathbb{N}$; (3) $n \neq m \Rightarrow (n + 1) \neq (m + 1)$; (4) $n \in \mathbb{N} \Rightarrow n + 1 \neq 1$; (5) für jede Teilmenge $A \subset \mathbb{N}$ gilt: $(1 \in A) \wedge (\forall n : [n \in A \Rightarrow (n + 1) \in A]) \Rightarrow A = \mathbb{N}$.

nicht immer in \mathbb{N} lösbar (eine **Subtraktion** wird benötigt). Die Lösungen solcher Gleichungen erweitern \mathbb{N} auf die **ganzen Zahlen \mathbb{Z}** :²

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

Die ganzen Zahlen haben eine besondere Struktur: $(\mathbb{Z}, +)$ ist eine **kommutative** (oder **Abelsche**)³ **Gruppe**, d.h. es gelten die folgenden Rechenregeln:

Definition 2.1.1

Unter einer **Gruppe** (\mathbb{G}, \circ) versteht man eine Menge \mathbb{G} versehen mit einer Verknüpfung \circ , die jedem geordneten Paar von Elementen g_1 und g_2 aus \mathbb{G} ein Element $g_1 \circ g_2$ **aus \mathbb{G}** (**Abgeschlossenheit**) zuordnet, sodass die folgenden Regeln gelten:

i) Es gibt genau ein **neutrales Element** $e \in \mathbb{G}$, sodass für alle $g \in \mathbb{G}$

$$e \circ g = g \circ e = g.$$

ii) Zu jedem $g \in \mathbb{G}$ existiert genau ein Element $\tilde{g} \in \mathbb{G}$ (**inverses Element**) mit

$$g \circ \tilde{g} = \tilde{g} \circ g = e.$$

iii) Für alle g_1, g_2, g_3 aus \mathbb{G} ist (**Assoziativgesetz**)

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3.$$

Eine Gruppe heißt **kommutativ** (oder **Abelsch**), falls **zusätzlich** gilt: Für alle g_1, g_2 aus \mathbb{G} ist

$$g_1 \circ g_2 = g_2 \circ g_1.$$

²Genauer: Vergleiche Axiomatik aus Paragraph 2.3.

³N.H. Abel, 1802-1829; Berlin, Paris.

Bemerkungen.

- i)* Im Fall $(\mathbb{Z}, +)$ ist in Definition 2.1.1 \mathbb{G} durch \mathbb{Z} und \circ durch $+$ zu ersetzen, man erhält wohl bekannte Rechenregeln.
- ii)* Das neutrale Element e ist in diesem Fall die 0.
- iii)* Das inverse Element \tilde{z} einer ganzen Zahl z wird als $-z$ bezeichnet (Schreibweise: $m + (-z) = m - z$). Die (eindeutige) Lösung der Gleichung $m + z = n$ ist $z = -m + n$.
- iv)* Auch in den Anwendungen treten zahlreiche Beispiele unterschiedlicher Gruppen auf (Symmetriegruppen, Transformationsgruppen, Permutationsgruppen Welche Eigenschaften sollte man von Symmetrieoperationen erwarten?).

Die natürlichen Zahlen (\mathbb{Z} ebenso) sind zudem mit einer **Multiplikation**⁴ versehen, als Abbildung geschrieben:

$$\cdot : \mathbb{N} \times \mathbb{N} \ni (m, n) \mapsto m \cdot n \in \mathbb{N}$$

(der Malpunkt wird oft weggelassen). Es gelten die bekannten Rechenregeln. Zudem gibt es eine **Ordnungsrelation**⁵ “ $<$ ” mit den Eigenschaften

- i)* $n < n + 1$ (n kleiner als $n + 1$);
- ii)* je zwei ganze Zahlen m, n sind **vergleichbar**, d.h.: Entweder gilt $m < n$ oder $n < m$ oder $m = n$ (“kleiner gleich”, “größer”, “größer gleich”: $m \leq n :\Leftrightarrow m < n$ oder $m = n$; $m > n :\Leftrightarrow n < m$; $m \geq n :\Leftrightarrow n \leq m$).

Eine wesentliche Eigenschaft von \mathbb{N} ist: Eine Teilmenge von \mathbb{N} , die die 1 enthält und mit jedem Element auch dessen Nachfolger, ist notwendig gleich \mathbb{N} . Auf dieser Eigenschaft basiert das **Beweisprinzip der vollständigen Induktion**. Es erinnert an eine Reihe von Dominosteinen, die so aufgestellt sind, dass, wenn irgendein Stein fällt (etwa der mit der Nummer n), auch der nachfolgende (also mit der Nummer $n + 1$) fällt (**Induktionsschluss**). Stößt man den ersten Stein an (**Induktionsanfang**), so werden alle Steine fallen.

⁴Präzise Einführung in Paragraph 2.3.

⁵Präzise

Satz 2.1.1 (Prinzip der vollständigen Induktion)

Es sei $n_0 \in \mathbb{N}$ (meist $n_0 = 1$). Für jede natürliche Zahl $n \geq n_0$ sei eine Aussage $A(n)$ gegeben. **Alle** Aussagen $A(n)$, $n \geq n_0$, sind wahr, wenn man zeigen kann:

- i) **Induktionsanfang**: $A(n_0)$ ist richtig.
- ii) **Induktionsschluss**: Ist $n \geq n_0$, so folgt aus der **Hypothese** “ $A(n)$ ist wahr”, dass auch $A(n + 1)$ wahr ist.

Beispiel. *Behauptung.* Es gilt für beliebiges $n \in \mathbb{N}$

$$1 + 2 + 3 + 4 + \dots + n = \frac{n(n+1)}{2} .^6$$

Beweis. Für Zahlen $a_m, \dots, a_n \in \mathbb{N}$ (später analog aus \mathbb{R}) vereinbart man die Notation

$$\sum_{k=m}^n a_k := a_m + \dots + a_n \quad (\text{Summe von } a_m \text{ bis } a_n, m \leq n)$$

($\sum_{k=m}^n a_k = 0$ falls $m > n$). Die Aussage $A(n)$, die für beliebiges n mit dem Induktionsprinzip zu beweisen ist, lautet mit dieser Notation

$$A(n) : \sum_{k=1}^n k = \frac{n(n+1)}{2} .$$

Induktionsanfang: Die Aussage $A(1)$, ist offensichtlich wahr:

$$\sum_{k=1}^1 k = \frac{1(1+1)}{2}$$

Induktionsschluss ($A(n) \Rightarrow A(n+1)$): Man nehme nun an, dass $A(n)$ für ein n (beliebig aber fest) wahr ist. Was folgt dann für $A(n+1)$?

⁶In \mathbb{N} müsste man streng genommen zeigen: $2(1 + 2 + 3 + 4 + \dots + n) = n(n+1)$.

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \underbrace{\sum_{k=1}^n k}_{=n(n+1)/2 \text{ da } A(n) \text{ nach Annahme wahr ist}} + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) = (n+1) \left(\frac{n}{2} + 1 \right) = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Das ist genau die Aussage $A(n+1)$, die also automatisch richtig ist, falls $A(n)$ richtig ist.

Nach dem Induktionsprinzip ist die Behauptung bewiesen. \square

Mit Hilfe der vollständigen Induktion können insbesondere viele Aussagen aus der **Kombinatorik** gezeigt werden.

Frage 1: Wie viele Teilmengen hat eine n -elementige Menge?

Satz 2.1.2 ⁷

Eine n -elementige Menge $A = \{a_1, \dots, a_n\}$ besitzt 2^n verschiedene Teilmengen (Die Potenzmenge $\mathcal{P}(A)$ hat 2^n Elemente).

Beweis durch vollständige Induktion.

Induktionsanfang ($n = 1$): Eine Menge bestehend aus einem Element, $A = \{a_1\}$, hat zwei Teilmengen, nämlich \emptyset und $\{a_1\}$. Der Induktionsanfang ist demnach richtig.

Induktionsschluss ($A(n) \Rightarrow A(n+1)$): Es sei nun als richtig angenommen, dass eine n -elementige Menge 2^n Teilmengen hat. Zu zeigen ist: Eine $(n+1)$ -elementige Menge, $A = \{a_1, \dots, a_{n+1}\}$, hat 2^{n+1} Teilmengen.

Dazu teilt man die Teilmengen von A (die Potenzmenge) in zwei Klassen (Teilmengen) K_1 und K_2 :

$$T \in K_1 : \Leftrightarrow a_{n+1} \notin T, \quad T \in K_2 : \Leftrightarrow a_{n+1} \in T.$$

⁷ $2^n := \underbrace{2 \cdot 2 \cdot \dots \cdot 2}_n$.

Beispiel. Ist $A = \{1, 2, 3\}$, so sind

$$\begin{aligned} T \in K_1: & \quad \emptyset, \{1\}, \{2\}, \{1, 2\}, \\ T \in K_2: & \quad \{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}. \end{aligned}$$

Fortführung des Beweises. Es besteht K_1 aus den Teilmengen von $A' := \{a_1, \dots, a_n\}$, entsprechend hat K_1 nach **Induktionsannahme** 2^n Elemente.

Ist T aus K_2 , so sieht T wie folgt aus:⁸

$$T = \{a_{i_1}, \dots, a_{i_k}, a_{n+1}\} = B \cup \{a_{n+1}\}, \quad B \subset A'.$$

Da B eine beliebige Teilmenge aus A' sein kann, hat auch K_2 insgesamt 2^n Elemente.

Die Klassen K_1 und K_2 haben offensichtlich keine Mengen gemeinsam, es ergeben sich genau wie behauptet $2^n + 2^n = 2^{n+1}$ Teilmengen von A . \square

Frage 2: Wie viele Möglichkeiten gibt es im Fußballtoto?

Ein Spieltag von 11 Spielen kann als **geordnetes** Tupel der Länge 11 aufgefasst werden, für jeden Eintrag gibt es drei Möglichkeiten. Nach dem folgenden Satz kann man den Tippschein auf $3^{11} = 177147$ verschiedene Arten ausfüllen.

Satz 2.1.3

Es sei A eine Menge mit n Elementen. Dann gibt es n^k verschiedene geordnete k -Tupel (a_1, a_2, \dots, a_k) von Elementen $a_j \in A$, $j = 1, \dots, k$.

Beweisidee. Fixiere n , Induktion **nach k** (Übungsaufgabe). \square

Sprechweise. Das Tupel (a_1, \dots, a_k) aus Satz 2.1.3 heißt eine **k -Variation von n Elementen**.

Im letzten Beispiel wurden geordnete Tupel betrachtet:

⁸ $i_j \in \{1, 2, \dots, n\}$ für $1 \leq j \leq k$; $i_j < i_{j+1}$ für $1 \leq j < k$.

Frage 3: Was ist mit Vertauschungen innerhalb eines Tupels?

Sprechweise. Eine Vertauschung innerhalb eines Tupels nennt man **Permutation**. Zu $(1, 2, 3)$ gibt es beispielsweise die 6 Permutationen $(1, 2, 3)$, $(1, 3, 2)$, $(2, 3, 1)$, $(2, 1, 3)$, $(3, 1, 2)$, $(3, 2, 1)$.

Satz 2.1.4⁹

Es gibt

$$n! := \prod_{k=1}^n k := 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$$

*Permutationen des n -Tupels $(1, 2, 3, \dots, n)$ (bzw. von $(a_1, a_2, a_3, \dots, a_n)$ mit **paarweise verschiedenen** a_1, \dots, a_n).*

Beweisidee. Induktion: Wie viele Permutationen mit “ a_{n+1} an erster Stelle”, “ a_{n+1} an zweiter Stelle” etc. gibt es dabei nach Induktionsannahme? \square

Beispiel. Aus einer Urne mit n nummerierten, ansonsten aber gleichen Kugeln werden $m \leq n$ Kugeln gezogen und herausgelegt. Wie viele unterschiedliche Möglichkeiten gibt es, falls die Reihenfolge der Ziehungen keine Rolle spielt? Die Anzahl der Möglichkeiten ist gleich der Anzahl m -elementiger Teilmengen von $\{1, \dots, n\}$ (m Nummern der gezogenen Kugeln sind anzugeben).

Ist $1 \leq m < n$, so gibt es $n \cdot (n - 1) \cdot \dots \cdot (n - m + 1)$ verschiedene m -Tupel der Zahlen $1, 2, \dots, n$ mit paarweise verschiedenen Elementen (analog zu Satz 2.1.4; man überlege sich ein Beispiel).

Man teilt diese Tupel so in verschiedene Klassen ein, dass die Tupel in einer Klasse sich nur in der Reihenfolge der Eintragungen unterscheiden, jede Klasse entspricht damit einer m -elementigen Menge.

⁹Sprechweise: n -Fakultät; Vereinbarung: $0! := 1$; II: **Produktzeichen**.

Nach Satz 2.1.4 besteht jede Klasse aus $m!$ Tupeln, es gibt somit $n \cdot (n - 1) \cdot \dots \cdot (n - m + 1)/m!$ Klassen und damit m -elementige Teilmengen.

Zusammenfassend ist gezeigt:

Korollar 2.1.1 ¹⁰

Es gibt zu $1 \leq m \leq n$

$$\binom{n}{m} := \frac{n \cdot (n - 1) \cdot \dots \cdot (n - m + 1)}{m!} = \frac{n!}{(n - m)! \cdot m!}$$

m -elementige Teilmengen einer n -elementigen Menge.

Die natürlichen Zahlen $\binom{n}{m}$ heißen Binomialkoeffizienten.

Für das Rechnen mit Binomialkoeffizienten beachtet man, dass nach Definition gilt $\binom{n}{0} = \binom{n}{n} = 1$. Außerdem folgt mit der Konvention $\binom{l}{m} = 0$ falls $l < m$ die Rekursionsformel ($m, n \in \mathbb{N}$, $1 \leq m < n$)

$$\begin{aligned} \binom{n}{m} + \binom{n}{m-1} &= \frac{n!}{m!(n-m)!} + \frac{n!}{(m-1)!(n-m+1)!} \\ &= \frac{n!(n+1-m) + n!m}{m!(n+1-m)!} \\ &= \frac{n!(n+1-m+m)}{m!(n+1-m)!} \\ &= \frac{(n+1)!}{m!(n+1-m)!} \\ &= \binom{n+1}{m}. \end{aligned}$$

¹⁰Ohne Kommentar wird hier dividiert. Man beachte aber, dass die resultierenden Objekte natürliche Zahlen sind.

Damit können die Binomialkoeffizienten leicht mit Hilfe des **Pascalschen**¹¹ **Dreiecks** bestimmt werden (vgl. Abbildung 2.2).

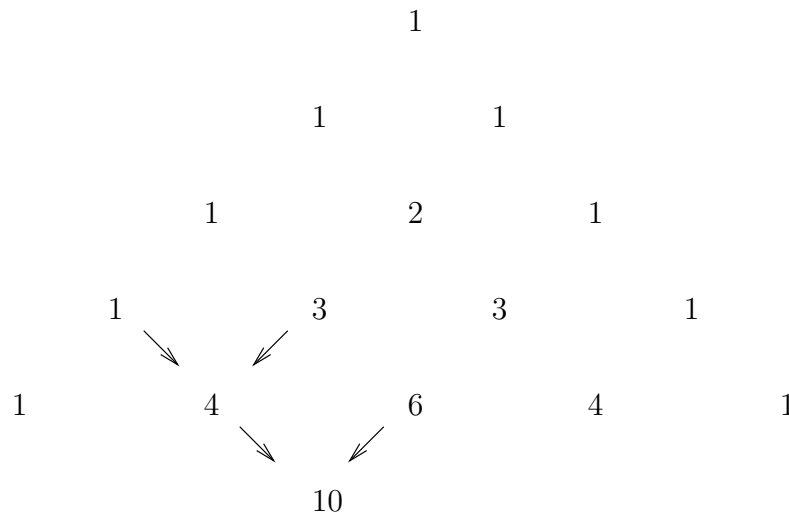


Abbildung 2.2: Das Pascalsche Dreieck.

Eine wichtige Folgerung ist der **binomische Lehrsatz**, der mit der Rekursionsformel und vollständiger Induktion bewiesen wird.

Satz 2.1.5 (*binomischer Lehrsatz*)

Für alle Zahlen a, b (auch reelle oder komplexe Zahlen) und für alle $n \in \mathbb{N}$ gilt

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k .$$

Als weitere Anwendung der vollständigen Induktion sei noch kurz auf die **Primzahlen** eingegangen:

Definition 2.1.2

*Eine natürliche Zahl $m \in \mathbb{N}$ heißt **Teiler** von $n \in \mathbb{N}$, falls es ein $k \in \mathbb{N}$ gibt mit $n = m \cdot k$. Ist $n > 1$ und besitzt n außer $n, 1$ keine Teiler, so heißt n eine **Primzahl**.*

Eine wesentliche Eigenschaft von Primzahlen ist

¹¹B. Pascal, 1623-1662; Paris.

Satz 2.1.6

Jede Zahl $n \in \mathbb{N}$, $n \geq 2$, lässt sich als *Produkt von Primzahlpotenzen* schreiben:

$$n = p_1^{r_1} \cdot p_2^{r_2} \cdot \cdots \cdot p_k^{r_k} .$$

Dabei ist $k \in \mathbb{N}$, für $j = 1, 2, \dots, k$ sind die p_j Primzahlen und $r_j \in \mathbb{N}$.

Beweis. Vollständige Induktion. □

Schließlich noch die

Sprechweise. Eine Menge A , die **gleichmächtig** wie \mathbb{N} ist, heißt **abzählbar (unendlich)**, d.h.: Es existiert eine bijektive Abbildung (**Bijektion**) $\Phi: \mathbb{N} \rightarrow A$, mit anderen Worten: Die natürlichen Zahlen und die Elemente der Menge A können eineindeutig einander zugeordnet werden.

2.2 \mathbb{Q} (Körper; indirekter Beweis)

Die Gleichung $m + z = n$ für gegebene $m, n \in \mathbb{N}$ motiviert die Erweiterung der natürlichen Zahlen auf die ganzen Zahlen. Analog ist für gegebene $z, w \in \mathbb{Z}$ die Gleichung

$$z \cdot q = w$$

i.A. nicht in \mathbb{Z} lösbar (eine **Division** wird benötigt). Man erweitert daher \mathbb{Z} auf die Menge der **rationalen Zahlen** \mathbb{Q} :¹²

$$\mathbb{Q} := \left\{ q : q = \frac{m}{n} = m \cdot n^{-1}, m \in \mathbb{Z}, n \in \mathbb{N} \right\} .$$

Hier spielt zusätzlich die Multiplikation bei der Definition eine besondere Rolle, als spezielle Struktur ist festzuhalten: $(\mathbb{Q}, +, \cdot)$ ist ein **Körper**, d.h. es gelten die Rechenregeln aus Definition 2.2.1.

¹²Wieder sei auf Abschnitt 2.3 verwiesen.

Definition 2.2.1

Unter einem **Körper** $(\mathbb{K}, +, \cdot)$ versteht man eine Menge K versehen mit einer Addition $+$ und einer Multiplikation \cdot , sodass gilt:

- i) \mathbb{K} ist **bzgl. der Addition** eine kommutative Gruppe mit neutralem Element (**Nullelement**) 0 .
- ii) $\mathbb{K} - \{0\}$ ist **bzgl. der Multiplikation** eine kommutative Gruppe mit neutralem Element (**Einselement**) 1 .
- iii) Für alle a, b, c aus \mathbb{K} gilt (**Distributivgesetz**):

$$a \cdot (b + c) = \underbrace{a \cdot b + a \cdot c}_{\text{Vereinbarung: Punkt- vor Strichrechnung}}$$

Bemerkungen.

- i) Punkt i) der Definition 2.2.1 ist völlig analog zu der Aussage “ $(\mathbb{Z}, +)$ ist kommutative Gruppe” (siehe Definition 2.1.1 und die anschließenden Bemerkungen).
- ii) **Zusätzlich** ist eine ebenfalls analoge Gruppeneigenschaft bzgl. der Multiplikation gefordert, wobei das neutrale Element bzgl. der Multiplikation suggestiv als 1 bezeichnet wird ($q \cdot 1 = q$). Das inverse Element einer rationalen Zahl q bzgl. der Multiplikation wird als q^{-1} oder $\frac{1}{q}$ bezeichnet ($\frac{q}{q} = 1$). Zur 0 (neutrales Element bzgl. der Addition) gibt es kein inverses Element, deshalb kann lediglich $\mathbb{K} - \{0\}$ eine Gruppe bzgl. der Multiplikation sein.
- iii) Die (eindeutige) Lösung der Gleichung $z \cdot q = w$ für gegebene $z \neq 0, w \in \mathbb{Q}$ ist $q = z^{-1}w$.
- iv) Typische Beispiele anderer Körper sind \mathbb{R} (vgl. Paragraph 2.3) oder \mathbb{C} (vgl. Kapitel 7). Ein weiteres Beispiel wird in den Übungen diskutiert.

v) Die Anordnungsrelationen übertragen sich auf \mathbb{Q} .

Natürlich enthält \mathbb{Q} viel mehr Elemente als \mathbb{N} , ist \mathbb{Q} aber auch mächtiger als \mathbb{N} ? Mit anderen Worten:

Frage: Ist \mathbb{Q} abzählbar unendlich?

Nach Abbildung 2.3 lautet die

Antwort: Ja, d.h. \mathbb{N} , \mathbb{Z} und \mathbb{Q} sind gleichmächtig.

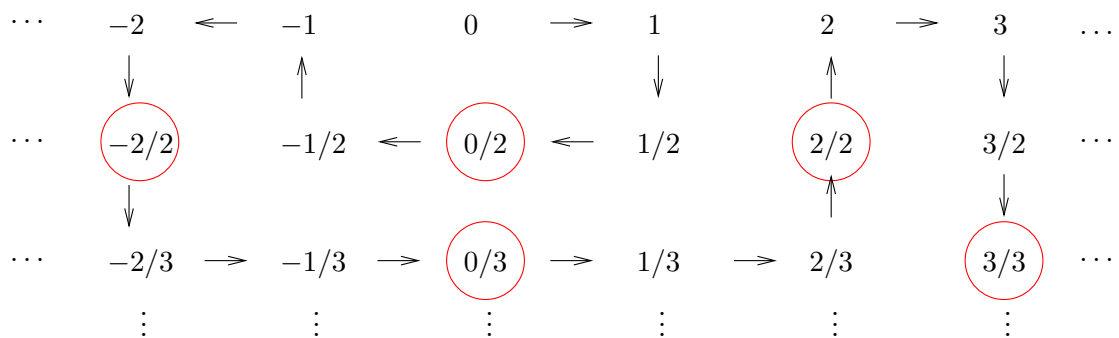


Abbildung 2.3: \mathbb{Q} ist abzählbar.

Man kann \mathbb{Q} wie angedeutet durchnummerieren, wobei die rot umkreisten Brüche nicht mitgezählt werden, da sie bereits vorgekommen sind. Beispielsweise erhält man in dem Schema aus Abbildung 2.3 für $-2/3$ die Nummer 7.

Bemerkung. Allgemein gilt: Die Vereinigung von **abzählbar vielen** abzählbaren Mengen ist abzählbar.

Um zu zeigen, dass in den rationalen Zahlen die Gleichung $q^2 = 2$ nicht lösbar ist, dass also noch die **reellen Zahlen** einzuführen sind, benutzt man eine weitere Beweismethode, den **indirekten Beweis**. Dabei **wird das Gegenteil der Behauptung angenommen** und daraus ein **Widerspruch** hergeleitet.

Behauptung. Es gibt keine rationale Zahl q mit $q^2 = 2$.

Beweis. Angenommen es gibt doch eine solche rationale Zahl, so schreibe man diese als $q = m/n$, wobei m und n **teilerfremd** gewählt werden können (kürzen!). Aus der **Annahme**

$$\left(\frac{m}{n}\right)^2 = 2$$

folgt: $m^2 = 2 \cdot n^2$ ist eine gerade Zahl, also ist auch m selbst eine gerade Zahl, d.h. für ein $k \in \mathbb{N}$ gilt $m = 2 \cdot k$. Setzt man dies ein, so ergibt sich

$$4 \cdot k^2 = 2 \cdot n^2, \text{ also auch } n^2 = 2 \cdot k^2.$$

Damit ist n^2 gerade, dies impliziert: n ist gerade.

Wenn m und n gerade sind, dann haben sie den gemeinsamen Teiler 2, was ein Widerspruch zur Annahme “teilerfremd” ist. \square

2.3 \mathbb{R} (Axiomatik; Anordnung; Vollständigkeit; Rechnen mit Ungleichungen; Längenmessung; Überabzählbarkeit)

Als typisches Beispiel dient in diesem Kapitel die kontinuierlich ablaufende Zeit. Dabei geht es nicht um die philosophische Frage “Was ist die Zeit?” sondern darum, welchen grundsätzlichen Gesetzmäßigkeiten die Zeit gehorcht, d.h. nach welchen Spielregeln kann man mit der Zeit operieren? Diese Spielregeln sollten **möglichst**

- knapp,
- vollständig,
- widerspruchsfrei

sein, um anschließend Vorhersagen für die Zukunft abzuleiten. Man nennt sie **Axiome**. Die Axiome sind nicht herzuleiten, sie werden als grundlegende Aussagen als **wahr angenommen**.

Betrachtet man die Weg- Zeitabhängigkeit des freien Falls, d.h.

$$\text{Weg} = \frac{1}{2} \cdot \text{Erdbeschleunigung} \cdot \text{Zeit}^2,$$

so müssen geeignete Axiome über die Eigenschaften der rationalen Zahlen hinausgehen, um die Lösbarkeit quadratischer Gleichungen zu sichern.

Vorgegeben werden: Die **Axiome der reellen Zahlen**.

Es gibt eine Menge \mathbb{R} von Elementen, genannt die **reellen Zahlen**, die die folgenden **Axiome** erfüllen.

- i*) (**Algebraische Axiome**) Die reellen Zahlen sind ein **Körper**.
- ii*) (**Anordnungsaxiome**) Es existiert eine **Ordnungsrelation** “ $<$ ” (vgl. Paragraph 2.1), die den folgenden Axiomen genügt:
 - (a) Aus $x < y$ und $y < z$ folgt $x < z$ (**Transitivität**).
 - (b) Aus $x < y$ folgt $x + z < y + z$ für alle $z \in \mathbb{R}$ (**Verträglichkeit mit der Addition**).
 - (c) Aus $x < y$ und $z > 0$ folgt $x \cdot z < y \cdot z$ (**Verträglichkeit mit der Multiplikation**).
- iii*) (**Vollständigkeitsaxiom**) Jede **nicht-leere**, nach **oben beschränkte** Teilmenge A besitzt eine **kleinste obere Schranke**, genannt das **Supremum** von A , $\sup A$. Dabei heißt eine nicht-leere Teilmenge $A \subset \mathbb{R}$ **nach oben beschränkt** (bzw. **nach unten beschränkt**), wenn es eine Zahl $k \in \mathbb{R}$ gibt, sodass gilt

$$x \leq k \quad (\text{bzw. } k \leq x) \quad \forall x \in A.$$

Eine Zahl $k \in \mathbb{R}$ heißt **kleinste obere Schranke** (bzw. **größte untere Schranke**) einer nicht-leeren Teilmenge $A \subset \mathbb{R}$, wenn sie eine obere (bzw. untere) Schranke ist und wenn es keine kleinere obere (bzw. größere untere) Schranke gibt.

Bemerkungen.

- i) Die Bedeutung obiger Axiome mache man sich anhand des Beispiels “Zeit” klar.
- ii) Die rationalen Zahlen \mathbb{Q} sind ebenfalls ein angeordneter Körper, der aber nicht vollständig¹³ ist. Beispielsweise hat die Menge $\{q \in \mathbb{Q} : q < \sqrt{2}\}$ keine kleinste obere Schranke in \mathbb{Q} . Anschaulich sieht man das am **Zahlenstrahl**¹⁴ (vgl. Abbildung 2.4).

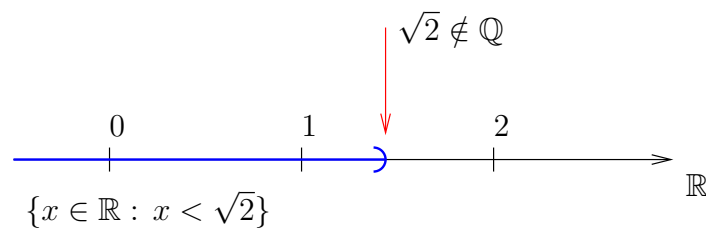


Abbildung 2.4: Zum Vollständigkeitsaxiom.

- iii) Die reellen Zahlen umfassen die rationalen Zahlen, im Vergleich zu den rationalen Zahlen kommen Wurzeln und **transzendente Zahlen**¹⁵ hinzu, es werden so die “Lücken im Zahlenstrahl” gefüllt.
- iv) Die Elemente aus $\mathbb{R} - \mathbb{Q}$ heißen die **irrationalen Zahlen**.
- v) \mathbb{R} ist ein **archimedisch**¹⁶ angeordneter Körper, d.h. zu jedem $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$, sodass $x < n$ gilt.
- vi) Die größte untere Schranke einer nichtleeren Menge A heißt das **Infimum** von A , $\inf A$.
- vii) Es wird nicht verlangt, dass das Supremum (Infimum) einer Menge zu dieser gehört. Ist das jedoch der Fall, so spricht man vom **Maximum** (**Minimum**) von A , $\max A$ ($\min A$).
- viii) Ist A nach oben und unten beschränkt, so heißt A **beschränkt**.

¹³R. Dedekind, 1831-1916; Göttingen, Braunschweig.

¹⁴Die Richtung des Pfeils auf dem Zahlenstrahl symbolisiert die Anordnungsrelation.

¹⁵Diese Zahlen sind nicht Lösungen einer **algebraischen Gleichung**. So zeigte etwa Ferdinand von Lindemann (1852-1939; Erlangen Würzburg, Freiburg, Königsberg und München) die Transzendenz der **Kreiszahl** π .

¹⁶Archimedes, um 285 v. Chr.; Syrakus.

Die Vorstellung, dass der Zahlenstrahl keine ‐Löcher‐ hat, entspricht dem **Intervallschachtelungsprinzip**, welches aus dem Vollständigkeitsaxiom folgt. Ein **Intervall** ist eine Menge der Form ($a, b \in \mathbb{R}$ mit $a < b$)

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} \quad (\text{abgeschlossen}) ; \\ (a, b] &:= \{x \in \mathbb{R} : a < x \leq b\} \quad (\text{halboffen}) ; \\ [a, b) &:= \{x \in \mathbb{R} : a \leq x < b\} \quad (\text{halboffen}) ; \\ (a, b) &:= \{x \in \mathbb{R} : a < x < b\} \quad (\text{offen}) . \end{aligned}$$

Die Idee des Intervallschachtelungsprinzips ist in Abbildung 2.5 verdeutlicht.

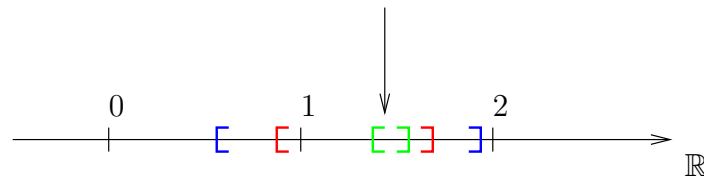


Abbildung 2.5: Zum Intervallschachtelungsprinzip.

Anschaulich gesprochen bedeutet das: Betrachtet man abzählbar unendlich viele abgeschlossene Intervalle I_n , die sich enthalten ($I_{n+1} \subset I_n \forall n \in \mathbb{N}$), und werden die Intervalle für große n ‐beliebig klein‐, so existiert genau eine reelle Zahl x im Durchschnitt all dieser Intervalle (vgl. wieder Abbildung 2.5).

Wie rechnet man mit Ungleichungen und Beträgen?

Aus den Axiomen folgen Rechenregeln für die reellen Zahlen, z.B.:

- i) $x \cdot 0 = 0$;
- ii) $x \leq y \Rightarrow -x \geq -y$ (analog folgt $x^{-1} \geq y^{-1}$);
- iii) $(x \leq y) \wedge (z \leq 0) \Rightarrow x \cdot z \geq y \cdot z$;
- iv) $x^2 \geq 0$;
- v) $(x \leq y) \wedge (u \leq v) \Rightarrow x + u \leq y + v$;

$$vi) (0 \leq x \leq y) \wedge (0 \leq u \leq v) \Rightarrow x \cdot u \leq y \cdot v.$$

Exemplarisch seien die ersten beiden Aussagen bewiesen:¹⁷

Ad *i*): Die Null ist das neutrale Element der Addition, und aus dem Distributivgesetz folgt

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0.$$

Das inverse Element bzgl. der Addition zu $x \cdot 0$ ist $-x \cdot 0$ und man erhält mit dem Assoziativgesetz:

$$0 = x \cdot 0 - x \cdot 0 = (x \cdot 0 + x \cdot 0) - x \cdot 0 = x \cdot 0 + (x \cdot 0 - x \cdot 0) = x \cdot 0.$$

Ad *ii*): Gilt in der Voraussetzung die Gleichheit, so ist die Aussage sicher richtig. Es sei also $x < y$. Aus dem Anordnungsaxiom (b) folgt (Assoziativität und Kommutativität beachten):

$$\begin{aligned} -y &= (x - x) - y = x + (-x - y) < y + (-x - y) \\ &= y + (-y - x) = (y - y) - x = -x, \end{aligned}$$

also die Behauptung. Die weiteren Regeln sind ähnlich zu zeigen. \square

Eine wesentliche Operation in der **Analysis**¹⁸ ist das **Messen von Längen und Abständen** (zunächst auf dem Zahlenstrahl):

Definition 2.3.1

Der **Betrag** oder **Absolutbetrag** einer reellen Zahl x ist definiert als

$$|x| := \begin{cases} x & \text{falls } x \geq 0, \\ -x & \text{falls } x < 0. \end{cases}$$

Für $x, y \in \mathbb{R}$ ist $|x - y|$ der (nicht-negative) **Abstand** auf dem Zahlenstrahl.

¹⁷Die erste Aussage beispielsweise klingt absolut trivial, aber folgt sie wirklich, wenn man nur mit den Axiomen argumentiert?

¹⁸Grundaufgabe der Analysis: Untersuchung von Funktionen einer oder mehrerer Veränderlicher.

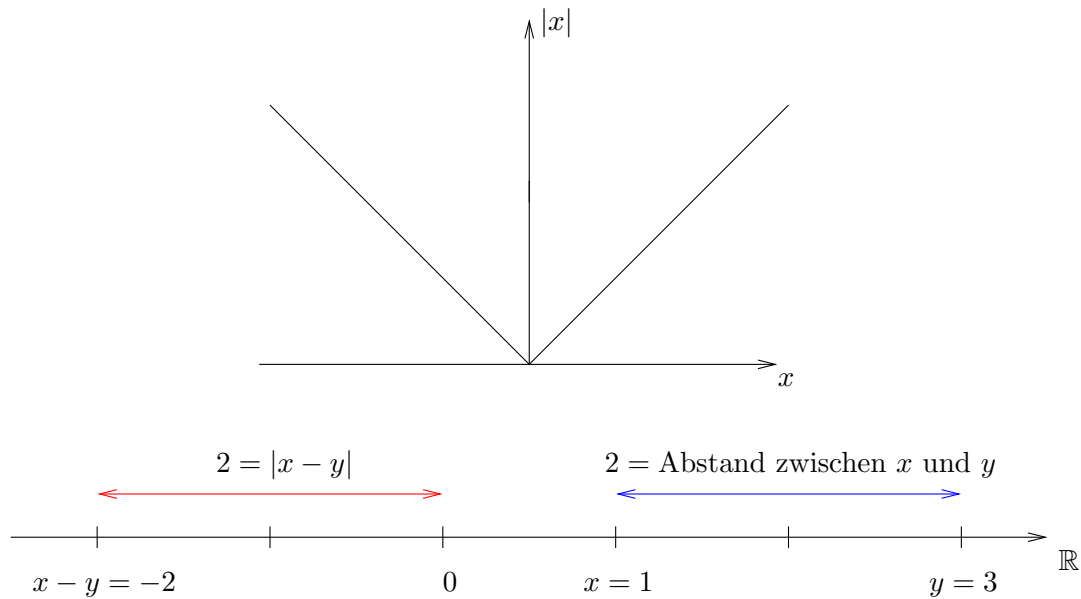


Abbildung 2.6: Zur Abstandsfunktion.

Eigenschaften und Rechenregeln. Für alle $x, y \in \mathbb{R}$ gilt:

- i) $|x| \geq 0$;
- ii) $|x| = 0 \Leftrightarrow x = 0$;
- iii) $|x \cdot y| = |x||y|$;
- iv) $|x + y| \leq |x| + |y|$ (**Dreiecksungleichung**);
- v) $||x| - |y|| \leq |x - y|$ (**umgekehrte Dreiecksungleichung**).

Beweis. Die ersten vier Aussagen sind direkt einsichtig, zum Beweis von v) beobachtet man (nach iv)):

$$|x| = |x - y + y| \leq |x - y| + |y| \Rightarrow |x| - |y| \leq |x - y| .$$

Ebenso gilt

$$|y| = |y - x + x| \leq |x - y| + |x| \Rightarrow |y| - |x| \leq |x - y| ,$$

d.h. v) ist richtig. □

Wie sieht die **Verteilung** der rationalen und der irrationalen Zahlen **auf dem Zahlenstrahl** aus?

Ohne Beweis sei dazu der folgende Satz angegeben.

Satz 2.3.1

*i) Sind $x, y \in \mathbb{R}$ und gilt $x < y$, so gibt es eine **rationale** Zahl $q \in \mathbb{Q}$ mit $x < q < y$. Man sagt: **Die rationalen Zahlen liegen dicht in \mathbb{R} .***

*ii) \mathbb{R} ist **nicht abzählbar**, man sagt auch \mathbb{R} ist **überabzählbar**.*

Bemerkungen. Anschaulich besagt Satz 2.3.1:

- i)* Egal wie dicht zwei reelle Zahlen beieinander liegen, dazwischen liegt immer noch eine rationale Zahl.
- ii)* Andererseits ist die Menge der rationalen Zahlen verschwindend klein gegenüber der Menge der irrationalen Zahlen.

Kapitel 3

Reelle Funktionen, Polynominterpolation

3.1 Reelle Funktionen (Grundbegriffe; Polynome; rationale Funktionen)

In diesem Paragraphen werden einige Grundbegriffe reeller Funktionen rekapituliert. Dazu sei im Folgenden $A \subset \mathbb{R}$ und $f: A \rightarrow \mathbb{R}$ eine Abbildung von A in die reellen Zahlen.

- i)* f heißt eine **gerade Funktion**, falls mit $x \in A$ auch $-x \in A$ gilt und falls $f(-x) = f(x)$ für alle $x \in A$. (Beispiel: $A = \mathbb{R}$, $f(x) = x^2$.)
- ii)* f heißt eine **ungerade Funktion**, falls mit $x \in A$ auch $-x \in A$ gilt und falls $f(-x) = -f(x)$ für alle $x \in A$. (Beispiel: $A = \mathbb{R}$, $f(x) = x$.)
- iii)* Ein Punkt $x_0 \in A$ heißt **Nullstelle** von f , falls $f(x_0) = 0$.
- iv)* Die Funktion f heißt **nach oben beschränkt**, falls eine Konstante $k \in \mathbb{R}$ existiert mit $f(x) \leq k$ für alle $x \in A$ (k ist eine **obere Schranke**). Die **kleinste obere Schranke** von f heißt das **Supremum** von f , $\sup_{x \in A} f(x)$. Existiert ein $x_0 \in A$ mit

$$f(x_0) = \sup_{x \in A} f(x) ,$$

so heißt x_0 ein **Maximierer** (oder eine **Maximalstelle**) und $f(x_0)$ das **Maximum** von f , $\max_{x \in A} f(x)$.

- v)* Die Funktion f heißt **nach unten beschränkt**, falls eine Konstante $k \in \mathbb{R}$ existiert mit $f(x) \geq k$ für alle $x \in A$ (k ist eine **untere**

Schranke). Die **größte untere Schranke** von f heißt das **Infimum** von f , $\inf_{x \in A} f(x)$. Existiert ein $x_0 \in A$ mit

$$f(x_0) = \inf_{x \in A} f(x) ,$$

so heißt x_0 ein **Minimierer** (oder eine **Minimalstelle**) und $f(x_0)$ das **Minimum** von f , $\min_{x \in A} f(x)$.

vi) Die Funktion f heißt **beschränkt**, falls f nach oben und nach unten beschränkt ist.

vii) Ganz **wesentlich** (und nie zu vergessen!) ist dabei, dass selbst bei “glatten” Funktionen (ohne “Sprünge” oder “Knicke”) das **Supremum oder das Infimum nicht angenommen werden muss**, falls die Funktion auf einem **offenen** Intervall (oder auf ganz \mathbb{R}) definiert ist. Anders ist die Situation im Fall eines **beschränkten und abgeschlossenen** Intervalls (vgl. Abbildung 3.1).

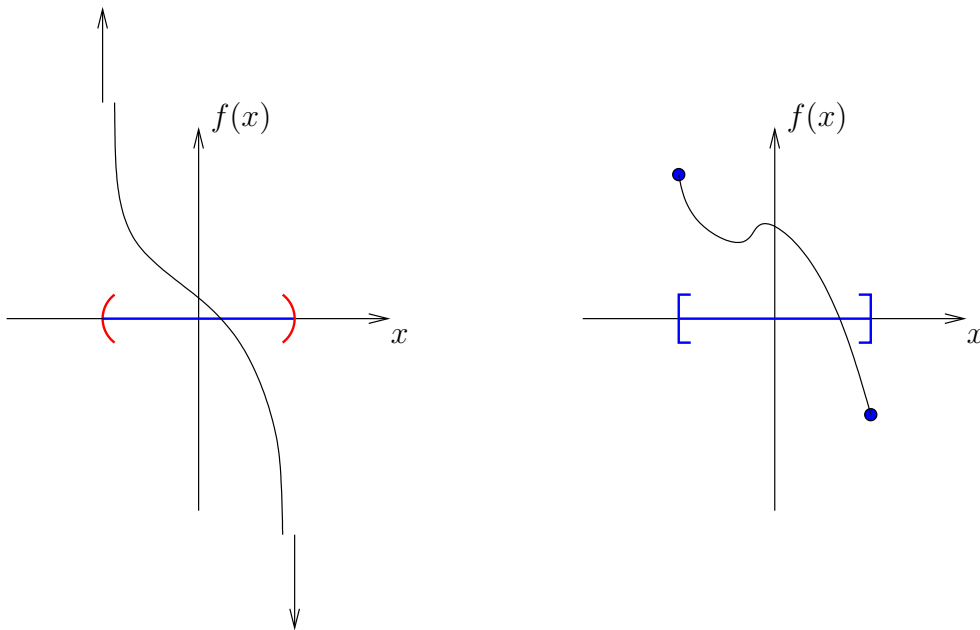


Abbildung 3.1: Zum möglichen Verhalten von Funktionen über offenen und abgeschlossenen Intervallen.

viii) Eine reelle Funktion p_n (mit festen **Koeffizienten** $a_i \in \mathbb{R}$, $i = 0, 1, 2, \dots, n$ ($a_n \neq 0$)),

$$p_n(x) := a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \dots + a_1 \cdot x + a_0 ,$$

heißt **Polynom vom Grad n** , $n \in \mathbb{N}_0$.

ix) Es seien p_m und q_n zwei Polynome vom Grad m bzw. n . Es sei weiter $A := \{x \in \mathbb{R} : q_n(x) \neq 0\}$. Dann heißt die Funktion $r(x) : A \rightarrow \mathbb{R}$,

$$r(x) := \frac{p_m(x)}{q_n(x)},$$

eine **rationale Funktion**.

x) Nullstellen von Polynomen, das **Horner-Schema** zur Auswertung von Polynomen etc. wurden bereits in den Übungen diskutiert. Es sei hier lediglich an Folgendes erinnert: Es ist

$$\begin{aligned} p_n(x) &= a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \dots + a_1 \cdot x + a_0 \\ &= (x - x_0) \cdot [b_{n-1} \cdot x^{n-1} + \dots + b_1 \cdot x + b_0] + B_0. \end{aligned}$$

Die Koeffizienten b_0, b_1, \dots, b_{n-1} und B_0 ergeben sich dabei durch Ausmultiplizieren und **Koeffizientenvergleich**, man sieht, dass B_0 gerade der Funktionswert $p_n(x_0)$ ist.

Ein Polynom vom Grad n ($p_n(x) \not\equiv 0$) hat also maximal n Nullstellen. Hat es tatsächlich n Nullstellen x_1, \dots, x_n (nicht notwendig verschieden), so lässt es sich in **Linearfaktoren** zerlegen:

$$p_n(x) = a_n \cdot (x - x_1) \cdot (x - x_2) \dots (x - x_n).$$

3.2 Polynominterpolation (Darstellung nach Lagrange und Newton; Algorithmus von Neville; Splines)

Beispiele.

i) **Auswertung von Messreihen:** Eine Größe y werde in Abhängigkeit von x an gewissen Messpunkten ermittelt, es entsteht eine Messreihe von geordneten Paaren (x_i, y_i) .

Interpolationsproblem: Gesucht ist eine möglichst einfache Kurve $y = f(x)$, die dieser unbekanntes "Gesetzmäßigkeit" annähernd gerecht wird. Man kann beispielsweise versuchen, ein geeignetes Polynom durch die Messpunkte zu legen.

ii) Eine gegebene Funktion sei (etwa auf einer Rechenanlage) als Tabelle mit endlich vielen Funktionsauswertungen abgelegt. Wie

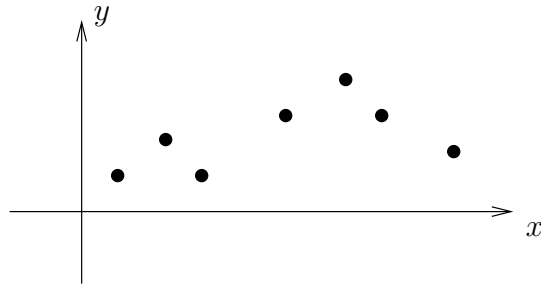


Abbildung 3.2: Eine Messreihe.

kann die Funktion an beliebigen Stellen (möglichst gut) approximiert werden?

Interpolationsaufgabe von Lagrange:

Gegeben seien zu $n \in \mathbb{N}$

$(n + 1)$ paarweise verschiedene **Stützstellen** x_0, x_1, \dots, x_n ,
 $(n + 1)$ **Werte** y_0, y_1, \dots, y_n .

Gesucht ist ein **Polynom p_n von möglichst kleinem Grad** mit der Eigenschaft

$$p_n(x_i) = y_i \quad \text{für alle } i = 0, 1, \dots, n. \quad (*)$$

Bemerkung. Man möchte beispielsweise Messpunkte durch ein Polynom minimalen Grades verbinden.

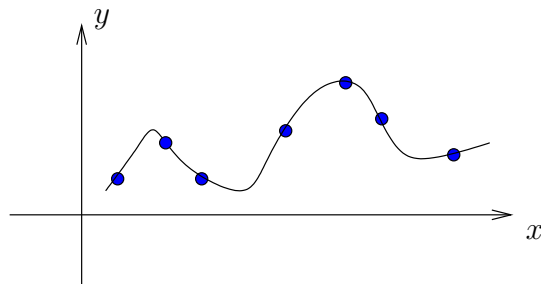


Abbildung 3.3: Zur Interpolation durch Polynome.

Satz 3.2.1¹

Es gibt **genau ein** Polynom vom Grad kleiner oder gleich n , das den Gleichungen (*) genügt. Es lautet

$$p_n(x) = \sum_{j=0}^n y_j \cdot L_j(x) \quad \text{mit} \quad L_j(x) := \prod_{k=0, k \neq j}^n \frac{x - x_k}{x_j - x_k}.$$

Diese Darstellung heißt die **Lagrangesche Darstellung** des Interpolationspolynome p_n .

Beweis. Siehe Übungen. □

Beispiel. Die Daten der Interpolationsaufgabe seien wie in Tabelle 3.1 gegeben.

| | | | | |
|-------|----|----|----|---|
| j | 0 | 1 | 2 | 3 |
| x_j | -2 | -1 | 1 | 2 |
| y_j | 1 | 3 | -3 | 2 |

Tabelle 3.1: Daten einer Interpolationsaufgabe.

Die Aufgabe laute: Man berechne den Wert $p_3(0)$ des Interpolationspolynoms $p_3(x)$.

Die Lösung der Aufgabe ergibt sich aus Tabelle 3.2:

$$p_3(0) = 1 \cdot \left(-\frac{1}{6}\right) + 3 \cdot \frac{2}{3} + (-3) \cdot \frac{2}{3} + 2 \cdot \left(-\frac{1}{6}\right) = -\frac{1}{2}.$$

Bemerkungen.

- i) Der **Nachteil** dieser Methode ist: Die Hinzunahme eines weiteren Messpaares (x_{n+1}, y_{n+1}) erfordert eine **komplette Neuberechnung**.
- ii) Das Interpolationspolynom ist zwar eindeutig bestimmt, es gibt aber **verschiedene Darstellungen** (Schreibweisen) dieses Polynoms. Mit anderen Worten stellt sich die Frage:

¹J.L. Lagrange, 1736-1813; Turin, Paris.

| j | $L_j(x)$ | $L_j(0)$ |
|-----|--|----------------|
| 0 | $\frac{(x+1)(x-1)(x-2)}{(-1) \cdot (-3) \cdot (-4)}$ | $-\frac{1}{6}$ |
| 1 | $\frac{(x+2)(x-1)(x-2)}{1 \cdot (-2) \cdot (-3)}$ | $\frac{2}{3}$ |
| 2 | $\frac{(x+2)(x+1)(x-2)}{3 \cdot 2 \cdot (-1)}$ | $\frac{2}{3}$ |
| 3 | $\frac{(x+2)(x+1)(x-1)}{4 \cdot 3 \cdot 1}$ | $-\frac{1}{6}$ |

Tabelle 3.2: Schema zur Lösung der Interpolationsaufgabe nach Lagrange.

Kann das Interpolationspolynom “geschickter” berechnet werden?

Es werden zunächst **Newtonschen² Basisfunktionen** eingeführt ($n \in \mathbb{N}$ fixiert, Stützstellen und Werte wie oben):

$$N_0^{(n)}(x) := 1, \quad N_j^{(n)}(x) := \prod_{k=0}^{j-1} (x - x_k), \quad j = 1, 2, \dots, n.$$

Das Interpolationspolynom der Aufgabe von Lagrange soll dann dargestellt werden in der Form:

$$p_n(x) = \sum_{j=0}^n a_j \cdot N_j^{(n)}(x)$$

mit reellen Koeffizienten a_0, a_1, \dots, a_n .

Dazu sei $0 \leq j \leq n$, $0 \leq k \leq n - j$ und $p_{j,j+k}$ sei das eindeutig bestimmte Interpolationspolynom vom **Grad $\leq k$** zu den Stützstellen/Werten $(x_j, y_j), (x_{j+1}, y_{j+1}), \dots, (x_{j+k}, y_{j+k})$.

²I. Newton, 1643-1727; Cambridge

Durch die sukzessive Berechnung dieser Polynome soll schließlich $p_n = p_{0,n}$, die Lösung der ursprünglichen Interpolationsaufgabe, berechnet werden. Dies geschieht mit Hilfe der **dividierten Differenzen**:

$$y[x_j] := y_j, \quad j = 0, 1, \dots, n;$$

$$y[x_j, \dots, x_{j+k}] := \frac{y[x_{j+1}, \dots, x_{j+k}] - y[x_j, \dots, x_{j+k-1}]}{x_{j+k} - x_j},$$

$$j = 0, 1, \dots, n-1, \quad k = 1, 2, \dots, n-j.$$

Bemerkung. Man beachte, dass die dividierten Differenzen tatsächlich sukzessive berechnet werden können.

Aus einer Induktion nach k und einer Indexverschiebung folgt

Satz 3.2.2 (*Interpolationspolynom in Newtonscher Darstellung*)

Mit obiger Notation gilt:

i) Für $0 \leq j \leq j+k \leq n$ ist

$$p_{j,j+k} = y[x_j] + y[x_j, x_{j+1}] \cdot (x - x_j) + \dots \\ + y[x_j, \dots, x_{j+k}](x - x_j) \dots (x - x_{j+k-1}).$$

ii) Das Interpolationspolynom in **Newtonscher Darstellung** lautet:

$$p_n(x) = \sum_{j=0}^n y[x_0, \dots, x_j] \cdot N_j^{(n)}(x).$$

Ein Schema zur Berechnung der dividierten Differenzen in der Newtonschen Darstellung ist in Tabelle 3.3 angegeben. Hierbei ergibt sich die Spalte mit der Nummer $l+1$ wie folgt: Zwei übereinanderstehende Eintragungen aus Spalte l werden subtrahiert (Vorzeichen beachten), das Ergebnis wird durch die entsprechende Argumentdifferenz dividiert. Die Koeffizienten der Newtonschen Darstellung stehen schließlich oben in jeder Spalte.

| | | | | | |
|-----------|------------------------|-----------------------|--------------------------------|---------|---------------------------|
| x_0 | $y[x_0] = y_0$ | | | | |
| | | $y[x_0, x_1]$ | | | |
| x_1 | $y[x_1] = y_1$ | | $y[x_0, x_1, x_2]$ | | |
| | | $y[x_1, x_2]$ | | | |
| x_2 | $y[x_2] = y_2$ | | $y[x_1, x_2, x_3]$ | | |
| \vdots | \vdots | \vdots | \vdots | \dots | $y[x_0, x_1, \dots, x_n]$ |
| x_{n-2} | $y[x_{n-2}] = y_{n-2}$ | | $y[x_{n-3}, x_{n-2}, x_{n-1}]$ | | |
| | | $y[x_{n-2}, x_{n-1}]$ | | | |
| x_{n-1} | $y[x_{n-1}] = y_{n-1}$ | | $y[x_{n-2}, x_{n-1}, x_n]$ | | |
| | | $y[x_{n-1}, x_n]$ | | | |
| x_n | $y[x_n] = y_n$ | | | | |

Tabelle 3.3: Schema zur Berechnung der dividierten Differenzen in der Newtonschen Darstellung.

Man beachte, dass **bei Hinzufügen eines neuen Interpolationspunktes nur eine weitere Schrägzeile berechnet werden muss.**

Beispiel. Man betrachte wieder die Daten aus Tabelle 3.1 und berechne $p_3(0)$ über die Newtonsche Darstellung.

Im konkreten Beispiel ergänzt Tabelle 3.4 die Tabelle 3.3 nach links, um die Argumentdifferenzen vor Augen zu haben: In der ersten Spalte steht der Wert $x_3 - x_0$, in der zweiten findet man die Werte $x_{j+2} - x_j$, $j = 0, 1$, in der dritten $x_{j+1} - x_j$, $j = 0, 1, 2$.

Aus dem Schema ergibt sich

$$p_3(x) = 1 + 2 \cdot (x + 2) - \frac{5}{3} \cdot (x + 2) \cdot (x + 1) + \frac{13}{12} \cdot (x + 2) \cdot (x + 1) \cdot (x - 1),$$

also insbesondere wieder

$$p_3(0) = -\frac{1}{2}.$$

| | | | | | | |
|---|---|------------|---------------|--------------------|-----------------------------------|---------------------------------|
| | | $x_0 = -2$ | $y[x_0] = 1$ | | | |
| | 1 | | | $y[x_0, x_1] = 2$ | | |
| | 3 | $x_1 = -1$ | $y[x_1] = 3$ | | $y[x_0, x_1, x_2] = -\frac{5}{3}$ | |
| 4 | 2 | | | $y[x_1, x_2] = -3$ | | $y[x_0, \dots] = \frac{13}{12}$ |
| | 3 | $x_2 = 1$ | $y[x_2] = -3$ | | $y[x_1, x_2, x_3] = \frac{8}{3}$ | |
| | 1 | | | $y[x_2, x_3] = 5$ | | |
| | | $x_3 = 2$ | $y[x_3] = 2$ | | | |

Tabelle 3.4: Schema zur Lösung der Interpolationsaufgabe nach Newton.

Will man (wie in obigen Beispielen) nur die **Auswertung des Interpolationspolynoms an einer festen Stelle** (und nicht das Polynom selbst) wissen, so ist der **Algorithmus von Neville** eine geeignete Methode. Anhand des Induktionsbeweises von Satz 3.2.2 wird deutlich: Mit der Notation (für fixiertes x , Interpolationsaufgabe wie oben)

$$p_{j,j}(x) := y_j, \quad j = 0, 1, \dots, n,$$

$$p_{j,j+k}(x) := p_{j,j+k-1}(x) + (x - x_j) \cdot \left[\frac{p_{j+1,j+k}(x) - p_{j,j+k-1}(x)}{x_{j+k} - x_j} \right],$$

$$0 \leq j < j+k \leq n,$$

ist $p_n(x) = p_{0,n}(x)$ der gesuchte Wert.

Bemerkung. Wieder wird bei der Berechnung sukzessive vorgegangen.

| x_j | y_j | $k = 1$ | $k = 2$ | $k = 3$ |
|------------|-------------------------|-------------------|-----------------------------|-----------------------------|
| $x_0 = -2$ | $y_0 = p_{0,0}(0) = 1$ | $p_{0,1}(0) = 5$ | $p_{0,2}(0) = \frac{5}{3}$ | $p_{0,3}(0) = -\frac{1}{2}$ |
| $x_1 = -1$ | $y_1 = p_{1,1}(0) = 3$ | $p_{1,2}(0) = 0$ | $p_{1,3}(0) = -\frac{8}{3}$ | |
| $x_2 = 1$ | $y_2 = p_{2,2}(0) = -3$ | $p_{2,3}(0) = -8$ | | |
| $x_3 = 2$ | $y_3 = p_{3,3}(0) = 2$ | | | |

Tabelle 3.5: Schema zum Algorithmus von Neville.

Wie man den Algorithmus von Neville schematisch angeht, ist in Tabelle 3.5 anhand des obigen Beispiels dargestellt (Daten aus Tabelle 3.1, gesucht $p_3(0)$).

Es wird spaltenweise vorgegangen:

$$k = 1 : \quad p_{0,1}(0) = 1 + (0 + 2) \cdot \frac{3 - 1}{-1 - (-2)} = 5 ;$$

$$p_{1,2}(0) = 3 + (0 - (-1)) \cdot \frac{-3 - 3}{1 - (-1)} = 0 ;$$

$$p_{2,3}(0) = -3 + (0 - 1) \cdot \frac{2 - (-3)}{2 - 1} = -8 ;$$

$$k = 2 : \quad p_{0,2}(0) = 5 + (0 - (-2)) \cdot \frac{0 - 5}{1 - (-2)} = \frac{5}{3} ;$$

$$p_{1,3}(0) = 0 + (0 - (-1)) \cdot \frac{-8 - 0}{2 - (-1)} = -\frac{8}{3} ;$$

$$k = 3 : \quad p_{0,3}(0) = \frac{5}{3} + (0 - (-2)) \cdot \frac{-\frac{8}{3} - \frac{5}{3}}{2 - (-2)} = -\frac{1}{2} .$$

Bemerkung. Wieder ist beim Hinzufügen eines neuen Messpunktes nur eine Schrägzeile zu ergänzen.

Bemerkung. Ein wesentlicher Vorteil der Polynominterpolation ist die geschlossene Form, in der die interpolierende Funktion (in diesem Fall das Polynom) angegeben werden kann. Liegen jedoch sehr viele Datenpaare (10 und mehr) vor, so wird das interpolierende Polynom sehr **stark oszillieren**, wie in Abbildung 3.4 angedeutet ist.

Einen Ausweg im Fall vieler Datenpaare bieten so genannte **Splines**. Sind x_0, x_1, \dots, x_n wieder die Stützstellen des Problems (auch **Knoten** genannt), jetzt mit $x_0 < x_1 < \dots < x_n$, so konstruiert man auf $[x_0, x_1]$ die **zu den Knoten gehörende Spline-Funktion vom Grade $k \geq 1$** , d.h. es wird eine Funktion $s_k(x)$ konstruiert mit den Eigenschaften:

- i) $s_k(x)$ ist eine “glatte” Funktion (keine “Sprünge” oder “Knickstellen”, genauer gesagt: s_k ist $(k - 1)$ mal stetig differenzierbar);

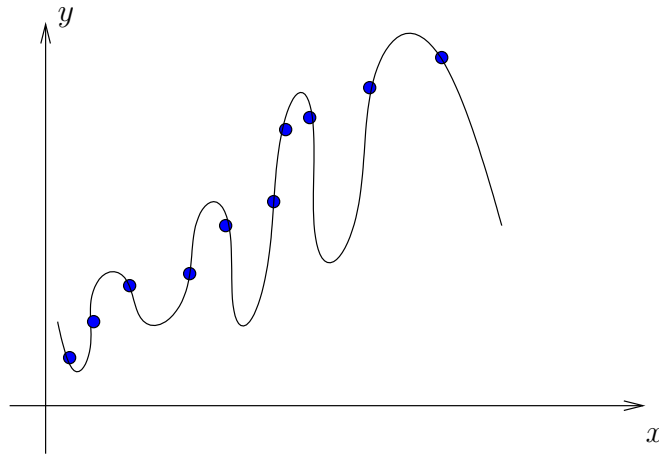


Abbildung 3.4: Oszillationen bei vielen Datenpaaren.

- ii) $s_k(x)$ ist auf allen Teilintervallen $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n - 1$, ein Polynom höchstens k -ten Grades;
- iii) s_k ist **interpolierende Spline-Funktion** zu gegebenen Werten y_0, y_1, \dots, y_n , mit anderen Worten:

$$s_k(x_i) = y_i \quad \text{für alle } i = 0, 1, \dots, n .$$

Dabei wird in der Praxis meist mit Spline-Funktionen niedrigen Grades gearbeitet, etwa $k \leq 3$.

In den Abbildungen 3.5 und 3.6 sind zu gleichen Daten bei 10 Stützstellen die Lösung der Interpolationsaufgabe nach Lagrange und die Spline-Interpolation dargestellt. Die Auswertung ist mit Hilfe von “Maple” erfolgt.

Man erkennt deutlich die starken Oszillationen bei der Lagrangeschen Interpolationsaufgabe (bitte die Skalierung beachten), sobald die Anzahl der Datenpaare zu groß wird.

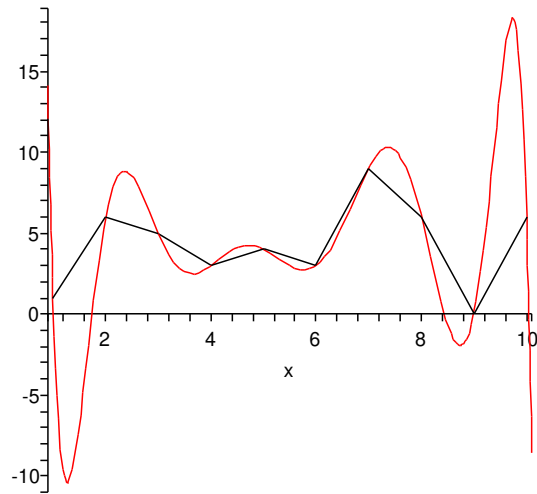


Abbildung 3.5: Zur Interpolationsaufgabe nach Lagrange.

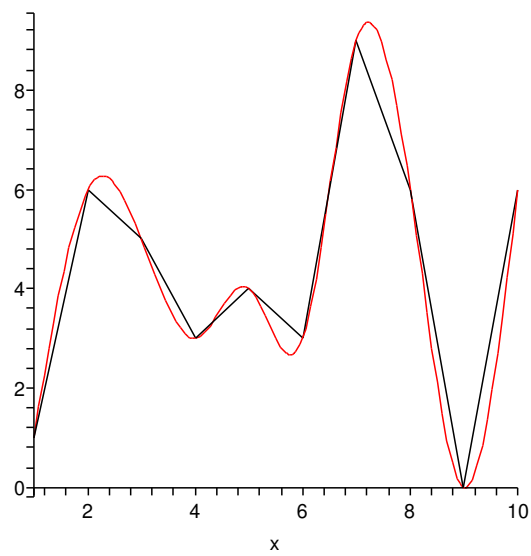


Abbildung 3.6: Zur Spline-Interpolation.

Kapitel 4

Folgen, Reihen, Maschinenzahlen

Beispiel 1. (*Zifferndarstellung*) Intuitiv wird eine Zifferndarstellung für Zahlen benutzt, beispielsweise¹

$$\frac{1}{3} = 0.\bar{3}, \quad \pi = 3.141592653589\dots$$

Was soll diese Schreibweise jedoch **präzise bedeuten**, wenn die Anzahl der Ziffern **nicht endlich** ist?

Beispiel 2. (vgl. “Achill und die Schildkröte”) A bewege sich mit einer Geschwindigkeit von $2m/s$. In einer Entfernung von $10m$ vor sich sieht er B , der sich mit $1m/s$ nach vorne bewegt. Bezeichnet t_0 die Zeit des Überholvorgangs und ist s_0 die bis dahin von A zurückgelegte Wegstrecke, so gilt (Weg = Geschwindigkeit · Zeit)

$$t_0 = \frac{s_0}{2m/s} = \frac{s_0 - 10m}{1m/s}.$$

A überholt B entsprechend nach $20m$.

Aber: Wenn A die $10m$ zurückgelegt hat, die B Vorsprung hatte, dann ist B um $5m$ weitergekommen und hat nach wie vor einen Vorsprung. Hat A diesen Vorsprung aufgeholt, dann ist B aber um $2.5m$ weitergekommen, ein Vorsprung den A erst wieder aufholen muss ... usw. usw. So kann A niemals B einholen?

Ist dieser scheinbare Widerspruch mathematisch aufzulösen?

¹ $0.\bar{3}$ bedeutet: Mit Periode 3.

4.1 Reelle Zahlenfolgen (Konvergenz; Divergenz; Konvergenzkriterien; Teilfolgen; Satz von Bolzano-Weierstraß)

Definition 4.1.1 ²

Es sei $A \neq \emptyset$ eine Menge von Elementen a, b, c, \dots . Eine **Folge** $\{a_n\}$ in A ist eine Abbildung $\mathbb{N} \rightarrow A$ (oft auch $\mathbb{N}_0 \rightarrow A$), die jedem $n \in \mathbb{N}$ ein Element $a_n \in A$, das **n^{te} Glied der Folge $\{a_n\}$** zuordnet. Mann schreibt für die Folge auch häufig a_1, a_2, a_3, \dots . Eine **reelle Zahlenfolge** ist eine Abbildung $\mathbb{N} \rightarrow \mathbb{R}$.

Bemerkung. Die **Glieder** der Folge sind **Elemente aus A** , die **Folge selbst** ist eine **Abbildung** wie unten angedeutet.

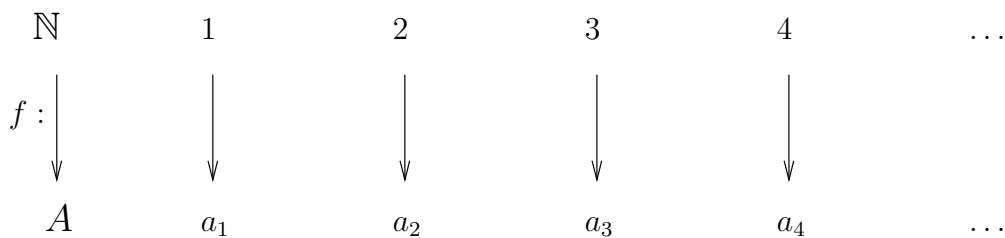


Abbildung 4.1: Eine Folge ist eine Abbildung.

In diesem Paragraphen geht es von nun an **ausschließlich um reelle Zahlenfolgen**, auch wenn dies nicht immer explizit erwähnt ist.

Beispiele.

- i) Ist $a_n := a \in \mathbb{R}$ für alle $n \in \mathbb{N}$, so erhält man die konstante Folge a, a, a, \dots
- ii) Ist $a_n := (-1)^{n+1}$, $n \in \mathbb{N}$, so entsteht die **alternierende** Folge $1, -1, 1, -1, \dots$
- iii) Ist a_n für alle n gegeben durch $1/n$, so handelt es sich um die Folge: $1, 1/2, 1/3, 1/4, \dots$

²Schreibweise: $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$

iv) Eine Folge kann auch **rekursiv** definiert sein, etwa die Folge der **Fibonaccizahlen**³.

Rekursionsanfang: $a_0 := 0$, $a_1 := 1$.

Rekursionsvorschrift: $a_n := a_{n-1} + a_{n-2}$ für $n \geq 2$.

Die Folge lautet: 0, 1, 1, 2, 3, 5, 8, ...

(Ein Kaninchenpaar befinde sich in einem geschlossenen Gehege. Man überlege sich, wie viele Kaninchenpaare es nach einem Jahr gibt, wenn jedes Paar jeden Monat zwei Junge zeugt, welche sich ihrerseits vom zweiten Monat an vermehren.)

Für “kleine” n können die Folgenglieder explizit aufgelistet werden. Der wesentliche Unterschied zwischen einer Folge und einem geordneten k -Tupel ist aber, dass es im Falle einer Folge “**unendlich viele geordnete Eintragungen**” gibt.

Die entscheidende Frage ist:

Wie kann man das Verhalten einer Folge für beliebig große n charakterisieren? Mit anderen Worten: Welches **Konvergenzverhalten** hat eine Folge?

Zur Veranschaulichung einer **konvergenten** Folge sei $\{1/n\}$ betrachtet. Die Vorstellung, dass die Folgenglieder sich immer mehr der 0 annähern, ist in Abbildung 4.2 illustriert.

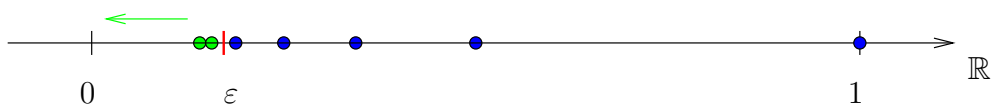


Abbildung 4.2: Bei diesem ε sind ab $n = 6$ alle Folgenglieder näher bei der 0 als ε .

Dass die a_n der 0 beliebig nahe kommen bedeutet: Wenn man sich auf dem Zahlenstrahl in **irgendeinem festen Abstand** $\varepsilon > 0$ befindet, dann sind irgendwann die a_n , die in dem Beispiel auf die 0 zulaufen, näher bei der 0 als man selbst (vgl. nochmals Abbildung 4.2).

Präzisiert wird das in

³L. Fibonacci, ca. 1180-1250; Pisa.

Definition 4.1.2

Eine Folge $\{a_n\}$ heißt **konvergent**, wenn es ein $a \in \mathbb{R}$ gibt, sodass gilt: Für jedes $\varepsilon > 0$ existiert eine (von ε abhängende) Zahl $N(\varepsilon) \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

Falls eine Folge konvergiert, so heißt a der **Grenzwert** der Folge (die Folge **konvergiert gegen** a ; die Folge hat den **Grenzwert** oder **Limes** a). Notation:

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a \quad \text{für } n \rightarrow \infty.$$

Eine Folge, die nicht konvergiert, heißt **divergent**.

Bemerkungen und Beispiele.

- i) Der Konvergenzbegriff gehört zu den **Grundbausteinen der Analysis**.
- ii) Betrachtet sei die Folge $\{1/n\}$ und $a = 0$. Zu gegebenem $\varepsilon > 0$ wählt man (irgend-) eine natürliche Zahl $N(\varepsilon)$ mit $N(\varepsilon) > \varepsilon^{-1}$ (Satz des Archimedes!). Für $n \geq N(\varepsilon)$ folgt

$$|1/n - 0| = \frac{1}{n} \leq \frac{1}{N(\varepsilon)} < \frac{1}{\varepsilon^{-1}} = \varepsilon,$$

die Konvergenz gegen 0 ist damit auch formal bestätigt. Man sagt auch, die Folge ist eine **Nullfolge**.

- iii) Divergenz einer Folge bedeutet: Für alle $a \in \mathbb{R}$ existiert ein $\varepsilon > 0$, sodass für alle $N \in \mathbb{N}$ ein $n \geq N$ existiert mit $|a_n - a| \geq \varepsilon$ (man mache sich klar, dass dies genau die **Negation** des Konvergenzbegriffes ist: Unter Beibehaltung der Reihenfolge ist “ \forall ” zu ersetzen durch “ \exists ” und vice versa, die übrigen Aussagen sind zu negieren).
- iv) Betrachtet sei die Folge $\{n\}$. Dann existiert für alle $a \in \mathbb{R}$ ein ε (man kann z.B. $\varepsilon = 1$ wählen), sodass für alle $N \in \mathbb{N}$ ein $n \geq N$

existiert (man wähle etwa $n = \max\{N, a + 1\}$ ⁴) mit

$$|a_n - a| = |n - a| \geq 1 = \varepsilon .$$

Die Folge ist also divergent. Auf die gleiche Art zeigt man

Satz 4.1.1

Jede *konvergente* reelle Zahlenfolge $\{a_n\}$ ist *beschränkt*. Dabei heißt die Folge beschränkt, wenn es ein $k \in \mathbb{R}$ gibt mit $|a_n| \leq k$ für alle n (nach oben beschränkt: $a_n \leq k$; nach unten beschränkt: $k \leq a_n$).

Weitere Bemerkungen.

i) Am Beispiel $1, -1, 1, -1, \dots$ sieht man, dass umgekehrt eine **beschränkte Folge nicht konvergent sein muss**.

ii) Kann für $n \rightarrow \infty$ sowohl $a_n \rightarrow a$ als auch $a_n \rightarrow b$ mit $a \neq b$ gelten?

Dann **gäbe** es zu jedem $\varepsilon > 0$ Zahlen $N_a(\varepsilon)$ und $N_b(\varepsilon) \in \mathbb{N}$ mit

$$\begin{aligned} |a_n - a| < \varepsilon & \quad \text{für alle } n \geq N_a(\varepsilon) , \\ |a_n - b| < \varepsilon & \quad \text{für alle } n \geq N_b(\varepsilon) . \end{aligned}$$

Ist $n > \max\{N_a(\varepsilon), N_b(\varepsilon)\}$, so gilt (Dreiecksungleichung nie vergessen!)

$$|a - b| = |a - a_n + a_n - b| \leq |a_n - a| + |a_n - b| \leq 2\varepsilon .$$

Es sind a und b reelle Zahlen mit einem festen Abstand $|a - b|$. Wählt man nun ε so klein, dass $2\varepsilon < |a - b|$ (z.B. $\varepsilon = |a - b|/3$), so erhält man einen Widerspruch, also ist zwingend $a = b$.

Mit anderen Worten: **Der** Grenzwert einer konvergenten Folge ist **eindeutig bestimmt**.

iii) Das Konvergenzverhalten einer Folge wird durch Ändern, Hinzufügen oder Weglassen einer **endlichen Anzahl** von Folgengliedern nicht geändert. Insbesondere bleibt im Falle der Konvergenz der Grenzwert unverändert.

⁴Das Maximum zweier Zahlen ist definiert als: $\max\{s, t\} := s$ falls $s \geq t$, $\max\{s, t\} := t$ falls $s < t$; Minimum analog.

iv) Für ein fixiertes $x \in \mathbb{R}$ sei eine **geometrische Folge** betrachtet: $a_n := x \cdot a_{n-1}$ für $n > 1$. Mit der Wahl $a_1 := x$ erhält man $a_n = x^n$. Ist $x = 0$ oder $x = 1$, so handelt es sich um eine konstante Folge, die Folge ist also konvergent.

Ist $x = -1$, so divergiert die Folge.

Ist $0 < |x| < 1$, so ist $|x|^{-1} > 1$, also gilt $h := |x|^{-1} - 1 > 0$ und $|x| = (1+h)^{-1}$. Die Bernoullische⁵ Ungleichung $((1+h)^n \geq 1+nh$ für $h > -1$, siehe Übungen) liefert zu $\varepsilon > 0$ und $N(\varepsilon) = 1/(h\varepsilon)$ für alle $n \geq N(\varepsilon)$

$$|a_n - 0| = |x^n| = (1+h)^{-n} \leq \frac{1}{1+nh} < \frac{1}{nh} = \frac{1}{h} \cdot \frac{1}{n} < \varepsilon,$$

die Folge konvergiert also gegen 0.

Im Fall $|x| > 1$ kann die Folge nach Satz 4.1.1 nicht konvergieren.

Wegen der besonderen Bedeutung dieser Folge sei das Ergebnis in einem Satz zusammengefasst:

Satz 4.1.2

Ist $x \in \mathbb{R}$ fixiert, so gilt für die Folge $\{x^n\}$:

- i) Für $|x| < 1$ ist die Folge **konvergent** mit $\lim_{n \rightarrow \infty} x^n = 0$.*
- ii) Für $x = 1$ ist die Folge **konvergent** mit $\lim_{n \rightarrow \infty} x^n = 1$.*
- iii) Für $|x| > 1$ und $x = -1$ ist die Folge **divergent**.*

Ein nächstes bekanntes Beispiel besagt, dass **die Fakultät $n!$ stärker wächst als x^n für fixiertes $x \in \mathbb{R}$** . Man betrachte dazu die Folgen $\{x^n\}$ und $\{n!\}$. Wächst n , so kommt in beiden Fällen ein Faktor hinzu: Im Falle $\{x^n\}$ pro Schritt der Faktor x , im Fall $\{n!\}$ werden die neuen Faktoren dagegen immer größer.

Mit dieser Idee kann relativ leicht gezeigt werden:

⁵Jacob Bernoulli, 1654-1705; Basel.

Satz 4.1.3

Es sei $x \in \mathbb{R}$ fest. Die für alle $n \in \mathbb{N}$ durch

$$a_n = \frac{x^n}{n!}$$

definierte Folge ist eine Nullfolge.

Rechenregeln für Folgen?

Satz 4.1.4

Es seien $\{a_n\}$ und $\{b_n\}$ zwei **konvergente** Folgen mit den Grenzwerten a bzw. $b \in \mathbb{R}$. Dann gilt:

- i) Ist $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist $a \leq b$.
- ii) Die Folgen $\{|a_n|\}$, $\{c \cdot a_n\}$ für $c \in \mathbb{R}$, $\{a_n \pm b_n\}$ und $\{a_n \cdot b_n\}$ konvergieren mit

$$\begin{aligned} \lim_{n \rightarrow \infty} |a_n| &= |a|, \\ \lim_{n \rightarrow \infty} c \cdot a_n &= c \cdot a, \\ \lim_{n \rightarrow \infty} (a_n \pm b_n) &= a \pm b, \\ \lim_{n \rightarrow \infty} a_n \cdot b_n &= a \cdot b. \end{aligned}$$

Ist $b \neq 0$ und gilt $b_n \neq 0$ für alle $n \in \mathbb{N}$, so konvergiert auch die Folge $\{a_n/b_n\}$ mit

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

- iii) (**Einschließungskriterium**) Ist $a = b$ und gilt $a_n \leq c_n \leq b_n$, wobei $\{c_n\}$ eine weitere Folge bezeichne, so ist auch $\{c_n\}$ konvergent mit $\lim_{n \rightarrow \infty} c_n = a$.

Beweis. Als ein typisches Beispiel sei hier die Konvergenz der Produktfolge bewiesen: Die Folgen $\{a_n\}$ und $\{b_n\}$ sind nach Voraussetzung konvergent und nach Satz 4.1.1 insbesondere beschränkt. Es existiert damit ein $K \in \mathbb{R}$ mit $|a_n| \leq K$, $|b_n| \leq K$, $|a| \leq K$, $|b| \leq K$. (Warum gilt das automatisch auch für die Grenzwerte a , b ?)

Zu gegebenem $\varepsilon > 0$ sei jetzt $N(\varepsilon)$ **hinreichend groß** (wird gleich präzisiert) und $n \geq N(\varepsilon)$. Dann gilt (Dreiecksungleichung)

$$\begin{aligned} |a_n b_n - ab| &\leq |a_n b_n - ab_n + ab_n - ab| \\ &\leq |a_n - a| |b_n| + |a| |b_n - b|. \end{aligned}$$

Wegen der Konvergenz der Folgen $\{a_n\}$, $\{b_n\}$ existiert aber ein $\tilde{N}(\varepsilon)$, sodass $|a_n - a| < \varepsilon/2K$ und $|b_n - b| < \varepsilon/2K$ für $n \geq \tilde{N}(\varepsilon)$. Wählt man nun $N(\varepsilon) > \tilde{N}(\varepsilon)$, so ist für alle $n \geq N(\varepsilon)$ gezeigt:

$$|a_n b_n - ab| < \frac{\varepsilon}{2K} K + \frac{\varepsilon}{2K} K = \varepsilon,$$

die Konvergenz der Folge ist bewiesen. □

Bemerkungen und Beispiele.

i) Diese Regeln dürfen **nur angewandt werden, wenn die Konvergenz der Folgen $\{a_n\}$ und $\{b_n\}$ überprüft ist**. Zudem folgt z.B. aus der Konvergenz von $\{a_n + b_n\}$ nicht die Konvergenz von $\{a_n\}$ und $\{b_n\}$. Beispiel. Ist für alle $n \in \mathbb{N}$ $a_n := n + 1/n$, $b_n := -n + 1/n$, so ist $\{a_n + b_n\} = \{2/n\}$ konvergent, die Folgen $\{a_n\}$ und $\{b_n\}$ sind aber divergent.

ii) Man betrachte die Folge $\{c_n\}$ mit

$$c_n = \frac{2n^2 - 3n + 1}{3n^2 + 15n + 4} \quad \text{für alle } n \in \mathbb{N},$$

bzw. umgeformt

$$c_n = \frac{2 - \frac{3}{n} + \frac{1}{n^2}}{3 + \frac{15}{n} + \frac{4}{n^2}} =: \frac{a_n}{b_n}.$$

Hierbei ist $\{a_n\}$ eine Summe aus drei konvergenten Folgen (warum konvergiert $1/n^2$?) mit den Grenzwerten 2, 0 und 0, d.h. $\{a_n\}$ konvergiert gegen $a = 2$ für $n \rightarrow \infty$.

$\{b_n\}$ ist eine Summe aus drei konvergenten Folgen mit den Grenzwerten 3, 0 und 0, $\{b_n\}$ konvergiert gegen $b = 3$ für $n \rightarrow \infty$.

Weiterhin gilt $b_n \neq 0$ für alle $n \in \mathbb{N}$, $b \neq 0$, aus Satz 4.1.4 folgt

$$c_n \xrightarrow{n \rightarrow \infty} \frac{2}{3}.$$

Konvergenzkriterien?

Neben dem Einschließungskriterium seien hier noch der [Satz von der monotonen Folge](#) und das [Konvergenzkriterium von Cauchy](#)⁶ diskutiert.

Dazu wird zunächst der Begriff Monotonie definiert.

Definition 4.1.3

Eine reelle Zahlenfolge $\{a_n\}$ heißt

- i) **monoton wachsend**, falls $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$;*
- ii) **monoton fallend**, falls $a_n \geq a_{n+1}$ für alle $n \in \mathbb{N}$.*

*Gelten die strikten Ungleichungen (d.h. mit “<” und “>”), so spricht man von einer **streng monoton wachsenden** bzw. **streng monoton fallenden Folge**.*

Satz 4.1.5 (*monotone Konvergenz*)

*Eine **monoton wachsende** reelle Zahlenfolge ist **genau dann** konvergent, wenn sie nach **oben beschränkt** ist.*

*Eine **monoton fallende** reelle Zahlenfolge ist **genau dann** konvergent, wenn sie nach **unten beschränkt** ist.*

Beweisidee. Existenz einer kleinsten oberen Schranke; Monotonie ausnutzen. □

⁶A.L. Cauchy, 1789-1857; Paris.

Beispiele.

i) Die Folge $\{n^{-1}\}$ ist monoton fallend und nach unten beschränkt, also konvergent (wie bereits bekannt).

ii) Es sei $\{a_n\}$ die monoton wachsende Folge mit

$$a_n = \sum_{j=0}^n \frac{1}{j!} \quad \text{für alle } n \in \mathbb{N}.$$

Satz 4.1.3 besagt für große j : $1/j! < 2^{-j}$. Da zudem $\sum_{j=0}^n 2^{-j} < 2$ für alle $n \in \mathbb{N}$, ist $\{a_n\}$ nach oben beschränkt und konvergiert nach Satz 4.1.5. Der Grenzwert wird nicht mitgeliefert.

iii) Die rekursiv definierte Folge $\{a_n\}$ mit

$$a_1 = 3, \quad a_{n+1} = \frac{1}{5}a_n^2 + \frac{4}{5} \quad \text{für alle } n \in \mathbb{N}$$

ist beschränkt (als Übung ist zu zeigen: $1 \leq a_n \leq 3$ für alle n) und monoton fallend. Dazu zeige man, dass für alle n folgt

$$\frac{a_{n+1}}{a_n} = \frac{\frac{1}{5}a_n^2 + \frac{4}{5}}{a_n} < 1.$$

Demnach ist die Folge konvergent. Um den Grenzwert zu bestimmen, hilft ein Trick: In der rekursiven Darstellung für a_{n+1} geht man zum Grenzwert über, d.h. **wenn** die Folge konvergiert und der Grenzwert mit a bezeichnet ist, dann muss gelten

$$a = \frac{1}{5}a^2 + \frac{4}{5}, \quad \text{also } a^2 - 5a + 4 = 0.$$

Diese quadratische Gleichung hat die Wurzeln 4 und 1, da die Folge aber nach oben durch 3 beschränkt ist, muss der Grenzwert 1 sein.

iv) In einer rekursiven Vorschrift darf **nur zur Grenze übergegangen werden, wenn die Konvergenz der Folge überprüft ist.**

Beispiel.

$$a_1 = 1, \quad a_{n+1} = 1 - a_n \quad \text{für alle } n \in \mathbb{N}.$$

Dürfte man hier zur Grenze übergehen, so **wäre**

$$a = 1 - a, \quad \text{d.h. } a = \frac{1}{2}.$$

Die Folge lautet aber 1, 0, 1, 0, ... und ist keinesfalls konvergent.

- v) Ein Kapital K_0 sei zu einem Jahreszinsfuß p ein Jahr lang verzinst. Nach diesem Jahr beläuft sich das Kapital auf

$$K_1 = K_0(1 + p) \quad \text{bei jährlicher Verzinsung;}$$

$$K_2 = K_0 \left(1 + \frac{p}{2}\right)^2 \quad \text{bei halbjährlicher Verzinsung;}$$

$$K_4 = K_0 \left(1 + \frac{p}{4}\right)^4 \quad \text{bei vierteljährlicher Verzinsung.}$$

Am Beispiel der Verzinsung wird deutlich, dass die Folge

$$a_n := \left(1 + \frac{1}{n}\right)^n \quad \text{für alle } n \in \mathbb{N}$$

monoton wächst (halbjährliche Verzinsung ist günstiger als jährliche etc.) und beschränkt ist (sonst wäre die Bank pleite) – beide Eigenschaften werden mit Hilfe der Bernoullischen Ungleichung bewiesen. Damit ist die Folge konvergent.

Wegen der besonderen Bedeutung wird dies als Satz formuliert:

Satz 4.1.6 ⁷

*Die Folge $\{a_n\}$, $a_n := (1 + 1/n)^n$ für alle $n \in \mathbb{N}$, ist konvergent. Der Grenzwert heißt die **Eulersche Zahl e** .*

Bemerkungen.

- i) Die transzendente Zahl e ist ungefähr $e = 2.71828182845 \dots$
- ii) Die Folge aus obigem Beispiel ii) konvergiert ebenfalls gegen e .

Zum Kriterium von Cauchy:

Definition 4.1.4

*Eine reelle Zahlenfolge $\{a_n\}$ heißt **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt, sodass gilt:*

$$|a_n - a_m| < \varepsilon \quad \text{für alle } n, m \geq N(\varepsilon) .$$

⁷L. Euler, 1707-1783; Basel, Berlin, St. Petersburg.

Satz 4.1.7 (*Cauchys Konvergenzkriterium*)

Eine reelle Zahlenfolge ist *genau dann konvergent*, wenn sie eine *Cauchy-Folge* ist.

Beweisidee. Der Satz folgt aus dem [Satz von Bolzano⁸-Weierstraß⁹](#), siehe Satz 4.1.8. □

Bemerkung. Der Satz gilt auch für nicht-monotone Folgen und wieder muss der Grenzwert nicht bekannt sein, um Konvergenz zu zeigen.

Beispiel. Die Folge $\{a_n\}$ sei rekursiv gegeben durch

$$a_1 = 1, \quad a_2 = \frac{1}{2}, \quad a_{n+2} = \frac{1}{2}(a_{n+1} + a_n) \quad \text{für } n \in \mathbb{N}$$

(a_{n+2} ist das [arithmetische Mittel](#) der beiden Vorgänger). Die Folge ist nicht monoton, trotzdem liefert das Cauchysche Konvergenzkriterium die Konvergenz (vgl. Übungen).

Teilfolgenauswahl?

Die Folge $1, -1, 1, -1, \dots$ divergiert, lässt man aber jedes zweite Glied weg, so erhält man eine konvergente Folge, nämlich die konstante Folge 1 oder die konstante Folge -1 .

Hier wurden so genannte [Teilfolgen](#) ausgewählt.

Definition 4.1.5

Eine Folge $\{\tilde{a}_j\}$ heißt *Teilfolge* der Zahlenfolge $\{a_n\}$, wenn es eine Folge $\{n_j\}$ natürlicher Zahlen n_j mit $n_1 < n_2 < n_3 < \dots$ gibt, sodass

$$\tilde{a}_j = a_{n_j} \quad \text{für alle } j \in \mathbb{N}.$$

⁸B. Bolzano, 1781-1848; Prag.

⁹K. Weierstraß, 1815-1897; Münster, Braunschweig, Berlin.

Beispiele.

- i) Hat $1, -1, 1, -1, \dots$ noch andere Teilfolgen als die oben angegebenen?
- ii) Veranschaulicht wird der Begriff der Teilfolge am Beispiel aus Abbildung 4.3.

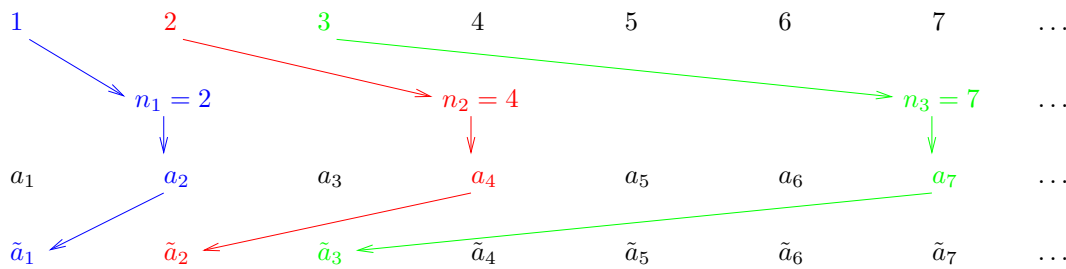


Abbildung 4.3: Zum Begriff der Teilfolge.

Satz 4.1.8 (*Satz von Bolzano-Weierstraß*)

Jede *beschränkte* reelle Zahlenfolge enthält eine *konvergente Teilfolge*.

Beweisidee. Sukzessive Intervallhalbierung, wobei in jedem Schritt jeweils die Hälfte ausgesucht wird, in der unendlich viele Folgenglieder enthalten sind. □

Bemerkungen.

- i) Die Folge $\{n\}$ besitzt **keine konvergente Teilfolge**, die Voraussetzung “beschränkt” ist also notwendig.
- ii) Der Satz ist ein wichtiges Hilfsmittel in der Analysis. Als eine Anwendung war schon das Cauchysche Konvergenzkriterium genannt.

Definition 4.1.6

Es sei $\{a_n\}$ eine reelle Zahlenfolge. Dann heißen die Grenzwerte konvergenter Teilfolgen von $\{a_n\}$ **Häufungspunkte** der Folge.

Bemerkungen und Beispiele.

- i) Nach dem Satz von Bolzano-Weierstraß besitzt jede beschränkte reelle Zahlenfolge mindestens einen Häufungspunkt.
- ii) Die Folge $1, -1, 1, -1 \dots$ hat die Häufungspunkte 1 und -1 .
- iii) Lässt man formal auch die Werte $\pm\infty$ zu, so hat jede reelle Zahlenfolge einen **größten Häufungspunkt**, den **limes superior**, und einen **kleinsten Häufungspunkt**, den **limes inferior**. Notation:

$$\limsup_{n \rightarrow \infty} a_n, \quad \liminf_{n \rightarrow \infty} a_n.$$

Vorsicht. Besondere Vorsicht ist stets bei der **Vertauschung von Grenzwerten** angeraten. Ist etwa für jedes fixierte $k \in \mathbb{N}$ die Folge $\{k/n\}$ zu untersuchen (deren Grenzwert ebenfalls wieder eine reelle Zahlenfolge ist), so gilt

$$\lim_{k \rightarrow \infty} \left[\lim_{n \rightarrow \infty} \frac{k}{n} \right] = \lim_{k \rightarrow \infty} 0 = 0.$$

Ist andererseits für jedes fixierte $n \in \mathbb{N}$ die Folge $\{k/n\}$ zu untersuchen, so ist

$$\lim_{k \rightarrow \infty} \frac{k}{n} = \infty$$

(was bedeutet: Für alle $K \in \mathbb{N}$ existiert ein $k_0(K, n)$, sodass für $k > k_0$ gilt $k/n > K$), demnach gilt

$$\lim_{n \rightarrow \infty} \left[\lim_{k \rightarrow \infty} \frac{k}{n} \right] = \infty.$$

4.2 Reelle Zahlenreihen (Konvergenz; Divergenz; g -adische Zifferndarstellung; Konvergenzkriterien; absolute Konvergenz)

Eine besonders bedeutende Art von Folgen (siehe die Beispiele am Anfang dieses Kapitels) ist von der Form:

Definition 4.2.1

i) Ist $\{a_n\}$ eine reelle Zahlenfolge, so heißt die Folge $\{s_k\}$, gegeben durch

$$s_k := \sum_{n=1}^k a_n$$

eine (unendliche) **Reihe**, s_k heißt die **k^{te} Partialsumme**.

ii) Eine Reihe heißt **konvergent**, wenn die Folge der Partialsummen gegen einen Grenzwert $s \in \mathbb{R}$ konvergiert. Konvergiert die Reihe nicht, so heißt sie **divergent**.

Notation. Man benutzt die Symbole

$$a_1 + a_2 + a_3 + \dots \quad \text{oder} \quad \sum_{n=1}^{\infty} a_n \quad \text{oder} \quad \sum_1^{\infty} a_n$$

in **doppelter Bedeutung**. Sie stehen einerseits für die Folge der Partialsummen, ist diese andererseits konvergent gegen $s \in \mathbb{R}$, so bezeichnen sie auch den **Grenzwert**

$$\sum_{n=1}^{\infty} a_n = s = \lim_{k \rightarrow \infty} s_k .$$

Bemerkungen.

- i) Eine Reihe ist **keine unendliche Summe** (Was soll das auch sein?), sondern ein **reelle Zahlenfolge von Partialsummen**.
- ii) Betrachtet man die Folge $a_m, a_{m+1}, a_{m+2}, \dots, m \in \mathbb{Z}$, so erhält man analog $\sum_{n=m}^{\infty} a_n$.
- iii) Es ist egal, wie der Summationsindex genannt wird.

Das wichtigste Beispiel. (Geometrische Reihe) Es sei $x \in \mathbb{R}$ fixiert und für alle $n \in \mathbb{N}_0$ sei $a_n := x^n$.

In den Übungen wurde bereits gezeigt, dass im Fall $|x| < 1$ gilt

$$s_k = \sum_{n=0}^k x^n = \frac{1 - x^{k+1}}{1 - x}.$$

Für $|x| < 1$ ist die geometrische Reihe demnach konvergent mit

$$\lim_{k \rightarrow \infty} s_k = \sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}.$$

Für $|x| > 1$ ist die Folge der Partialsummen nicht beschränkt, nach Satz 4.1.1 divergiert die Reihe in diesem Fall.

Gleiches gilt für $x = 1$, für $x = -1$ ist die Folge der Partialsummen $1, -1, 1, -1, \dots$, die Reihe divergiert in diesem Fall ebenfalls.

Satz 4.2.1

Die geometrische Reihe *konvergiert* für $|x| < 1$ und es ist

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}.$$

Für $|x| \geq 1$ *divergiert* die Reihe.

Anwendungen.

i) Die *g-adische Zifferndarstellung* einer reellen Zahl x lautet

$$x = n + x_0, \quad n \in \mathbb{N}_0, \quad 0 \leq x_0 < 1, \quad x_0 \in \mathbb{R}.$$

Dabei wurde **o.B.d.A.**¹⁰ $x \geq 0$ angenommen, für $x < 0$ kommt lediglich ein **Vorzeichen** “-” hinzu.

Der **ganze Anteil** von x ist in dieser Darstellung mit n abgespalten, wobei n geschrieben wird als endliche Summe ($k \in \mathbb{N}$)

$$n = \sum_{j=0}^k r_j \cdot g^j, \quad r_j \in \{0, 1, 2, \dots, g - 1\}.$$

¹⁰“Ohne Beschränkung der Allgemeinheit”; oder: “Ohne Einschränkung” (o.E.).

Dabei heißen die r_j die **Ziffern**, $g \in \mathbb{N}$ die **Basis** und k (falls $r_k \neq 0$) die **Stellenzahl**.

Im Fall $g = 2$ spricht man von der **Dualdarstellung** (Ziffern 0, 1), für $g = 10$ erhält man die **Dezimaldarstellung**, $g = 16$ entspricht der **Hexadezimaldarstellung**. Beispielsweise lautet die Zifferndarstellung von 10 im Dualsystem

$$10 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0$$

(Schreibweise: 1010).

Der Anteil x_0 wird als (unendliche) Reihe geschrieben:

$$x_0 = \sum_{j=1}^{\infty} s_j \cdot g^{-j}, \quad s_j \in \{0, 1, 2, \dots, g-1\}$$

(Schreibweise: $0.s_1s_2s_3\dots$).

Im Folgenden wird **o.E. stets die Dezimaldarstellung** einer reellen Zahl betrachtet – die Umrechnung in andere Systeme ist oben angedeutet.

Es gilt: Die Dezimaldarstellung jeder **rationalen Zahl** ist **endlich oder periodisch**, und jede periodische Zahl stellt eine rationale Zahl dar. Dezimaldarstellungen von **irrationalen Zahlen** sind **nicht periodisch**.

Dabei kann mit Hilfe der geometrischen Reihe auch definiert werden, was beispielsweise $0.\bar{3}$ bedeutet:

$$\begin{aligned} 0.\bar{3} &:= \sum_{j=1}^{\infty} 3 \cdot 10^{-j} = 3 \cdot \sum_{j=1}^{\infty} (10^{-1})^j \\ &= 3 \cdot \left(\left(\sum_{j=0}^{\infty} (10^{-1})^j \right) - 1 \right) = 3 \cdot \left(\frac{1}{1 - 10^{-1}} - 1 \right) \\ &= 3 \cdot \frac{1}{9} = \frac{1}{3}. \end{aligned}$$

(In der zweiten Gleichung durfte die 3 herausgezogen werden, da die geometrische Reihe in diesem Fall ($|10^{-1}| < 1$) konvergent ist, vgl. Satz 4.1.4.)

Man beachte, dass im Folgenden **Darstellungen, die mit $\bar{9}$ enden, ausgeschlossen werden**, denn mit der gleichen Rechnung würde folgen

$$0.\bar{9} = 9 \cdot \frac{1}{9} = 1 ,$$

$0.\bar{9}$ ist also bereits durch 1 repräsentiert (analog $0.356\bar{9}$ durch 0.357).

ii) Im zweiten Beispiel vom Anfang dieses Kapitels steckt kein Widerspruch, man rechnet in der Tat genau den Punkt des Überholvorgangs aus: A legt die Teilstrecken $10m$, $5m$, $2.5m$, ... zurück, also $10m \cdot 1$, $10m \cdot 1/2$, $10m \cdot 1/4$, ..., und es ist

$$10m \cdot \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j = 10m \cdot \frac{1}{1 - 1/2} = 20m .$$

Konvergenzkriterien für weitere Reihen?

Der Satz von der monotonen Folge impliziert

Satz 4.2.2

*Eine reelle Zahlenreihe $\sum_{n=1}^{\infty} a_n$ mit **nicht-negativen Gliedern a_n** ist **genau dann** konvergent, wenn es eine Zahl $k > 0$ gibt, sodass*

$$\sum_{j=1}^n a_j \leq k \quad \text{für alle } n \in \mathbb{N} .$$

Das Cauchysche Konvergenzkriterium liefert

Satz 4.2.3 (*Cauchys Konvergenzkriterium für Reihen*)

Eine reelle Zahlenreihe $\sum_{n=1}^{\infty} a_n$ ist **genau dann** konvergent, wenn es zu jedem $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt, sodass

$$|a_{n+1} + a_{n+2} + \cdots + a_{n+p}| < \varepsilon \quad \forall n \geq N(\varepsilon), \quad \forall p \geq 1.$$
Korollar 4.2.1 (*Wende Satz 4.2.3 an mit $p = 1$.*)

Wenn eine Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert, so bilden ihre Glieder a_n eine **Nullfolge**.

Die **Umkehrung** von Korollar 4.2.1 ist aber **nicht richtig**, wie ebenfalls aus dem Cauchyschen Konvergenzkriterium folgt.

Beispiel. (*harmonische Reihe*) $\sum_{n=1}^{\infty} \frac{1}{n}$. Für die Partialsummen gilt

$$\begin{aligned} s_{2k} - s_k &= \sum_{n=1}^{2k} \frac{1}{n} - \sum_{n=1}^k \frac{1}{n} = \frac{1}{k+1} + \frac{1}{k+2} + \cdots + \frac{1}{2k} \\ &> k \cdot \frac{1}{2k} = \frac{1}{2}, \end{aligned}$$

nach dem Cauchyschen Konvergenzkriterium **divergiert die harmonische Reihe**.

Anders ist die Situation im Fall $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$. Eine solche Reihe heißt **alternierend**, da für alle $n \in \mathbb{N}_0$ die Vorzeichen von a_n und a_{n+1} unterschiedlich sind. Aus geschicktem Zusammenfassen der Glieder der Reihe folgt die Konvergenz.

Satz 4.2.4 (*Konvergenzkriterium von Leibniz¹¹*)

Ist $\{a_n\}$ eine **monoton fallende Nullfolge**, so konvergiert die **alternierende Reihe** $a_1 - a_2 + a_3 - a_4 + \dots$.

Am Beispiel der harmonischen Reihe erkennt man also den Unterschied zwischen Konvergenz und **absoluter Konvergenz**:

¹¹G.W. Leibniz, 1646-1716; Paris, Hannover, Wien.

Definition 4.2.2

Eine Reihe $\sum_{n=1}^{\infty} a_n$ heißt **absolut konvergent**, wenn die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergent ist.

Bemerkungen und Beispiele.

- i) Die geometrische Reihe ist absolut konvergent.
- ii) Die Reihe $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ ist konvergent (nach Leibniz), aber nicht absolut konvergent. Aus absoluter Konvergenz folgt zwar die Konvergenz, nach diesem Beispiel ist die **Umkehrung aber falsch**.
- iii) Eine Reihe darf man nur dann **umordnen** (d.h. die Reihenfolge der Glieder vertauschen), wenn sie absolut konvergent ist. An dieser Stelle wird besonders deutlich, dass der Begriff “unendliche Summe” (in der die bekannten Rechenregeln gelten) nicht geeignet ist: **Wäre dem so, so folgte**

$$\begin{aligned}
 \sum_{k=0}^{\infty} (-1)^k \frac{1}{k+1} &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \pm \dots \\
 &\stackrel{(*)}{=} 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \\
 &\quad + \left(\frac{1}{5} + \frac{1}{7} \right) - \frac{1}{6} \\
 &\quad + \left(\frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} \right) - \frac{1}{8} \\
 &\quad + \dots \\
 &\quad + \left(\frac{1}{2^n + 1} + \frac{1}{2^n + 3} + \dots + \frac{1}{2^{n+1} - 1} \right) - \frac{1}{2n + 2} \\
 &\quad + \dots
 \end{aligned}$$

Dabei ist zu beachten, dass tatsächlich alle Glieder der “unendlichen Summe” auf der linken Seite der Gleichung auch in der “unendlichen Summe” auf der rechten Seite der Gleichung vorkommen

und vice versa, also scheint die Gleichheit (*) plausibel.

Es gilt jedoch

$$\left(\frac{1}{2^n + 1} + \frac{1}{2^n + 3} + \cdots + \frac{1}{2^{n+1} - 1} \right) > 2^{n-1} \cdot \frac{1}{2^{n+1}} = \frac{1}{4}$$

(es kommen 2^{n-1} Terme vor, wobei jeder einzelne größer als $1/2^{n+1}$ ist), die Reihe auf der rechten Seite der Gleichheit konvergiert nach dem Cauchyschen Kriterium nicht.

Die Reihe auf der linken Seite der Gleichheit konvergiert aber nach dem Kriterium von Leibniz, (*) kann nicht richtig sein, man darf mit Reihen nicht ohne Weiteres so rechnen, wie man es von Summen her kennt.

iv) Für eine absolut konvergente Reihe gilt

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n| .$$

Sehr häufig kann die Konvergenz einer Reihe über einen Vergleich mit einer bekannten Reihe (oft der geometrischen Reihe) gezeigt werden:

Definition 4.2.3

Eine reelle Zahlenreihe $\sum_{n=1}^{\infty} b_n$, $b_n \geq 0$ für alle $n \in \mathbb{N}$, heißt eine **Majorante** einer reellen Zahlenreihe $\sum_{n=1}^{\infty} a_n$, wenn es einen Index $n_0 \in \mathbb{N}$ gibt, sodass gilt

$$|a_n| \leq b_n \quad \text{für alle } n \geq n_0 .$$

Aus Satz 4.2.2 folgt

Satz 4.2.5 (*Majorantenkriterium*)

Besitzt eine reelle Zahlenreihe $\sum_{n=1}^{\infty} a_n$ eine **konvergente Majorante** $\sum_{n=1}^{\infty} b_n$, so ist sie absolut konvergent und damit insbesondere konvergent. Es gilt dann

$$\sum_{n=1}^{\infty} |a_n| \leq \sum_{n=1}^{\infty} b_n .$$

Bemerkungen.

- i) Die Divergenz einer Reihe kann über die **Divergenz einer Minorante** gezeigt werden. Gilt beispielsweise $\frac{1}{n} \leq b_n$ für alle $n \in \mathbb{N}$, so muss $\sum_{n=1}^{\infty} b_n$ divergent sein (warum?).
- ii) Der Vergleich mit der geometrischen Reihe liefert die folgenden expliziten Konvergenzkriterien:

Satz 4.2.6 (*Quotientenkriterium; Wurzelkriterium*)¹²

Es sei $\sum_{n=1}^{\infty} a_n$ eine reelle Zahlenreihe mit $a_n \neq 0$ für alle $n \in \mathbb{N}$. Die Reihe ist **absolut konvergent** (und damit konvergent), wenn eine der beiden folgenden Bedingungen erfüllt ist.

- i) Es existieren ein $0 < q < 1$ und ein $N \in \mathbb{N}$ derart, dass

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q < 1 \quad \text{für alle } n \geq N .$$

- ii) Es existieren ein $0 < q < 1$ und ein $N \in \mathbb{N}$ derart, dass

$$\sqrt[n]{|a_n|} \leq q < 1 \quad \text{für alle } n \geq N .$$

Die Reihe ist **divergent**, falls ein $N \in \mathbb{N}$ existiert, sodass für alle $n \geq N$ gilt

$$\left| \frac{a_{n+1}}{a_n} \right| \geq 1 \quad \text{oder für das Wurzelkriterium } \sqrt[n]{|a_n|} \geq 1 .$$

Bemerkungen und Beispiele.

- i) Im Fall $q = 1$ liefern weder das Quotienten- noch das Wurzelkriterium eine Aussage.

¹²Ist $x \geq 0$ und $n \in \mathbb{N}$, so ist $\sqrt[n]{x} = y \Leftrightarrow x = y^n$.

ii) Existiert der Grenzwert

$$q = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \quad \text{oder} \quad q = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|},$$

so hat man für $q < 1$ absolute Konvergenz, für $q > 1$ Divergenz und für $q = 1$ keine Aussage.

iii) Die Reihe

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}$$

ist absolut konvergent. Es gilt nämlich

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

und das Quotientenkriterium liefert die Aussage mit $q = 1/2$. Analog kann man argumentieren

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{e} < 1,$$

woraus die absolute Konvergenz der Reihe folgt.

iv) Betrachtet sei die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Es ist

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

Hier gibt es **kein festes** $q < 1$ mit $|a_{n+1}/a_n| \leq q < 1$ und das Quotientenkriterium liefert keine Aussage. Gleiches gilt für das Wurzelkriterium. Mit Hilfe des so genannten **Cauchyschen Verdichtungskriterium** kann das Problem aber auf eine geometrische Reihe reduziert werden, es gilt: Für alle $\alpha > 1$ konvergiert die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha},$$

für alle $\alpha \leq 1$ divergiert diese Reihe.

Schließlich sei noch als **Rechenregel** festgehalten:

Satz 4.2.7

Es seien $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ **konvergente** reelle Zahlenreihen. Dann ist auch für alle $\alpha, \beta \in \mathbb{R}$ die Reihe

$$\sum_{n=1}^{\infty} (\alpha \cdot a_n + \beta \cdot b_n)$$

konvergent, der Grenzwert ist

$$s = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n .$$

4.3 Maschinenzahlen (Festpunktdarstellung; Gleitpunktdarstellung; relativer Rundungsfehler; Maschinengenauigkeit; Rundungsabbildung; Maschinenoperationen)

Viele Probleme sind sinnvoller Weise mit computergestützten Rechenverfahren anzugehen.

Beispiele.

i) Eine **exakte Lösung** eines Problems **existiert** zwar, kann aber nicht **explizit berechnet werden**. Gedacht ist dabei z.B. an die Fragestellungen:

- Wie lautet der exakte Wert von

$$\int_0^{10} e^{t^2} dt ?$$

- Wie lautet die (eindeutig bestimmte) Lösung des **Anfangswertproblems**

$$y'(t) = t \cdot e^{\sin(y(t))} , \quad y(0) = 0 ?$$

- ii)* Eine exakte Lösung eines Problems kann zwar **prinzipiell bestimmt werden**, der Aufwand ist aber zu groß. Typischerweise denke man an die Lösung eines linearen Gleichungssystems mit 1000 Unbekannten und 1000 Gleichungen.
- iii)* Interpolationsaufgaben (Auswertung von Messreihen): Die Auswertung von Meßreihen wurde bereits in Paragraph 3.2 diskutiert (Interpolationsaufgabe von Lagrange, Splines). Das Problem kann aber auch lauten, eine Kurve nach der **Methode der kleinsten Quadrate** “anzufitten” (vgl. Abschnitt 8.4).

Entscheidend bei **numerischen Verfahren** ist die **Genauigkeit** einer numerischen Rechnung. Hierbei berücksichtigt der **Gesamtfehler**:

i) **Modellfehler**, d.h.

- **Idealisierungsfehler** (ein physikalischer Sachverhalt muss in ein mathematisches Modell “übersetzt” werden, dabei sind oft Vereinfachungen (z.B. **Linearisierungen**) notwendig, um das Modell aufstellen und behandeln zu können);
- **Datenfehler** (die Daten des Modells sind etwa aufgrund ungenauer Kenntnis von Materialeigenschaften bzw. aufgrund fehlerhafter Messwerte ebenfalls fehlerhaft).

ii) **Numerische Fehlern**, d.h.

- **Diskretisierungsfehler** (**kontinuierliche** Prozesse müssen in Rechenmaschinen durch **diskrete** ersetzt werden);
- **Abbruchsfehler** (unendliche **Algorithmen** werden nach endlich vielen Schritten abgebrochen);
- **Rundungsfehler** (ein Computer kann nur **endlich viele Zahlen darstellen**).

In diesem Abschnitt wird einführend auf den letzten Punkt eingegangen und der Begriff der **Maschinenzahlen** vorgestellt.

Die Darstellung reeller Zahlen in Rechenmaschinen erfolgt im **Dualsystem** mit den Ziffern 0 und 1 (vgl. die Diskussion g -adischer Zifferndarstellungen in Paragraph 4.2). Dies ist jedoch kein grundsätzlicher Unterschied zur Darstellung im Dezimalsystem und **im Folgenden**

wird o.E. das Dezimalsystem betrachtet.

Zur Realisierung in einer Maschine stehen für eine reelle Zahl **endlich viele Stellen** zur Verfügung. Diese Stellen werden auf zwei verschiedene Arten verwaltet:

- i) **Festpunktdarstellung**: Hier ist die Anzahl der Stellen vor und nach dem Komma fixiert. Stehen etwa drei Vor- und fünf Nachkommastellen zur Verfügung, so erhält man folgende Festpunktdarstellungen:

$$\begin{aligned} 27.3025 &\rightarrow 027|30250 ; \\ 0.103 &\rightarrow 000|10300 . \end{aligned}$$

- ii) **Gleitpunktdarstellung**: Diese ist in der Regel vorzuziehen und basiert darauf, dass eine reelle Zahl x in der Form

$$x = a \cdot 10^E, \quad a \in \mathbb{R}, \quad |a| < 1, \quad E \in \mathbb{Z},$$

geschrieben werden kann. Dabei heißt a die **Mantisse**, E der **Exponent**. Es ist beispielsweise

$$\begin{aligned} x = 27.3025 &= 0.273025 \cdot 10^2 \\ &= 0.00273025 \cdot 10^4, \end{aligned}$$

und die Darstellung ist noch **nicht eindeutig**. Deshalb wird zur **normalisierten Gleitpunktdarstellung** übergegangen. Für $x \neq 0$ fordert man dabei

$$|a| \geq 10^{-1}.$$

Damit ist die Darstellung eindeutig (zur Erinnerung: In Paragraph 4.2 wurde o.E. angenommen, dass eine Zahl (in diesem Fall a) nicht mit einer Periode 9 dargestellt wird). Für $x = 0$ ist $a = 0$ und E beliebig.

Im Folgenden wird **stets von der Darstellung einer Zahl in normalisierter Gleitpunktdarstellung** ausgegangen.

Bisher wurden beliebige reelle Zahlen betrachtet. Die **Situation auf einer konkreten Rechenanlage** ist die Folgende: Es stehen $(t, s \in \mathbb{N})$

t Ziffern und 1 Vorzeichen für a ;

s Ziffern und 1 Vorzeichen für E

für die Darstellung einer Zahl zur Verfügung, als Konsequenz ist nur eine **endliche Menge** reeller Zahlen auf der Maschine exakt darstellbar. Sie heißt die Menge der **Maschinenzahlen** $A = A(t, s)$. Alle anderen Zahlen werden **fehlerhaft** dargestellt, es kommt zu **Rundungsfehlern**.

Bemerkung. Die Menge der Maschinenzahlen hängt über die Parameter t und s von der konkret betrachteten Rechenanlage ab.

$A(t, s)$ ist eine endliche Menge, insbesondere gibt es die Größen (Berechnung als Übungsaufgabe)

c_{max} : **Größte exakt darstellbare Zahl;**

c_{min} : **Kleinste exakt darstellbare Zahl.**

Ebenso kann die Null nicht beliebig gut mit Maschinenzahlen approximiert werden (außer mit der Null selbst), und es gibt (wie immer in normalisierter Gleitpunktdarstellung)

$c_{+,min}$: **Kleinste positive Maschinenzahl;**

$c_{-,max}$: **Größte negative Maschinenzahl.**

Definiert man schließlich (zu beachten ist, dass es sich um kontinuierliche Intervalle und nicht um Teilmengen der Maschinenzahlen handelt)

$$D := [c_{min}, c_{-,max}] \cup \{0\} \cup [c_{+,min}, c_{max}] ,$$

so können alle $x \in D$ in sinnvoller Weise **gerundet** werden, d.h. man kann x eine Maschinenzahl "in relativer Nähe" zuordnen: Es sei $x \in D$, $x = a \cdot 10^E$ und o.E. sei $x > 0$ (Vorzeichen bleiben bei der Rundung natürlich erhalten, $x = 0$ wird exakt dargestellt). Die Mantisse

$$a = 0.a_1a_2a_3 \dots a_t a_{t+1} \dots \quad (a_1 \neq 0)$$

wird rechnerintern dargestellt als

$$\tilde{a} = \begin{cases} 0.a_1a_2a_3 \dots a_t , & \text{falls } 0 \leq a_{t+1} \leq 4 , \\ 0.a_1a_2a_3 \dots a_t + 10^{-t} , & \text{falls } 5 \leq a_{t+1} \leq 9 . \end{cases}$$

Der **relative Rundungsfehler** ist ($\tilde{x} := \tilde{a} \cdot 10^E$)

$$|\varepsilon_x| := \left| \frac{\tilde{x} - x}{x} \right| \leq 5 \cdot 10^{-t} =: \text{eps} ,$$

eps heißt die **Maschinengenauigkeit**.

Die oben skizzierte Umwandlung einer reellen Zahl in eine Maschinenzahl wird beschrieben durch die **Rundungsabbildung**,

$$x \mapsto \text{rd}(x) := \tilde{x} = x \cdot (1 + \varepsilon_x) .$$

Bemerkung. Die Rundungsabbildung ist definiert für $x \in D$. Ist dagegen $x \notin D$, so kann es zum **Exponentenüberlauf** oder zum **Exponentenunterlauf** kommen.

Beispiel. Ist $t = 4$, $s = 2$ und

$$x = 0.99998 \cdot 10^{99} ,$$

so würde die Rundung

$$\tilde{x} = 0.1000 \cdot 10^{100}$$

lauten. Dies ist jedoch **keine Maschinenzahl** mehr, wodurch es zum **Abbruch** des Verfahrens kommt.

Schließlich müssen auch die Rechenoperationen $+$, $-$, \cdot , $/$ als **Maschinenoperationen** \oplus , \ominus , \odot , \oslash ausgeführt werden. Für die Addition gilt beispielsweise (alle anderen Operationen analog)

$$x \oplus y := \text{rd}(x + y) = (x + y)(1 + \varepsilon) , \quad \text{mit } |\varepsilon| < \text{eps} .$$

Dabei wird (eine Näherung für) den Wert $x + y$ in der Regel maschinenintern mit erhöhter Stellenzahl berechnet, erst anschließend wird gerundet. Deshalb ist die obige Fehlerabschätzung richtig.

Beispiel. Ist ($t = 3$, $s = 1$) $x = 0.106 \cdot 10^2$, $y = 0.612 \cdot 10^0$, so ist

$$x \oplus y = \text{rd}(0.11212 \cdot 10^2) = 0.112 \cdot 10^2 .$$

Bemerkung. Die **üblichen Rechengesetze gelten nicht mehr!** (Ein Beispiel wird in den Übungen diskutiert.)

Kapitel 5

Funktionenfolgen, Potenzreihen, Exponentialfunktion

5.1 Funktionenfolgen, Funktionenreihen (punktweise und gleichmäßige Konvergenz; gleichmäßige absolute Konvergenz)

Über **Funktionenfolgen** bzw. **Funktionenreihen** werden einerseits viele elementare Funktionen definiert. Andererseits spielen diese Begriffe eine wesentliche Rolle, um praktische Probleme handhaben zu können (Stichworte: Taylor- oder Fourier-Reihen, Numerik, Regularisierungen von Variationsproblemen ...).

Die Definition einer Folge (vgl. Definition 4.1.1) beschränkt sich nicht auf Folgen reeller Zahlen. Sind $f_n: \mathbb{R} \supset U \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, reelle Funktionen, so liefert Definition 4.1.1 eine Funktionenfolge $\{f_n\}$. Für festes $x \in U$ ergibt sich die **reelle Zahlenfolge** $\{f_n(x)\}$, und es ist wohl definiert:

Definition 5.1.1

*Eine Funktionenfolge $\{f_n\}$ wie oben heißt **punktweise konvergent**, wenn für jedes $x \in U$ die reelle Zahlenfolge $\{f_n(x)\}$ konvergiert. Für jedes $x \in U$ ist dann die **Grenzfunktion** definiert durch*

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) .$$

Beispiele.

i) Es sei $U = (0, 1)$, für alle $n \in \mathbb{N}$ und für alle $x \in U$ sei

$$f_n(x) = \frac{x}{n}.$$

Dann konvergiert $f_n(x)$ punktweise gegen die Grenzfunktion $f \equiv 0$.

ii) Wieder sei $U = (0, 1)$, für alle $n \in \mathbb{N}$ und für alle $x \in U$ sei jetzt

$$f_n(x) := \begin{cases} 1 - n \cdot x & \text{für } 0 < x \leq 1/n, \\ 0 & \text{für } 1/n < x < 1. \end{cases}$$

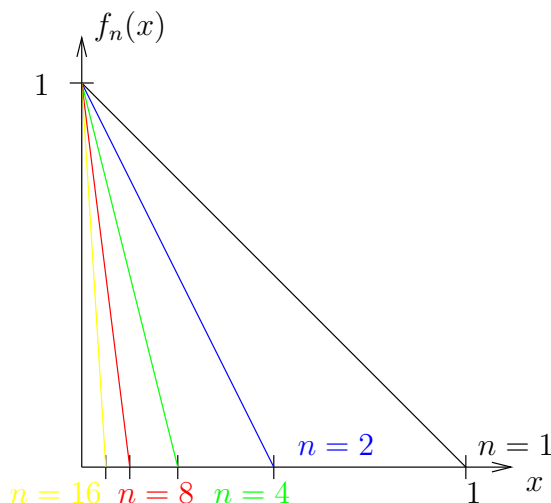


Abbildung 5.1: Die Funktionen $f_n(x)$.

Ist $x > 0$ fixiert, so gilt für alle $n > \frac{1}{x}$: $f_n(x) = 0$, die Folge ist also punktweise konvergent mit Grenzfunktion $f(x) = 0$ für alle $x \in (0, 1)$.

Die Folge $\{f_n(x)\}$ konvergiert im zweiten Beispiel aber **“nicht für alle x mit gleicher Geschwindigkeit”** gegen die Null, egal wie groß man n wählt gibt es beispielsweise immer ein $x \in U$ mit $f(x) > 1/2$. Dies wird mit der so genannten **Supremumsnorm** (vgl. Übungen) gemessen,

$$\|f\|_\infty := \sup_{x \in U} |f(x)|,$$

die den (“maximalen”) **Abstand** zweier beschränkter Funktionen auf U angibt:

$$\|f - g\|_\infty := \sup_{x \in U} |f(x) - g(x)| .$$

Zurück zu den Beispielen.

i) In diesem Beispiel ist

$$\sup_{x \in U} |f_n(x) - f(x)| = \sup_{x \in (0,1)} \left| \frac{x}{n} - 0 \right| \leq \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0 ,$$

auch die Supremumsnorm von $f_n - f$ konvergiert also gegen Null.

ii) Im zweiten Beispiel ist

$$\|f_n(x) - f(x)\|_\infty = 1 \quad \text{für alle } n \in \mathbb{N} ,$$

das Konvergenzverhalten ist hier (zumindest in der Nähe des Nullpunktes) nicht sehr gut.

Definition 5.1.2

Eine Funktionenfolge $\{f_n\}$ auf U beschränkter Funktionen heißt **gleichmäßig konvergent** gegen eine Grenzfunktion f auf U , wenn

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0 .$$

Bemerkungen.

- i) Gleichmäßige Konvergenz impliziert punktweise Konvergenz, nach obigem Beispiel ii) ist die **Umkehrung** aber **falsch**.
- ii) Es gilt wieder ein **Cauchy-Kriterium** zur Überprüfung gleichmäßiger Konvergenz ($\|f_n - f_m\|_\infty < \varepsilon$ für n, m hinreichend groß).
- iii) Im Fall gleichmäßiger Konvergenz benutzt man häufig die Notation:

$$f_n \rightrightarrows f \quad \text{für } n \rightarrow \infty .$$

Mit einer Funktionenfolge $\{f_n\}$, $f_n: \mathbb{R} \supset U \rightarrow \mathbb{R}$ für alle $n \in \mathbb{N}$, sind (für festes $x \in U$) analog zu Paragraph 4.2 die **Partialsommen**

$$s_k(x) := \sum_{n=1}^k f_n(x), \quad k \in \mathbb{N},$$

definiert. Wie oben kann $\{s_k\}$ als **Funktionenfolge**, $s_k: U \rightarrow \mathbb{R}$ für alle $k \in \mathbb{N}$, interpretiert werden.

Eine **Funktionenreihe** ist die **Folge der Partialsommen einer Funktionenfolge** $\{s_k\}$, und man definiert:

Definition 5.1.3

Die Reihe $\sum_{n=1}^{\infty} f_n$, f_n wie oben, heißt **punktweise** bzw. **gleichmäßig** konvergent, wenn die Folge $\{s_k\}$ der Partialsommen punktweise bzw. gleichmäßig konvergent ist, falls also mit einer Grenzfunktion $s: U \rightarrow \mathbb{R}$ gilt

$$s(x) = \lim_{k \rightarrow \infty} s_k(x) = \sum_{n=1}^{\infty} f_n(x) \quad \text{für alle } x \in U,$$

bzw.

$$s_k = \sum_{n=1}^k f_n \rightrightarrows s \quad \text{für } k \rightarrow \infty.$$

Bemerkungen.

- i) Hier wird stets die Beschränktheit der s_k bzw. von s angenommen.
- ii) Eine Funktionenreihe ist insbesondere dann gleichmäßig konvergent, wenn sie **gleichmäßig absolut konvergent** ist, d.h. nach Definition: Die reelle Zahlenreihe $\sum_{n=1}^{\infty} \|f_n\|_{\infty}$ konvergiert.

Nach dem Majorantenkriterium (analog zu Satz 4.2.5) ist eine Funktionenreihe z.B. dann gleichmäßig absolut konvergent, wenn es ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq n_0$ gilt $\|f_n\|_{\infty} \leq a_n$, wobei $\sum_{n=1}^{\infty} a_n$ konvergiere.

5.2 Potenzreihen (Konvergenzradius; Konvergenzintervall)

Hier werden lediglich einige einführende Betrachtungen gemacht, über die elementare Funktionen wie die [Exponentialfunktion](#) definiert werden. [Potenzreihen](#) werden später im Rahmen der [Funktionentheorie](#) noch eingehender untersucht.

Potenzreihen sind spezielle Funktionenreihen, bei denen die Glieder von der Form $a_n \cdot (x - x_0)^n$ sind, die Folge der Partialsummen ist also eine Folge von Polynomen.

Definition 5.2.1

Eine (Funktionen-) Reihe der Form

$$\sum_{n=0}^{\infty} a_n \cdot (x - x_0)^n$$

*heißt **Potenzreihe** um den (festen) **Entwicklungspunkt** $x_0 \in \mathbb{R}$ mit den **Koeffizienten** a_n .*

Beispiel. Die geometrische Reihe: $\sum_{n=0}^{\infty} (x - x_0)^n$. Nach Satz 4.2.1 konvergiert die Reihe (punktweise) für $|x - x_0| < 1$, sie divergiert für $|x - x_0| > 1$. Für $|x - x_0| = 1$ divergiert die Reihe ebenfalls, dazu mussten die “Randpunkte” $x_0 - 1$ und $x_0 + 1$ aber einzeln untersucht werden.

Zeigen andere Potenzreihen ein ähnliches Verhalten?

In der Tat beweist man über einen genauen Vergleich mit der geometrischen Reihe (Majorantenkriterium):

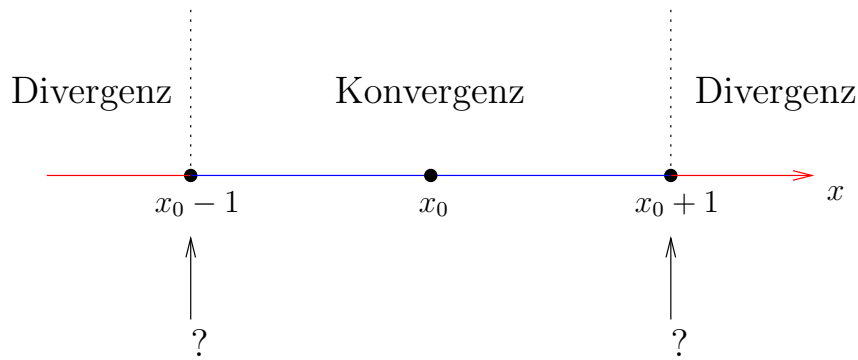


Abbildung 5.2: Zum Konvergenzverhalten der geometrischen Reihe.

Satz 5.2.1

Zu jeder Potenzreihe $\sum_{n=0}^{\infty} a_n \cdot (x - x_0)^n$ wie oben gibt es ein $\rho \geq 0$ und ein Intervall $(x_0 - \rho, x_0 + \rho)$ derart, dass gilt:

- i) Die Potenzreihe **konvergiert** für $x \in (x_0 - \rho, x_0 + \rho)$ punktweise. Sie konvergiert außerdem **gleichmäßig absolut** in jedem **abgeschlossenen Teilintervall** von $(x_0 - \rho, x_0 + \rho)$.
- ii) **Außerhalb von $[x_0 - \rho, x_0 + \rho]$ divergiert** die Potenzreihe.

Bemerkungen.

- i) Die Zahl ρ heißt der **Konvergenzradius** der Potenzreihe. Im Fall $0 < \rho < \infty$ heißt $(x_0 - \rho, x_0 + \rho)$ das **Konvergenzintervall**.
- ii) Über Konvergenz bzw. Divergenz der Potenzreihe **am Rande des Konvergenzintervalls** lassen sich **keine allgemeinen Aussagen** machen.

iii) Im Fall $\rho = 0$ konvergiert die Reihe nur für $x = x_0$, im Falle der Konvergenz auf ganz \mathbb{R} setzt man **formal** $\rho = \infty$. Dabei sind **uneigentliche Intervalle** wie $[0, \infty)$ definiert als $\{x \in \mathbb{R} : x \geq 0\}$.

iv) Der Konvergenzradius genügt der **Formel von Cauchy-Hadamard**¹ (Beweis: Wurzelkriterium, Satz 4.2.6; man setzt $\frac{1}{0} := \infty$, $\frac{1}{\infty} = 0$):

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

v) Zur Anwendung der Formel von Cauchy-Hadamard ist es beispielsweise nützlich zu wissen:

$$\sqrt[n]{n} \rightarrow 1, \quad \text{falls } n \rightarrow \infty.$$

Beweis. Ist h_n definiert mittels $\sqrt[n]{n} = 1 + h_n$, so ist insbesondere $h_n \geq 0$. Potenzieren liefert zusammen mit dem binomischen Lehrsatz

$$n = (1 + h_n)^n = \sum_{k=0}^n \binom{n}{k} h_n^k.$$

Werden bis auf $k = 2$ alle Terme weggelassen, so ergibt dies ($n \geq 2$)

$$n \geq \binom{n}{2} h_n^2 = \frac{n(n-1)}{2} h_n^2.$$

Es folgt (für $n \geq 2$)

$$h_n \leq \sqrt{\frac{2}{n-1}},$$

d.h. $h_n \rightarrow 0$ für $n \rightarrow \infty$, und die Behauptung ist bewiesen. \square

vi) Zur Überprüfung punktweiser Konvergenz können natürlich für festes x die Kriterien aus Paragraph 4.2 angewandt werden.

¹J. Hadamard, 1865-1963; Bordeaux, Paris.

5.3 Die Exponentialfunktion (Cauchy-Produkt; Logarithmus; allgemeine Potenzfunktion; Umkehrfunktion)

Definition 5.3.1

Die durch die Potenzreihe

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

definierte Funktion $\exp: \mathbb{R} \rightarrow \mathbb{R}$ heißt *Exponentialfunktion*.

Bemerkungen.

- i) Zunächst ist zu überprüfen, ob diese Definition überhaupt sinnvoll ist. Dazu fixiert man ein beliebiges $x \in \mathbb{R}$, $x \neq 0$ (für $x = 0$ ist $\exp(x) = 1$). Um etwa das Quotientenkriterium (Satz 4.2.6) anzuwenden, wird nachgerechnet:

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{|x|^{n+1} n!}{(n+1)! |x|^n} = \frac{|x|}{n+1} < \frac{1}{2}$$

für alle $n > 2|x| - 1$. Demnach konvergiert die Reihe für alle $x \in \mathbb{R}$ absolut und $\exp(x)$ ist wohl definiert.

- ii) Nach Satz 4.1.6 und der anschließenden Bemerkung *ii*) gilt $\sum_{n=0}^{\infty} \frac{1}{n!} = e$, d.h. $\exp(1) = e$.

Gilt mit dieser Definition die **Funktionalgleichung** (das **Additionstheorem**) $\exp(x) \cdot \exp(y) = \exp(x + y)$?

Zur Beantwortung der Frage muss zunächst das Produkt von Reihen studiert werden.

Satz 5.3.1 (*Cauchy-Produkt von Reihen*)

Die reellen Zahlenreihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ seien absolut konvergent. Dann gilt

$$\begin{aligned} \left(\sum_{n=0}^{\infty} a_n \right) \cdot \left(\sum_{n=0}^{\infty} b_n \right) &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n a_k \cdot b_{n-k} \right) \\ &= a_0 \cdot b_0 + (a_0 \cdot b_1 + a_1 \cdot b_0) \\ &\quad + (a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0) \\ &\quad + \dots \end{aligned}$$

Insbesondere ist die Reihe auf der rechten Seite absolut konvergent.

Korollar 5.3.1 (*Funktionalgleichung der Exponentialfunktion*)

Für alle $x, y \in \mathbb{R}$ gilt

$$\exp(x + y) = \exp(x) \cdot \exp(y) .$$

Beweis. Die absolute Konvergenz aller vorkommenden Reihen ist bereits gezeigt. Satz 5.3.1 impliziert (zusammen mit dem binomischen Lehrsatz 2.1.5):

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{x^k \cdot y^{n-k}}{k! \cdot (n-k)!} \right) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^n \binom{n}{k} \cdot x^k \cdot y^{n-k} \right) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (x + y)^n \\ &= \exp(x + y) , \end{aligned}$$

was zu beweisen war. □

Bemerkungen.

i) Es folgt

$$(\exp(x) \cdot \exp(-x)) = \exp(x + (-x)) = \exp(0) = 1 ,$$

also auch

$$\exp(-x) = \frac{1}{\exp(x)} .$$

ii) Der Graph der Exponentialfunktion ist in Abbildung 5.3 dargestellt.

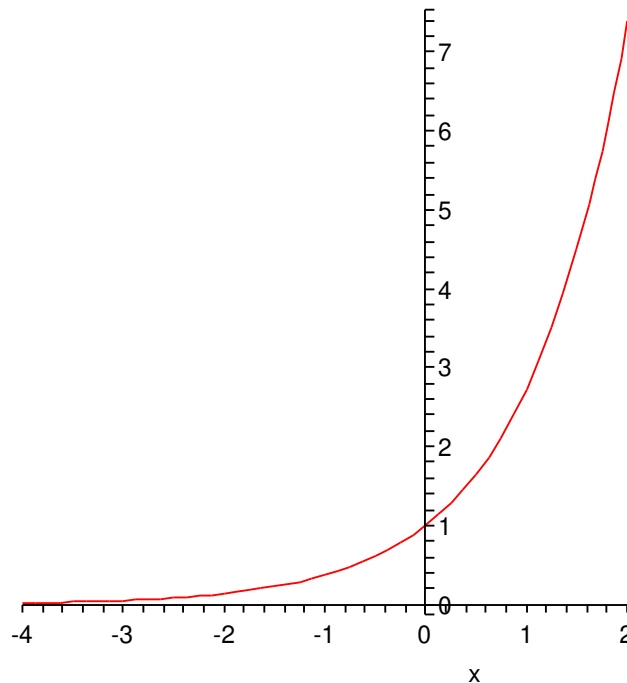


Abbildung 5.3: Der Graph der Exponentialfunktion.

- iii) Die Exponentialfunktion ist **streng monoton wachsend**, wie leicht anhand der Definition (im Fall $x \geq 0$) und der Funktionalgleichung (für $x < 0$) verifiziert werden kann (vgl. wieder Abbildung 5.3).
- iv) Wie ebenfalls in Abbildung 5.3 angedeutet, ist die Exponentialfunktion eine **bijektive Abbildung** $\exp: \mathbb{R} \rightarrow (0, \infty)$. Zu beweisen ist das mit Hilfe der strengen Monotonie und der gleichmäßigen Konvergenz der Potenzreihe (die Bijektivität wird später auch aus

der Stetigkeit und der strengen Monotonie folgen). Die Bijektivität wiederum zeigt die **Existenz einer Umkehrfunktion**.

Definition 5.3.2

Die *Umkehrfunktion der Exponentialfunktion* $\exp: \mathbb{R} \rightarrow (0, \infty)$ heißt *natürlicher Logarithmus*, $\ln: (0, \infty) \rightarrow \mathbb{R}$, $x \mapsto \ln(x)$.

Bemerkungen.

- i) Wie in Abbildung 5.4 dargestellt, geht der Graph des natürlichen Logarithmus aus dem der Exponentialfunktion durch eine **Spiegelung an der Winkelhalbierenden** hervor (Übung zu Kapitel 8.7).

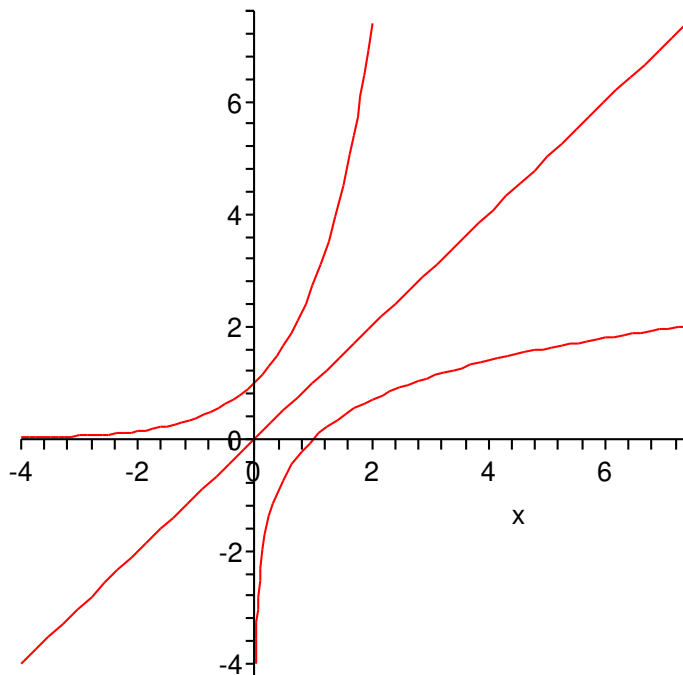


Abbildung 5.4: Der Graph des natürlichen Logarithmus geht aus dem der Exponentialfunktion durch Spiegelung an der Winkelhalbierenden hervor.

ii) Nach Definition gilt $\exp(\ln(x)) = x$ für alle $x > 0$ und $\ln(\exp(x)) = x$ für alle $x \in \mathbb{R}$.

iii) Es gilt $\ln(1) = 0$, $\ln(e) = 1$, $\ln(x)$ ist streng monoton wachsend.

iv) Aus Korollar 5.3.1 folgt für alle positiven x, y

$$\exp(\ln(x \cdot y)) = x \cdot y = \exp(\ln(x)) \cdot \exp(\ln(y)) = \exp(\ln(x) + \ln(y)) ,$$

also das **Logarithmengesetz**

$$\ln(x \cdot y) = \ln(x) + \ln(y)$$

sowie

$$\ln(x/y) = \ln(x) - \ln(y) , \quad \ln(1/x) = -\ln(x) .$$

Über den natürlichen Logarithmus lässt sich nun die **allgemeine Potenzfunktion** einführen.

Definition 5.3.3

Es sei $a > 0$ und $x \in \mathbb{R}$. Dann heißt die Funktion $f: \mathbb{R} \rightarrow (0, \infty)$,

$$f(x) := a^x := \exp(x \cdot \ln(a))$$

*Potenzfunktion zur Basis a . Diese Funktion ist bijektiv (warum?), es existiert eine Umkehrfunktion $g(x): (0, \infty) \rightarrow \mathbb{R}$, genannt die **Logarithmusfunktion zur Basis a** , $g(x) =: \log_a(x)$.*

Bemerkungen.

i) Es folgen unmittelbar die Regeln

$$a^{x+y} = a^x \cdot a^y , \quad a^{-x} = \frac{1}{a^x} , \quad (a^x)^y = a^{x \cdot y} , \quad a^0 = 1 ,$$

wobei aus der ersten Regel für $n \in \mathbb{N}$ induktiv folgt

$$a^n = \underbrace{a \cdot a \dots a}_{n\text{-mal}} ,$$

die Definition beinhaltet also die “anschauliche Definition” einer Potenz mit Exponenten aus \mathbb{N} .

ii) $a^{1/n}$ stimmt mit der bekannten n -ten Wurzel $\sqrt[n]{a}$ überein:

$$\begin{aligned} (a^{1/n})^n &= \exp(n \cdot \ln(a^{1/n})) = \exp(n \cdot \ln(\exp[(1/n) \cdot \ln(a)])) \\ &= \exp(\ln(\exp^n[(1/n) \cdot \ln(a)])) \\ &= \exp(\ln(\exp[\ln(a)])) \\ &= \exp(\ln(a)) = a . \end{aligned}$$

iii) Die Exponentialfunktion ist die Potenzfunktion zur Basis e , da gilt

$$e^x = \exp(x \cdot \ln(e)) = \exp(x \cdot \ln(\exp(1))) = \exp(x) .$$

iv) Per definitionem gibt der Logarithmus $\log_a(x)$ den Exponenten y an, für den $a^y = x$ gilt.

Ist beispielsweise $a = 4$, $x = 64$ so gilt $a^x = a^3 = 64$. Die Umkehrfunktion bildet aber 64 auf das Urbild 3 ab, also $\log_4(64) = 3$.

Allgemein wird für festes $a > 0$ die Funktion

$$f : \mathbb{R} \rightarrow (0, \infty) , \quad x \mapsto a^x ,$$

betrachtet. Die Umkehrfunktion f^{-1} (genannt \log_a) ist definiert als

$$f^{-1}(y) = \log_a(y) := x , \quad \text{falls gilt } y = f(x) = a^x .$$

Kapitel 6

Der \mathbb{R}^n

6.1 Der Vektorraum \mathbb{R}^n (Vektorräume; Funktionenräume; lineare Abhängigkeit; Dimension; Basis; Unterraum)

Unter einem (freien) Vektor im \mathbb{R}^n stellt man sich eine gerichtete Größe (einen Pfeil) vor (als typisches Beispiel eine Kraft im \mathbb{R}^3), die durch Länge und Richtung gekennzeichnet ist. Freie Vektoren, die durch Parallelverschiebung ineinander überführt werden können, werden als gleich angesehen.

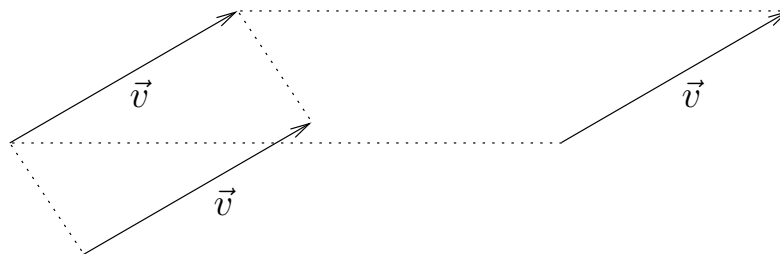


Abbildung 6.1: Parallelverschiebung freier Vektoren.

Dementsprechend kann ein Vektor so verschoben werden, dass sein Anfangspunkt im Koordinatenursprung, dem Nullpunkt, liegt. Damit ist ein Vektor im \mathbb{R}^n durch die Lage seines Endpunktes, d.h. durch die Koordinaten (oder Komponenten) seines Endpunktes, charakterisiert. Mit anderen Worten: Es handelt sich um ein Element \underline{x} des kartesischen Produktes

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}},$$

wobei \underline{x} in der Regel als Spaltenvektor (geordnetes n -Tupel) geschrieben

wird:

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x_1, x_2, \dots, x_n \in \mathbb{R}.$$

Die Komponenten werden geometrisch durch die **Projektionen auf die Koordinatenachsen** beschrieben.

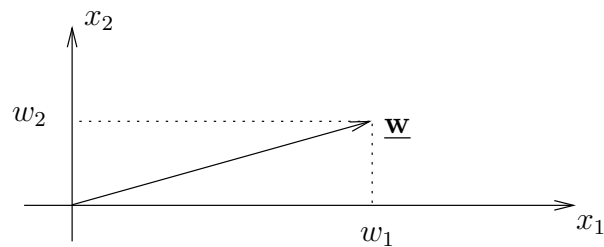


Abbildung 6.2: Der Vektor $\underline{\mathbf{w}}$ mit den Komponenten w_1, w_2 im \mathbb{R}^2 .

Anschaulich werden Vektoren in einem **Kräfteparallelogramm** addiert. Ebenso können Vektoren durch **Multiplikation mit einem Skalar** gestreckt werden.

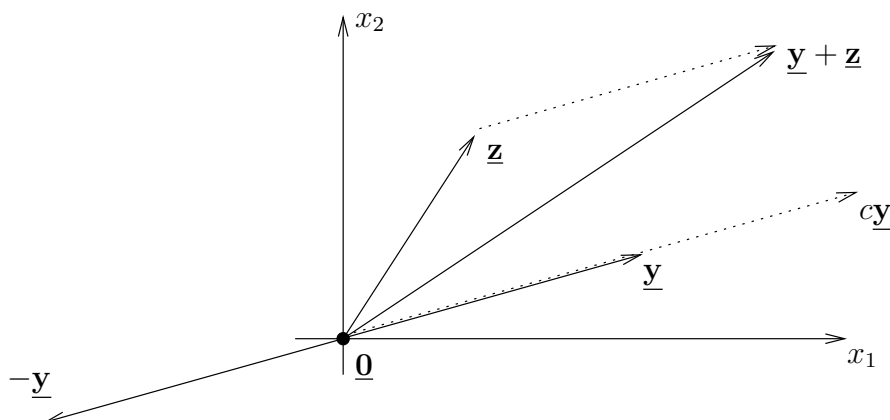


Abbildung 6.3: Ein Kräfteparallelogramm.

Es ist dabei ($\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n, c \in \mathbb{R}$)

$$c\underline{\mathbf{x}} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}, \quad -\underline{\mathbf{x}} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix},$$

$$\underline{\mathbf{x}} + \underline{\mathbf{y}} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \underline{\mathbf{0}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Die oben plausibel gemachte Struktur wird präzisiert in den [Vektorraumaxiomen](#):

Definition 6.1.1

*Es sei \mathbb{K} ein Körper und V eine Menge. Ferner existiere ein **Addition** $+$,*

$$+ : V \times V \rightarrow V, \\ (\underline{\mathbf{v}}, \underline{\mathbf{w}}) \mapsto \underline{\mathbf{v}} + \underline{\mathbf{w}}$$

*und eine **Multiplikation mit Skalaren** \cdot ,*

$$\cdot : \mathbb{K} \times V \rightarrow V, \\ (\lambda, \underline{\mathbf{v}}) \mapsto \lambda \underline{\mathbf{v}},$$

sodass gilt:

V1 $(V, +)$ ist eine **kommutative Gruppe** (mit neutralem Element $\underline{\mathbf{0}}$ und inverselem Element $-\underline{\mathbf{v}}$).

V2 Für das *multiplikative Einselement* des Körpers (neutrales Element bzgl. der Multiplikation im Körper) gilt

$$1 \underline{\mathbf{v}} = \underline{\mathbf{v}},$$

und es gelten die *Assoziativ-* und *Distributivgesetze* ($\lambda, \mu \in \mathbb{K}, \underline{\mathbf{v}}, \underline{\mathbf{w}} \in V$)

$$(\lambda \mu) \underline{\mathbf{v}} = \lambda (\mu \underline{\mathbf{v}}),$$

$$(\lambda + \mu) \underline{\mathbf{v}} = (\lambda \underline{\mathbf{v}}) + (\mu \underline{\mathbf{v}}),$$

$$\lambda (\underline{\mathbf{v}} + \underline{\mathbf{w}}) = (\lambda \underline{\mathbf{v}}) + (\lambda \underline{\mathbf{w}}).$$

Dann heißt $(V, +, \cdot)$ ein *Vektorraum* oder *linearer Raum* über dem Körper \mathbb{K} . Die *Elemente eines Vektorraums* heißen *Vektoren*.

Bemerkung. Mit dem Körper \mathbb{K} sind in diesem Kapitel stets die reellen Zahlen \mathbb{R} gemeint. Nach einer systematischen Einführung der komplexen Zahlen (Kapitel 7) ist aber klar, dass ebenso die Situation $\mathbb{K} = \mathbb{C}$ betrachtet werden kann.

Beispiele.

- i) Wie oben angedeutet, ist der \mathbb{R}^n ein \mathbb{R} -Vektorraum.
- ii) Zu den wichtigsten Beispielen linearer Räume gehören *Funktionsräume*. So ist beispielsweise der Raum der Polynome vom Grad $\leq n$, $n \in \mathbb{N}$, über \mathbb{R} (oder \mathbb{C}) ein Vektorraum:

$$\Pi_n(\mathbb{R}) := \left\{ p(x) = \sum_{k=0}^n a_k x^k : a_k \in \mathbb{R} \text{ für alle } k = 0, 1, \dots, n \right\}.$$

Funktionsräume (etwa der Raum der stetigen Funktionen) sind *i.A. nicht von endlicher Dimension* wie der \mathbb{R}^n .

Was ist die Dimension eines Vektorraumes?

Definition 6.1.2

Es sei V ein \mathbb{K} -Vektorraum, $m \in \mathbb{N}$ und $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(m)}$ seien Vektoren aus V .

i) Sind $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{K}$, so heißt die Summe

$$\sum_{k=1}^m \lambda_k \underline{\mathbf{v}}^{(k)}$$

eine **Linearkombination** der $\underline{\mathbf{v}}^{(k)}$. Die Linearkombination heißt **trivial**, wenn **alle** λ_k gleich Null sind. Andernfalls heißt sie **nicht-trivial**.

ii) Die Vektoren $\underline{\mathbf{v}}^{(k)}$, $k = 1, \dots, m$, heißen **linear unabhängig**, falls gilt:

$$\sum_{k=1}^m \lambda_k \underline{\mathbf{v}}^{(k)} = \underline{\mathbf{0}} \Rightarrow \lambda_k = 0 \text{ für alle } k = 1, 2, \dots, m.$$

Mit anderen Worten: Der Nullvektor ist durch die $\underline{\mathbf{v}}^{(k)}$ nur **trivial darstellbar**.

iii) Gibt es dagegen $\alpha_1, \alpha_2, \dots, \alpha_m \in \mathbb{K}$ – **nicht alle Null** – mit

$$\sum_{k=1}^m \alpha_k \underline{\mathbf{v}}^{(k)} = \underline{\mathbf{0}},$$

so nennt man die Vektoren $\underline{\mathbf{v}}^{(k)}$ **linear abhängig**.

Beispiele.

i) Betrachtet sei der \mathbb{R}^2 sowie die Vektoren

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Aus

$$1\underline{\mathbf{v}}^{(1)} + (-1)\underline{\mathbf{v}}^{(2)} + (-2)\underline{\mathbf{v}}^{(3)} = \underline{\mathbf{0}}$$

folgt unmittelbar, dass die Vektoren linear abhängig sind.

ii) Sind die Vektoren

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

im \mathbb{R}^3 gegeben, so gilt

$$\underline{\mathbf{0}} = \lambda_1 \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow 2\lambda_1 = 0, \quad \lambda_2 = 0 \Rightarrow \lambda_1 = \lambda_2 = 0,$$

was die lineare Unabhängigkeit der Vektoren beweist.

iii) Im Raum der Polynome vom Grad ≤ 2 sind

$$\underline{\mathbf{v}}^{(1)} := p(x) := 1 \quad \text{und} \quad \underline{\mathbf{v}}^{(2)} := q(x) := x \quad \text{für alle } x$$

linear unabhängig. Aus

$$\lambda_1 p(x) + \lambda_2 q(x) = \underline{\mathbf{0}} = (r(x) \equiv 0)$$

folgt nämlich

$$\lambda_1 + \lambda_2 x = 0 \quad \text{für alle } x,$$

also wieder

$$\lambda_1 = \lambda_2 = 0.$$

Wie behauptet sind die beiden Polynome linear unabhängig.

Bemerkung.

- i) Gleichbedeutend mit linearer Abhängigkeit von zwei und mehr Vektoren ist: Mindestens einer lässt sich als Linearkombination der anderen schreiben.
- ii) Ein einzelner Vektor $\underline{\mathbf{v}}$ ist per definitionem genau dann linear unabhängig, wenn $\underline{\mathbf{v}} \neq \underline{\mathbf{0}}$. Der Nullvektor $\underline{\mathbf{0}}$ ist linear abhängig.
- iii) Sind $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(k)}$, $k \in \mathbb{N}$, linear abhängig und nimmt man weitere Vektoren hinzu, so ist das erweiterte System erst recht linear abhängig.
- iv) Sind $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(k)}$, $k \in \mathbb{N}$, linear unabhängig und nimmt man Vektoren aus dieser Familie heraus, so sind die verbleibenden Vektoren linear unabhängig.

Definition 6.1.3

- i) Die *maximale Zahl linear unabhängiger Vektoren* eines Vektorraumes V heißt die *Dimension* des Vektorraumes, $\dim V$.
- ii) Es sei n die Dimension eines Vektorraumes V . Dann heißt jedes n -Tupel $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$ von linear unabhängigen Vektoren aus V eine *Basis* von V .

Bemerkung. Ist V ein Vektorraum mit Basis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$, so existieren zu jedem $\underline{\mathbf{v}} \in V$ **eindeutig bestimmte** Koeffizienten $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ mit

$$\underline{\mathbf{v}} = \sum_{k=1}^n \alpha_k \underline{\mathbf{v}}^{(k)}.$$

Diese Koeffizienten heißen **Koordinaten** von $\underline{\mathbf{v}}$ bzgl. der Basis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$.

Ist V ein \mathbb{K} -Vektorraum und $\underline{\mathbf{v}} \in V$, so ist $\{\lambda \underline{\mathbf{v}} : \lambda \in \mathbb{K}\}$ selbst wieder ein Vektorraum, der in V enthalten ist.

Es wird definiert:

Definition 6.1.4

*Ist V ein \mathbb{K} -Vektorraum und ist $U \subset V$ eine Teilmenge von V , die selbst ein Vektorraum ist, so heißt U **Unterraum** von V .*

Im Folgenden wird der (\mathbb{R} -) Vektorraum \mathbb{R}^n näher untersucht. Ist

$$\underline{\mathbf{e}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{e}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad \underline{\mathbf{e}}^{(n)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

so sind die Vektoren $\underline{\mathbf{e}}^{(i)}$ wegen

$$\underline{\mathbf{0}} = \sum_{i=1}^n \lambda_i \underline{\mathbf{e}}^{(i)} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

linear unabhängig, und es kann kein weiterer linear unabhängiger Vek-

tor hinzugefügt werden. Ist nämlich $\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ beliebig, so

gilt

$$1\underline{\mathbf{x}} - x_1\underline{\mathbf{e}}^{(1)} - x_2\underline{\mathbf{e}}^{(2)} - \dots - x_n\underline{\mathbf{e}}^{(n)} = \underline{\mathbf{0}}, \quad \text{bzw. } \underline{\mathbf{x}} = \sum_{i=1}^n x_i \underline{\mathbf{e}}^{(i)}.$$

Analog wird gezeigt, dass im \mathbb{R}^n nicht mehr als n Vektoren linear unabhängig sein können (warum?), es gilt $\dim \mathbb{R}^n = n$. Die Basis $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$ heißt die **kanonische Basis** oder die **Standardbasis** des \mathbb{R}^n .

Unterräume im \mathbb{R}^n kann man allgemein wie folgt erzeugen: Gegeben seien k Vektoren $\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)} \in \mathbb{R}^n$. Dann ist ihre **lineare Hülle**

$$L := \text{Spann}(\underline{\mathbf{x}}^{(1)}, \dots, \underline{\mathbf{x}}^{(k)}) := \left\{ \sum_{i=1}^k \alpha_i \underline{\mathbf{x}}^{(i)} : \alpha_i \in \mathbb{R}, i = 1, 2, \dots, k \right\}$$

ein Unterraum des \mathbb{R}^n (der von den $\underline{\mathbf{x}}^{(i)}$ **aufgespannte** Raum). Es gilt (wie allgemein in einem Unterraum):

$$\underline{\mathbf{0}} \in L, \quad 0 \leq \dim L \leq n,$$

und

$$\dim L = 0 \Leftrightarrow L = \{\underline{\mathbf{0}}\}, \quad \dim L = n \Leftrightarrow L = \mathbb{R}^n.$$

Beispiele.

$$i) \text{ Es sei } n = 3, \underline{\mathbf{v}} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ und}$$

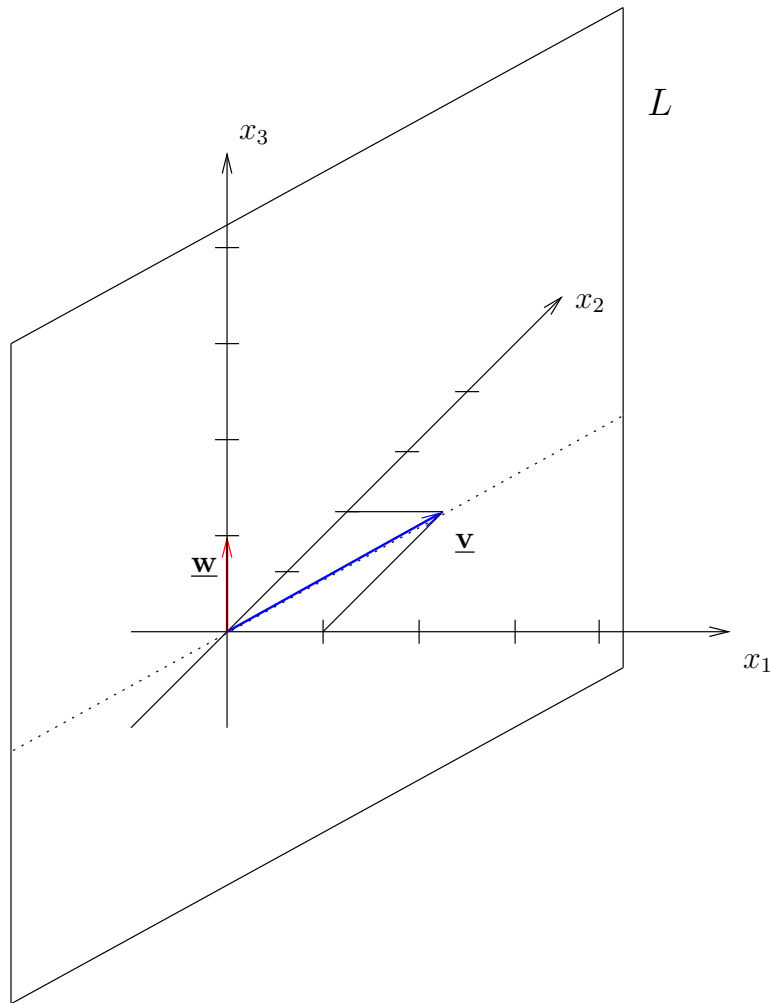
$$L = \{\alpha \underline{\mathbf{v}} + \beta \underline{\mathbf{w}} : \alpha, \beta \in \mathbb{R}\}.$$

Der Unterraum L ist die Ebene durch den Nullpunkt, die von den beiden Vektoren $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ aufgespannt wird (vgl. Abbildung 6.4).

ii) Es sei $n = 2$ und

$$L = \left\{ \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

In diesem Beispiel ist L die “ x_1 ”-Achse im \mathbb{R}^2 .

Abbildung 6.4: Die von \underline{v} und \underline{w} aufgespannte Ebene.

6.2 Die Geometrie des \mathbb{R}^n (Norm; Skalarprodukt; Cauchy-Schwarzsche Ungleichung; Kosinus; orthogonal; orthonormal; Kronecker-Symbol; Gram-Schmidtsches Orthogonalisierungsverfahren; orthogonale Projektion)

Im Vektorraum \mathbb{R}^n können (wie es der anschaulichen Vorstellung im \mathbb{R}^3 entspricht) **Längen** und **Winkel** gemessen werden. Dies geschieht mit der Hilfe der **Euklidischen Norm** bzw. des **Euklidischen Skalarproduktes**. Der Vektorraum \mathbb{R}^n versehen mit dem Euklidischen Skalarprodukt heißt der Euklidische Raum \mathbb{R}^n . Manchmal wird dies durch die Schreibweise \mathbb{E}^n verdeutlicht.

Zunächst wird hier zur Längenmessung der Begriff einer Norm eingeführt. Die (axiomatisch) geforderten Eigenschaften einer Norm entsprechen dabei genau den anschaulichen Erwartungen, die an eine Längenmessung zu stellen sind.

Definition 6.2.1

Es sei V ein Vektorraum über \mathbb{R} . Eine Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad \underline{\mathbf{x}} \mapsto \|\underline{\mathbf{x}}\|,$$

heißt **Norm** auf V , falls für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in V$ und für alle $\lambda \in \mathbb{R}$ gilt:

- i) $\|\underline{\mathbf{x}}\| \geq 0$ und $\|\underline{\mathbf{x}}\| = 0 \Leftrightarrow \underline{\mathbf{x}} = \underline{\mathbf{0}}$;
- ii) $\|\lambda \underline{\mathbf{x}}\| = |\lambda| \|\underline{\mathbf{x}}\|$;
- iii) (**Dreiecksungleichung**) $\|\underline{\mathbf{x}} + \underline{\mathbf{y}}\| \leq \|\underline{\mathbf{x}}\| + \|\underline{\mathbf{y}}\|$.

Bemerkung. In Abbildung 6.5 ist angedeutet, dass die Dreiecksungleichung der geometrischen Vorstellung entspricht.

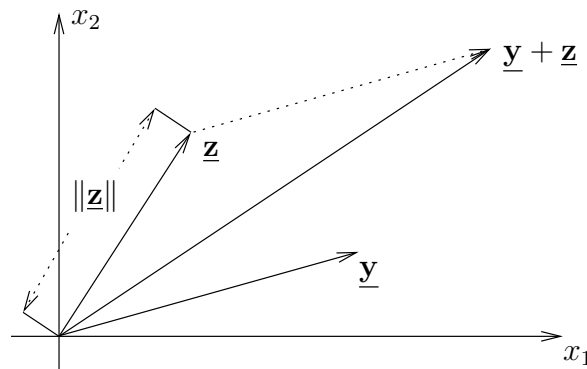


Abbildung 6.5: Klarerweise sollte gelten: $\|\underline{\mathbf{y}} + \underline{\mathbf{z}}\| \leq \|\underline{\mathbf{y}}\| + \|\underline{\mathbf{z}}\|$.

Beispiele. Im \mathbb{R}^n gibt es zahlreiche Normen wie etwa

$$\|\underline{\mathbf{x}}\|_\infty := \max\{|x_1|, |x_2|, \dots, |x_n|\} \quad (\text{Maximumsnorm});$$

$$\|\underline{\mathbf{x}}\|_2 := \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad (\text{Euklidische Norm});$$

$$\|\underline{\mathbf{x}}\|_1 := \sum_{i=1}^n |x_i| \quad (\text{Einsnorm}).$$

Eine besondere Bedeutung fällt dabei der Euklidischen Norm zu, die unsere Vorstellung im dreidimensionalen Raum widerspiegelt. Wesentlich ist, dass die Euklidische Norm von einem **Skalarprodukt induziert** ist:

Definition 6.2.2

Es sei V ein \mathbb{R} -Vektorraum. Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}, \quad (\underline{\mathbf{x}}, \underline{\mathbf{y}}) \mapsto \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle,$$

*heißt **Skalarprodukt** (oder **inneres Produkt**), wenn für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}} \in V$ und für alle $\lambda \in \mathbb{R}$ gilt:*

- i) (**Kommutativgesetz**) $\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \langle \underline{\mathbf{y}}, \underline{\mathbf{x}} \rangle$;*
- ii) (**Distributivgesetz**) $\langle \underline{\mathbf{x}} + \underline{\mathbf{y}}, \underline{\mathbf{z}} \rangle = \langle \underline{\mathbf{x}}, \underline{\mathbf{z}} \rangle + \langle \underline{\mathbf{y}}, \underline{\mathbf{z}} \rangle$;*
- iii) (**Assoziativgesetz**) $\lambda \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \langle \lambda \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \langle \underline{\mathbf{x}}, \lambda \underline{\mathbf{y}} \rangle$;*
- iv) (**positive Definitheit**) $\underline{\mathbf{x}} \neq \underline{\mathbf{0}} \Leftrightarrow \langle \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle > 0$.*

Bemerkung. Man beachte, dass das Skalarprodukt zweier Vektoren eine reelle Zahl ist.

Im \mathbb{R}^n existiert ein Skalarprodukt, wie folgender elementar zu beweisender Satz festhält.

Satz 6.2.1

Durch die Vorschrift

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle := \sum_{j=1}^n x_j y_j, \quad \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n,$$

wird im \mathbb{R}^n ein Skalarprodukt, genannt das *Euklidische* oder das *kanonische* oder das *Standardskalarprodukt* des \mathbb{R}^n , definiert.

Korollar 6.2.1

Bezeichnet $\langle \cdot, \cdot \rangle$ das Standardskalarprodukt des \mathbb{R}^n , so gilt für die Euklidische Norm oder den *Betrag* eines jeden Vektors $\underline{\mathbf{x}} \in \mathbb{R}^n$

$$\|\underline{\mathbf{x}}\|_2 = \langle \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle^{\frac{1}{2}}.$$

Vereinbarung. Wenn es nicht ausdrücklich anders betont ist, *stehen* $\langle \cdot, \cdot \rangle$ und $\|\cdot\|$ im Folgenden stets für das Standardskalarprodukt und die daraus induzierte Euklidische Norm.

Ein wesentliches Hilfsmittel zur weiteren Diskussion der Geometrie des \mathbb{R}^n ist

Satz 6.2.2 (*Cauchy-Schwarzsche Ungleichung*)¹

Für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$ gilt

$$|\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle| \leq \|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\|.$$

¹H.A. Schwarz, 1843-1921; Halle, Zürich, Göttingen, Berlin.

Beweis. Es seien $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$ (o.E. $\underline{\mathbf{x}}, \underline{\mathbf{y}} \neq \underline{\mathbf{0}}$). Dann gilt für beliebige $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} & \langle \alpha \underline{\mathbf{x}} + \beta \underline{\mathbf{y}}, \alpha \underline{\mathbf{x}} + \beta \underline{\mathbf{y}} \rangle \geq 0 \\ \Rightarrow & \alpha^2 \langle \underline{\mathbf{x}}, \underline{\mathbf{x}} \rangle + \beta^2 \langle \underline{\mathbf{y}}, \underline{\mathbf{y}} \rangle + 2\alpha\beta \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \geq 0 \\ \Rightarrow & \alpha^2 \|\underline{\mathbf{x}}\|^2 + \beta^2 \|\underline{\mathbf{y}}\|^2 \geq -2\alpha\beta \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle . \end{aligned}$$

Wird nun $\alpha := \|\underline{\mathbf{y}}\|$ und $\beta := \pm \|\underline{\mathbf{x}}\|$ gewählt, so ergibt sich

$$\begin{aligned} & 2\|\underline{\mathbf{x}}\|^2 \|\underline{\mathbf{y}}\|^2 \geq \pm 2\|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\| \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \\ \Rightarrow & \|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\| \geq \pm \langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle , \end{aligned}$$

und die Behauptung ist gezeigt. \square

Folgerung. Die Cauchy-Schwarzsche Ungleichung impliziert für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$:

$$\begin{aligned} \|\underline{\mathbf{x}} + \underline{\mathbf{y}}\|^2 &= \langle \underline{\mathbf{x}} + \underline{\mathbf{y}}, \underline{\mathbf{x}} + \underline{\mathbf{y}} \rangle = \|\underline{\mathbf{x}}\|^2 + \|\underline{\mathbf{y}}\|^2 + 2\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle \\ &\leq \|\underline{\mathbf{x}}\|^2 + \|\underline{\mathbf{y}}\|^2 + 2\|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\| = (\|\underline{\mathbf{x}}\| + \|\underline{\mathbf{y}}\|)^2 . \end{aligned}$$

Damit ist die Dreiecksungleichung für die Euklidische Norm verifiziert und Satz 6.2.1 ist bewiesen. \square

Durch das Skalarprodukt sind **Winkelmessungen** überhaupt erst möglich. Sind $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$, so **definiert man** den **Kosinus zwischen zwei Vektoren**

$$\cos(\underline{\mathbf{x}}, \underline{\mathbf{y}}) := \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle}{\|\underline{\mathbf{x}}\| \|\underline{\mathbf{y}}\|} .$$

Dabei impliziert die Cauchy-Schwarzsche Ungleichung

$$-1 \leq \cos(\underline{\mathbf{x}}, \underline{\mathbf{y}}) \leq 1 .$$

Bemerkung. Auf diese Art kann in beliebigen Vektorräumen, die mit einem Skalarprodukt versehen sind, ein Winkel zwischen zwei Vektoren erklärt werden.

Nun ist zu überprüfen, ob die Definition nicht im Widerspruch steht zur Definition des **Kosinus** über elementargeometrische Betrachtungen. Um einzusehen, dass das in der Tat nicht der Fall ist, sei ein rechtwinkliges Dreieck bzw. der Einheitskreis in der Ebene \mathbb{R}^2 betrachtet (wie in Abbildung 6.6 dargestellt).

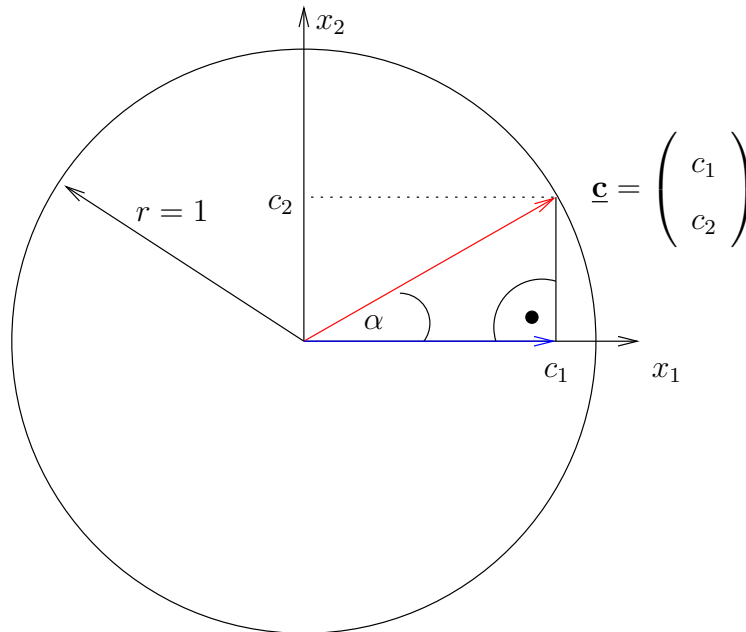


Abbildung 6.6: Elementargeometrische Betrachtungen zum Kosinus.

Der Radius des Kreises ist 1, also impliziert der Satz des Pythagoras $1 = c_1^2 + c_2^2$.

Der Kosinus des Winkels α wird über den Einheitskreis definiert als (“Ankathete durch Hypotenuse”, analog entspricht der Sinus dem Längenverhältnis “Gegenkathete durch Hypotenuse”)

$$\cos(\alpha) := c_1 .$$

Nach der “neuen” Definition des Winkels muss das Skalarprodukt des **blau** gekennzeichneten Vektors $\underline{\mathbf{x}} = \begin{pmatrix} c_1 \\ 0 \end{pmatrix}$ und des **rot** dargestellten Vektors $\underline{\mathbf{c}} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ berechnet werden:

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{c}} \rangle = c_1^2 + 0c_2 = c_1^2 .$$

Zusammen mit den Euklidischen Normen

$$\|\underline{\mathbf{x}}\| = \sqrt{c_1^2 + 0^2} = c_1, \quad \|\underline{\mathbf{c}}\| = \sqrt{c_1^2 + c_2^2} = \sqrt{1} = 1.$$

folgt in Übereinstimmung mit der “alten” Definition

$$\cos(\underline{\mathbf{x}}, \underline{\mathbf{c}}) = \frac{\langle \underline{\mathbf{x}}, \underline{\mathbf{c}} \rangle}{\|\underline{\mathbf{x}}\| \|\underline{\mathbf{c}}\|} = \frac{c_1^2}{c_1} = c_1.$$

Insbesondere ist über das Skalarprodukt definiert, was es heißt, dass zwei Vektoren **senkrecht** aufeinander stehen:

Definition 6.2.3

i) Zwei Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$ stehen senkrecht aufeinander (oder sind orthogonal), falls

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = 0.$$

ii) Ist $U \subset \mathbb{R}^n$ eine beliebige Menge, so heißt

$$U^\perp := \{ \underline{\mathbf{x}} \in \mathbb{R}^n : \langle \underline{\mathbf{x}}, \underline{\mathbf{u}} \rangle = 0 \text{ für alle } \underline{\mathbf{u}} \in U \}$$

das orthogonale Komplement von U .

Bemerkung.

i) Die Orthogonalität in Funktionenräumen, die mit einem Skalarprodukt versehen sind, spielt im Kapitel 26 eine wesentliche Rolle.

ii) U^\perp ist stets ein Unterraum, auch wenn U selbst kein Unterraum ist.

Beispiele.

i) Es sei $n = 3$,

$$\underline{\mathbf{N}} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \text{ und } U = \{ \underline{\mathbf{N}} \}.$$

Dann ist

$$U^\perp = \left\{ \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 + x_2 + 2x_3 = 0 \right\}.$$

U^\perp ist die Ebene durch den Nullpunkt, die auf dem Vektor $\underline{\mathbf{N}}$ senkrecht steht. In U^\perp liegen beispielsweise die Vektoren

$$\underline{\mathbf{v}} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Diese sind linear unabhängig und spannen die Ebene U^\perp auf (vgl. Abbildung 6.7).

ii) Betrachtet sei die Standardbasis $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$ des \mathbb{R}^n . Für alle $i, j = 1, 2, \dots, n$ ist

$$\langle \underline{\mathbf{e}}^{(i)}, \underline{\mathbf{e}}^{(j)} \rangle = \delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Dabei heißt δ_{ij} das **Kronecker-Symbol**². Die Basisvektoren haben die Länge 1 und stehen senkrecht aufeinander.

Umgekehrt gilt

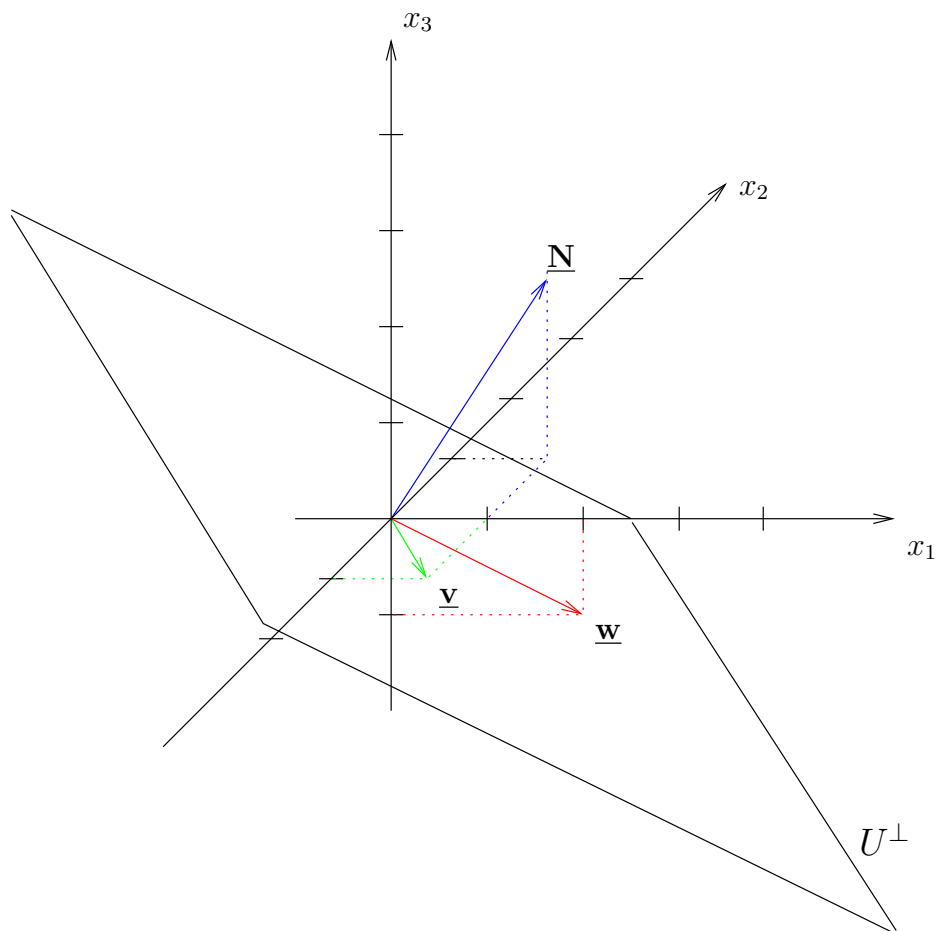
Satz 6.2.3

Es seien $\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(n)} \in \mathbb{R}^n$ und für alle $i, j = 1, 2, \dots, n$ gelte

$$\langle \underline{\mathbf{f}}^{(i)}, \underline{\mathbf{f}}^{(j)} \rangle = \delta_{ij}. \quad (*)$$

Dann ist $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(n)})$ eine **Basis** des \mathbb{R}^n . Jede Basis mit der Eigenschaft (*) heißt **Orthonormalbasis (ONB)** des \mathbb{R}^n .

²L. Kronecker, 1821-1891; Berlin, Breslau, Bonn.

Abbildung 6.7: Das orthogonale Komplement des Vektors \underline{N} .

Beweis. Multipliziere $\sum_{k=1}^n \lambda_k \underline{\mathbf{f}}^{(k)} = \underline{\mathbf{0}}$ skalar mit $\underline{\mathbf{f}}^{(i)}$. □

Bemerkung. Analog wird argumentiert: Es sei eine ONB nach Satz 6.2.3 gegeben. Alle $\underline{\mathbf{x}} \in \mathbb{R}^n$ können insbesondere dargestellt werden als $\underline{\mathbf{x}} = \sum_{i=1}^n \alpha_i \underline{\mathbf{f}}^{(i)}$, $\alpha_i \in \mathbb{R}$, $i = 1, \dots, n$.

Ist $1 \leq k \leq n$ und multipliziert man $\underline{\mathbf{x}}$ skalar mit $\underline{\mathbf{f}}^{(k)}$, so folgt aufgrund der Orthonormalität $\alpha_k = \langle \underline{\mathbf{x}}, \underline{\mathbf{f}}^{(k)} \rangle$, d.h.

$$\underline{\mathbf{x}} = \sum_{k=1}^n \langle \underline{\mathbf{x}}, \underline{\mathbf{f}}^{(k)} \rangle \underline{\mathbf{f}}^{(k)} .$$

Problem. Es seien $\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)}, \dots, \underline{\mathbf{g}}^{(m)}$, $1 \leq m \leq n$, linear unabhängige Vektoren, die den Unterraum U aufspannen. In U ist es aber viel ein-

facher, mit einer ONB zu arbeiten als mit den Vektoren $\underline{\mathbf{g}}^{(i)}$, $i = 1, 2, \dots, m$. Es stellt sich die Frage:

Wie kann eine ONB von U konstruiert werden?

Ein **konstruktives Verfahren** dazu ist das **Schmidtsche**³ oder auch **Gram-Schmidtsche**⁴ **Orthogonalisierungsverfahren**. Die **Idee** ist, von einem Basisvektor die **orthogonale Projektion** in den von den vorherigen Basisvektoren erzeugten Teilraum abzuziehen.

Beispiel. Es sei $n = 2$,

$$\underline{\mathbf{g}}^{(1)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{g}}^{(2)} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Um eine ONB des aufgespannten Raumes (in diesem Fall des ganzen \mathbb{R}^2) zu finden, wird zunächst der Vektor $\underline{\mathbf{g}}^{(1)}$ auf die Länge 1 **normiert**:

$$\underline{\mathbf{f}}^{(1)} := \frac{1}{\|\underline{\mathbf{g}}^{(1)}\|} \underline{\mathbf{g}}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Anschließend wird aus $\underline{\mathbf{g}}^{(2)}$ ein zu $\underline{\mathbf{f}}^{(1)}$ senkrechter Vektor konstruiert. Dabei zieht man von $\underline{\mathbf{g}}^{(2)}$ den Anteil ab, der schon im von $\underline{\mathbf{f}}^{(1)}$ aufgespannten Raum liegt, d.h. die orthogonale Projektion von $\underline{\mathbf{g}}^{(2)}$ auf den von $\underline{\mathbf{f}}^{(1)}$ aufgespannten Teilraum.

Zu bemerken ist, dass ein Anteil in die orthogonale Richtung verbleiben muss, da $\underline{\mathbf{g}}^{(1)}$ und $\underline{\mathbf{g}}^{(2)}$ linear unabhängig gewählt waren.

Die orthogonale Projektion von $\underline{\mathbf{g}}^{(2)}$ auf den von $\underline{\mathbf{f}}^{(1)}$ aufgespannten Teilraum ist aber (vgl. Abbildung 6.8, dort rot dargestellt)

$$\langle \underline{\mathbf{g}}^{(2)}, \underline{\mathbf{f}}^{(1)} \rangle \underline{\mathbf{f}}^{(1)},$$

³E. Schmidt, 1876-1959; Zürich, Erlangen, Breslau, Berlin.

⁴J.P. Gram, 1850-1916; Kopenhagen.

es verbleibt (in der Abbildung grün dargestellt):

$$\underline{\mathbf{g}}^{(2)} - \langle \underline{\mathbf{g}}^{(2)}, \underline{\mathbf{f}}^{(1)} \rangle \underline{\mathbf{f}}^{(1)} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} - \left\langle \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

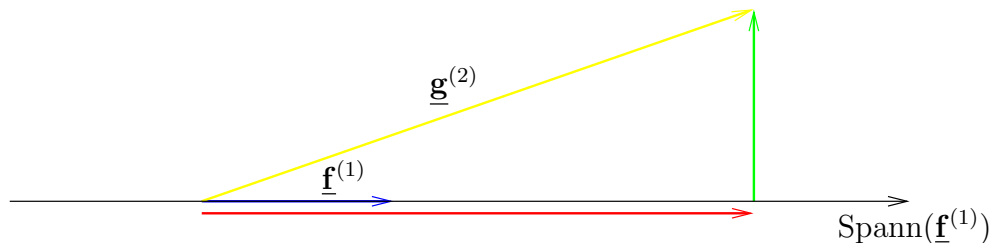


Abbildung 6.8: Zur orthogonalen Projektion von $\underline{\mathbf{g}}^{(2)}$ auf $\text{Spann}(\underline{\mathbf{f}}^{(1)})$.

Dieser hat in dem speziellen Beispiel bereits die Länge 1, ansonsten ist er noch auf die Länge 1 zu normieren.

Im allgemeinen Fall wird wie folgt vorgegangen: Es sei $1 \leq m \leq n$ und $\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)}, \dots, \underline{\mathbf{g}}^{(m)}$ seien linear unabhängige Vektoren im \mathbb{R}^n .

i) Zunächst wird gesetzt:

$$\underline{\mathbf{f}}^{(1)} := \frac{1}{\|\underline{\mathbf{g}}^{(1)}\|} \underline{\mathbf{g}}^{(1)} .$$

ii) Anschließend betrachtet man den Vektor

$$\tilde{\underline{\mathbf{f}}}^{(2)} := \underline{\mathbf{g}}^{(2)} - \langle \underline{\mathbf{g}}^{(2)}, \underline{\mathbf{f}}^{(1)} \rangle \underline{\mathbf{f}}^{(1)}$$

und normiert diesen auf die Länge 1, d.h.

$$\underline{\mathbf{f}}^{(2)} := \frac{1}{\|\tilde{\underline{\mathbf{f}}}^{(2)}\|} \tilde{\underline{\mathbf{f}}}^{(2)} .$$

iii) Analog wird

$$\tilde{\underline{\mathbf{f}}}^{(3)} := \underline{\mathbf{g}}^{(3)} - \langle \underline{\mathbf{g}}^{(3)}, \underline{\mathbf{f}}^{(1)} \rangle \underline{\mathbf{f}}^{(1)} - \langle \underline{\mathbf{g}}^{(3)}, \underline{\mathbf{f}}^{(2)} \rangle \underline{\mathbf{f}}^{(2)}$$

gesetzt und anschließend normiert:

$$\underline{\mathbf{f}}^{(3)} := \frac{1}{\|\tilde{\underline{\mathbf{f}}}^{(3)}\|} \tilde{\underline{\mathbf{f}}}^{(3)} .$$

iv) Sind nun für ein $1 \leq k \leq m$ die Vektoren $\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \dots, \underline{\mathbf{f}}^{(k-1)}$ wie oben konstruiert, so wird

$$\tilde{\underline{\mathbf{f}}}^{(k)} := \underline{\mathbf{g}}^{(k)} - \sum_{i=1}^{k-1} \langle \underline{\mathbf{g}}^{(k)}, \underline{\mathbf{f}}^{(i)} \rangle \underline{\mathbf{f}}^{(i)}$$

wieder normiert:

$$\underline{\mathbf{f}}^{(k)} := \frac{1}{\|\tilde{\underline{\mathbf{f}}}^{(k)}\|} \tilde{\underline{\mathbf{f}}}^{(k)}.$$

v) Im letzten Schritt ist ein Orthonormalsystem aus m Vektoren entstanden.

Beispiel. Es sei U im \mathbb{R}^3 der Unterraum $\text{Spann}(\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)})$,

$$\underline{\mathbf{g}}^{(1)} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \text{und} \quad \underline{\mathbf{g}}^{(2)} = \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}$$

Man wählt nach obigem Rezept zunächst

$$\underline{\mathbf{f}}^{(1)} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Anschließend wird gesetzt:

$$\begin{aligned} \tilde{\underline{\mathbf{f}}}^{(2)} &= \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix} - \left\langle \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\rangle \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}. \end{aligned}$$

Nach der Normierung von $\tilde{\mathbf{f}}^{(2)}$ ist eine ONB von U gefunden:

$$\frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

6.3 Folgen im \mathbb{R}^n (Übertragung des Konvergenzbegriffes)

In Definition 4.1.1 ist bereits der Begriff einer Folge $\{\mathbf{x}^{(k)}\}$ von Vektoren $\mathbf{x}^{(k)}$ im \mathbb{R}^n enthalten. Man stelle sich beispielsweise die Folge $\{\mathbf{x}(t_k)\}$ der Ortskoordinaten eines Massenpunktes im \mathbb{R}^3 zu einer Folge von Zeitpunkten $\{t_k\}$ vor.

Der Konvergenzbegriff für eine reelle Zahlenfolge ist in Abbildung 4.2 veranschaulicht. Die natürliche Vorstellung im höher dimensionalen Fall ist in Abbildung 6.9 dargestellt.

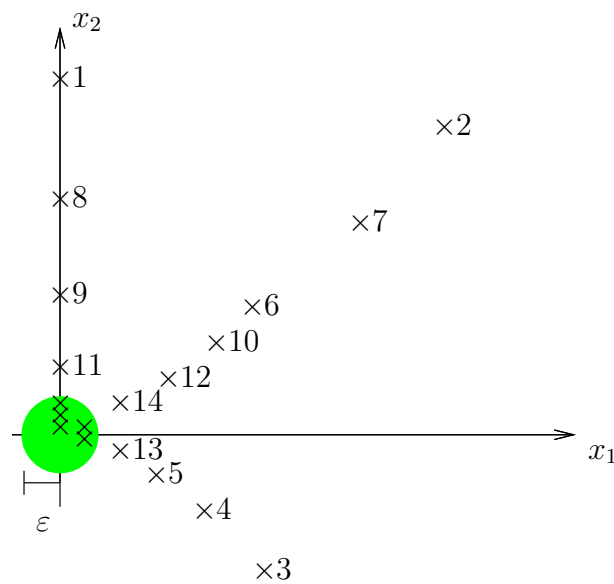


Abbildung 6.9: Für dieses ε sind ab $k = 15$ alle Folgenglieder näher bei $\mathbf{0}$ als ε .

“Näher bei der $\mathbf{0}$ ” (analog bei $\mathbf{a} \in \mathbb{R}^n$) ist für reelle Zahlen durch die Betragsfunktion $|\cdot|$ bestimmt, die den Abstand in \mathbb{R} misst. Im \mathbb{R}^n wird der Abstand durch die Euklidische Norm gemessen (andere Normen werden zur Zeit nicht betrachtet), also ist in der Definition

4.1.2 einer konvergenten reellen Zahlenfolge einfach **der Betrag durch die Euklidische Norm zu ersetzen**.

Dies führt auf

Definition 6.3.1

Es sei $n \geq 1$. Eine Folge $\{\underline{\mathbf{x}}^{(k)}\}$, $\underline{\mathbf{x}}^{(k)} \in \mathbb{R}^n$ für alle $k \in \mathbb{N}$, heißt **konvergent**, wenn es ein $\underline{\mathbf{x}} \in \mathbb{R}^n$ gibt, sodass gilt: Für jedes $\varepsilon > 0$ existiert eine (**von ε abhängende**) Zahl $N(\varepsilon) \in \mathbb{N}$ mit

$$\|\underline{\mathbf{x}}^{(k)} - \underline{\mathbf{x}}\| < \varepsilon \text{ für alle } k \geq N(\varepsilon).$$

Falls eine Folge konvergiert, so heißt $\underline{\mathbf{x}}$ der **Grenzwert** der Folge (die Folge **konvergiert gegen $\underline{\mathbf{x}}$** ; die Folge hat den **Grenzwert** oder **Limes $\underline{\mathbf{x}}$**).

Eine Folge, die nicht konvergiert, heißt **divergent**.

Bemerkungen.

i) Die Notation bleibt die Gleiche, d.h. man schreibt

$$\lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{x}} \quad \text{oder} \quad \underline{\mathbf{x}}^{(k)} \rightarrow \underline{\mathbf{x}} \text{ für } k \rightarrow \infty.$$

ii) Beispiele ergeben sich unmittelbar aus dem folgenden Satz, der den Zusammenhang mit reellen Zahlenfolge wiedergibt.

Satz 6.3.1

Es seien $\underline{\mathbf{x}}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}$, $k \in \mathbb{N}$, und $\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. Die

Folge $\{\underline{\mathbf{x}}^{(k)}\}$ ist **genau dann** konvergent gegen $\underline{\mathbf{x}}$, falls für jedes feste i , $i = 1, \dots, n$, gilt: $\lim_{k \rightarrow \infty} x_i^{(k)} = x_i$.

Bemerkungen.

- i)* Mit anderen Worten: $\lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{x}}$ genau dann, wenn **komponentenweise** Konvergenz (**für alle Komponenten!**) vorliegt (siehe Übungen).
- ii)* Nach *i)* können Grenzwertbetrachtungen von Folgen $\{\underline{\mathbf{x}}^{(k)}\}$ im \mathbb{R}^n komplett auf Paragraph 4.1 zurückgeführt werden: Es muss jede Komponente einzeln diskutiert werden.

Beispiele.

- i)* Es sei $n = 3$

$$\underline{\mathbf{x}}^{(k)} = \begin{pmatrix} \frac{2k^3+1}{k^3+2k-1} \\ \left(1 + \frac{1}{k}\right)^k \\ \frac{k!}{k^k} \end{pmatrix} \quad \text{für alle } k \in \mathbb{N} .$$

Die Folge $\{\underline{\mathbf{x}}^{(k)}\}$ konvergiert gegen $\underline{\mathbf{x}} = \begin{pmatrix} 2 \\ e \\ 0 \end{pmatrix}$.

- ii)* Ist für $n = 2$ und für alle $k \in \mathbb{N}$

$$\underline{\mathbf{x}}^{(k)} = \begin{pmatrix} 1/k \\ k \end{pmatrix} ,$$

so divergiert die Folge $\{\underline{\mathbf{x}}^{(k)}\}$.

6.4 Die Topologie des \mathbb{R}^n (beschränkte, offene, abgeschlossene und kompakte Mengen im \mathbb{R}^n)

Analysiert man Funktionen einer Veränderlichen (der Zeit), so ist der Definitionsbereich meist ein beschränktes/offenes/abgeschlossenes Intervall. Die essentielle Bedeutung der verschiedenen Arten von Intervallen als Definitionsbereich ist in Abbildung 3.1 verdeutlicht.

Im \mathbb{R}^n ist die Situation wesentlich vielfältiger, und es ist nicht unmittelbar evident, welche verschiedenen Eigenschaften einer Menge eine “natürliche Bedeutung” haben.

Da im \mathbb{R}^n **keine Ordnungsrelation** “ $<$ ” (im Sinne der Axiomatik aus Paragraph 2.3) existiert, ergeben die Begriffe “nach **oben** beschränkt” und “nach **unten** beschränkt” keinen Sinn

Die **Beschränktheit einer Menge im \mathbb{R}^n** kann aber mit Hilfe der Euklidischen Norm definiert werden. Die Vorstellung dabei ist, dass eine Menge in einer Kreisscheibe (im \mathbb{R}^2) oder einer Kugel (im \mathbb{R}^n) enthalten ist (zur Erinnerung: Eine beschränkte Menge reeller Zahlen ist per definitionem in einem Intervall $[-k, k]$ enthalten).

Definition 6.4.1

*Eine Menge U von Elementen des \mathbb{R}^n , $U \subset \mathbb{R}^n$, heißt **beschränkt**, wenn es eine reelle Zahl $M \in \mathbb{R}$ gibt mit*

$$\|\underline{\mathbf{x}}\| \leq M \quad \text{für alle } \underline{\mathbf{x}} \in U .$$

Beispiele.

- i) Die in Abbildung 6.10 dargestellte Menge $U \subset \mathbb{R}^2$ ist beschränkt, sie ist enthalten in einer Kreisscheibe vom Radius M .
- ii) Die Menge $U = \mathbb{R} \times [0, 1] \subset \mathbb{R}^2$ ist nicht beschränkt, da es zu beliebigem $M \in \mathbb{R}$ ein $\underline{\mathbf{x}} \in U$ gibt mit $\|\underline{\mathbf{x}}\| > M$, z.B.

$$\underline{\mathbf{x}} = \begin{pmatrix} M \\ 1/2 \end{pmatrix} .$$

Die wichtigsten Mengen sind (als Verallgemeinerung offener Intervalle in \mathbb{R}) offene⁵ Kreisscheiben (im \mathbb{R}^2) und offene Kugeln (im \mathbb{R}^n), hier auch als **Umgebungen** bezeichnet.

⁵Der Begriff “offen” wird in Definition 6.4.4 präzisiert.

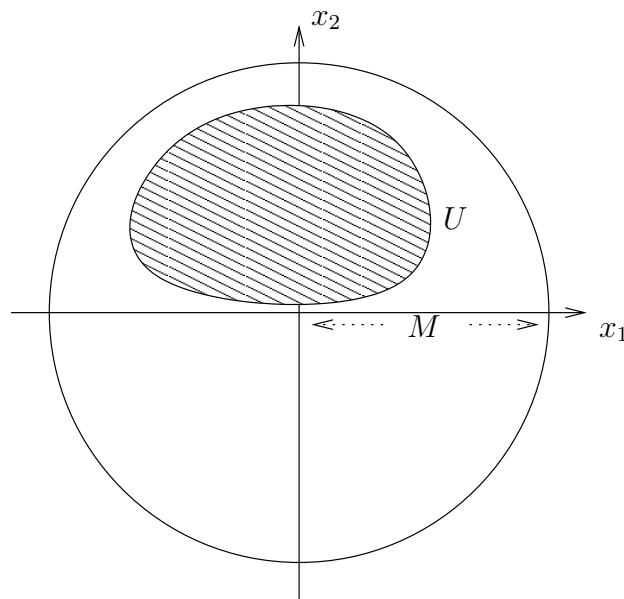


Abbildung 6.10: Ein beschränkte Menge.

Definition 6.4.2

Sind $\underline{\mathbf{x}}^{(0)} \in \mathbb{R}^n$ und $\varepsilon > 0$ gegeben, so heißt die Menge

$$B_\varepsilon(\underline{\mathbf{x}}^{(0)}) = \{\underline{\mathbf{x}} \in \mathbb{R}^n : \|\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}\| < \varepsilon\}$$

eine ε -Umgebung von $\underline{\mathbf{x}}^{(0)}$.

Beispiele.

- i) Im \mathbb{R}^1 ist $B_\varepsilon(x^{(0)})$ das offene Intervall $(x^{(0)} - \varepsilon, x^{(0)} + \varepsilon)$.
- ii) Im \mathbb{R}^2 ist $B_\varepsilon(\underline{\mathbf{x}}^{(0)})$ eine offene Kreisscheibe um den Punkt $\underline{\mathbf{x}}^{(0)}$ vom Radius ε .
- iii) Analog ist $B_\varepsilon(\underline{\mathbf{x}}^{(0)})$ im \mathbb{R}^3 eine offene Kugel um den Punkt $\underline{\mathbf{x}}^{(0)}$ vom Radius ε .

Mit Hilfe des Umgebungsbegriffes kann definiert werden: (Beispiel: Betrachtet man in Abbildung 6.11 die Vereinigung der offenen Kreisscheibe mit der rot dargestellten Kreislinie, so sind nach wie vor die **Randpunkte** zu unterscheiden von den blau dargestellten **inneren Punkten**.)

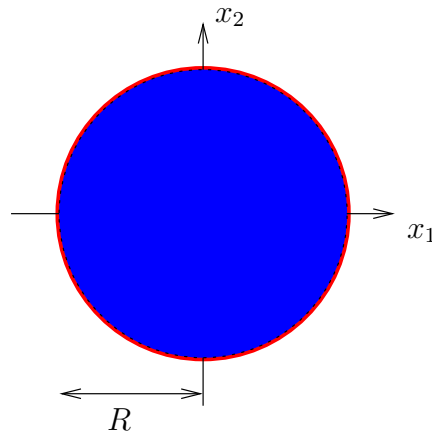


Abbildung 6.11: Die blau dargestellte Menge ist die offene Kreisscheibe vom Radius R um den Nullpunkt. Die rote Linie (der Rand) gehört nicht dazu.

Definition 6.4.3

- i) Ein Punkt $\underline{x} \in \mathbb{R}^n$ heißt **innerer Punkt** einer Teilmenge U des \mathbb{R}^n , wenn es eine Umgebung von \underline{x} gibt, die ganz in U enthalten ist.*
- ii) Ein Punkt $\underline{x} \in \mathbb{R}^n$ heißt **Häufungspunkt** einer Teilmenge U des \mathbb{R}^n , wenn in jeder Umgebung von \underline{x} ein Punkt von U liegt, der von \underline{x} verschieden ist (vgl. mit dem Begriff “Häufungspunkt einer Folge”).*
- iii) Ein Punkt $\underline{x} \in \mathbb{R}^n$ heißt **Randpunkt** einer Teilmenge U des \mathbb{R}^n , wenn in jeder Umgebung von \underline{x} mindestens ein Punkt von U und ein Punkt des Komplementes von U liegt.*

Bemerkung. Häufungspunkte und Randpunkte müssen selbst nicht zur Menge gehören.

Beispiele.

- i) Es sei $n = 1$, $a < b$ und $U = (a, b] \subset \mathbb{R}$.*

- Die Menge aller inneren Punkte ist in diesem Beispiel

$$\{x : a < x < b\} .$$

Ist nämlich $a < x < b$ und $\varepsilon < \min\{|x - a|, |x - b|\}$, so gilt $(x - \varepsilon, x + \varepsilon) \subset (a, b]$, es liegt also eine (und damit beliebig viele) ε -Umgebung von x komplett in $(a, b]$. Die Punkte a und b (und alle Punkte, die nicht zu dem Intervall gehören) können keine inneren Punkte sein (vgl. Abbildung 6.12).

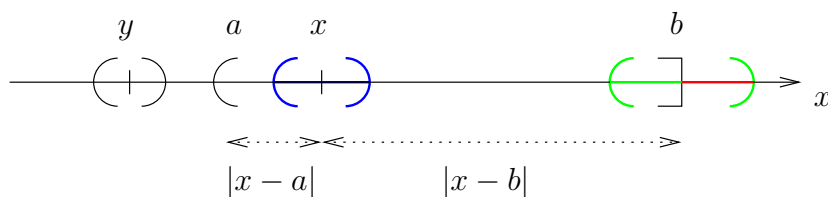


Abbildung 6.12: Das blau gekennzeichnete “ ε -Intervall” um x liegt ganz in $(a, b]$. Um b kann es kein solches Intervall geben, da die obere Grenze größer als b ist und die rot gekennzeichneten Punkte nicht zu $(a, b]$ gehören.

- Anhand von Abbildung 6.12 erkennt man ebenfalls, dass alle $a \leq x \leq b$ Häufungspunkte des Intervalls $(a, b]$ sind. Zu beachten ist: a gehört zwar nicht zur Menge U , ist aber trotzdem Häufungspunkt der Menge.

Ist $y < a$ oder $y > b$, so kann y kein Häufungspunkt sein. Die in der Abbildung angedeutete Umgebung von y enthält keinen Punkt der Menge.

- Schließlich sieht man, dass a und b die Randpunkte des Intervalls (im Sinne der Definition 6.4.3) sind.

ii) Analog sind die inneren Punkte der offenen Kreisscheibe die blau gekennzeichneten Punkte aus Abbildung 6.11, die Randpunkte bilden die roten Kreislinie, die Menge der Häufungspunkte ist die Vereinigung der inneren Punkte und der Randpunkte (vgl. auch die Menge U aus Abbildung 6.13).

iii) In Abbildung 6.9 ist eine Menge angedeutet, die nur $\mathbf{0}$ als Häufungspunkt hat.

Definition 6.4.3 führt schließlich zu

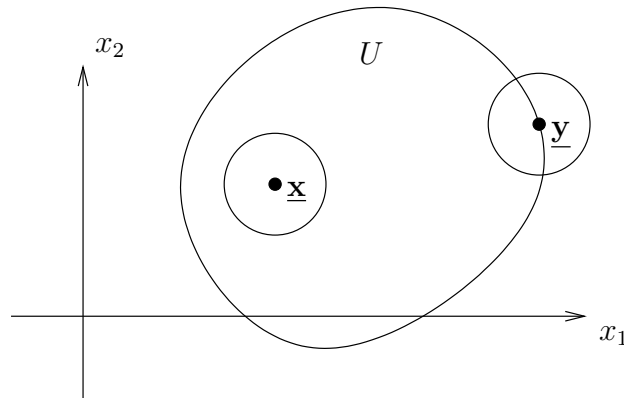


Abbildung 6.13: Hier ist \underline{x} ein innerer Punkt, \underline{y} ein Randpunkt von U .

Definition 6.4.4

- i) Eine Menge $U \subset \mathbb{R}^n$, die nur aus inneren Punkten besteht, heißt **offen**. Die Menge aller inneren Punkte heißt das **Innere** (oder der **offene Kern**) $\overset{\circ}{U}$.
- ii) Eine Menge $U \subset \mathbb{R}^n$, die alle ihre Häufungspunkte enthält, heißt **abgeschlossen**. Die Vereinigung einer Menge mit der Menge aller ihrer Häufungspunkte heißt der **Abschluß** (oder die **abgeschlossenen Hülle**) \overline{U} , der **Rand** ∂U einer Menge $U \subset \mathbb{R}^n$ ist die Menge $\overline{U} - \overset{\circ}{U}$.
- iii) Eine Menge $U \subset \mathbb{R}^n$, die abgeschlossen und beschränkt ist, heißt **kompakt**.

Beispiele.

- i) $U = \{\underline{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 < R^2\}$ (vgl. Abbildung 6.11): U ist beschränkt, nicht kompakt, offen, nicht abgeschlossen. Innere Punkte und Randpunkte sind bereits oben spezifiziert worden.
- ii) $U = \{\underline{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq R^2\}$ (vgl. Abbildung 6.14): U ist beschränkt, kompakt, nicht offen, abgeschlossen. Die inneren Punk-

te und die Randpunkte sind die gleichen wie im Fall der offenen Kreisscheibe.

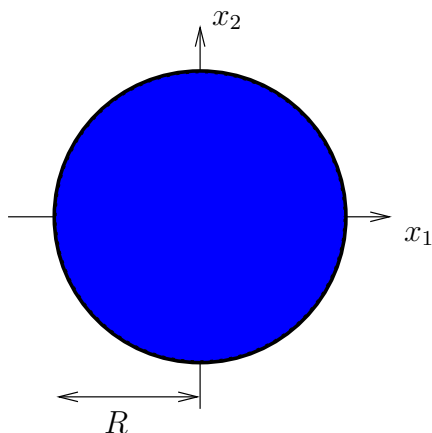


Abbildung 6.14: Die dunkler gekennzeichneten Punkte gehören zur abgeschlossenen Kreisscheibe vom Radius R .

iii) $U = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1^2 + x_2^2 = R^2\}$ (vgl. Abbildung 6.15): U ist beschränkt, kompakt, nicht offen, abgeschlossen. Die Menge besteht genau aus ihren Randpunkten, insbesondere gibt es keine inneren Punkte.

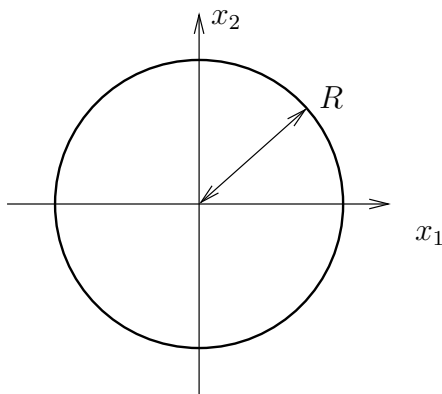


Abbildung 6.15: Die Kreislinie vom Radius R .

iv) $U = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$ (vgl. Abbildung 6.16): U ist nicht beschränkt, nicht kompakt, nicht offen, abgeschlossen. Das Innere ist die Menge

$$\{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1 > 0, x_2 > 0\},$$

die Menge der Randpunkte (gehört in diesem Fall zu U) ist

$$\{\underline{x} \in U : x_1 = 0 \text{ oder } x_2 = 0\}$$

und ist in Abbildung 6.16 rot dargestellt.

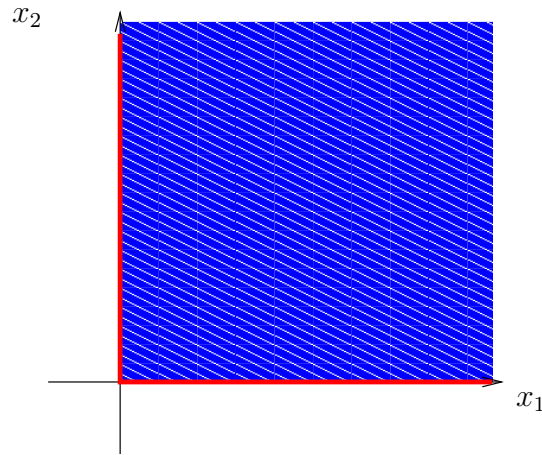


Abbildung 6.16: Ein Quadrant im \mathbb{R}^2 .

Bemerkungen.

- i) Das Komplement einer offenen Menge ist abgeschlossen, das Komplement einer abgeschlossenen Menge ist offen.
- ii) Wenn eine Menge offen ist, heißt das aber nicht, dass sie nicht abgeschlossen ist und vice versa. Es gilt jedoch:

Die einzigen **zugleich offenen und abgeschlossenen** Teilmengen des \mathbb{R}^n sind die leere Menge und der \mathbb{R}^n selbst (vgl. Übungen).

Was geschieht bei Vereinigungen bzw. beim Durchschnitt?

Beispiel.

- i) Für alle $n \in \mathbb{N}$ sei $U_n := \{x \in \mathbb{R} : 1/n < x < 1\}$. Die Mengen U_n sind offen und es ist

$$\bigcup_{n=1}^{\infty} U_n = (0, 1),$$

die **abzählbare (unendliche) Vereinigung**⁶ der offenen Mengen U_n ist wieder offen.

⁶ $x \in \bigcup_{n=1}^{\infty} U_n : \Leftrightarrow x \in U_k$ für ein $k \in \mathbb{N}$; analog: $x \in \bigcap_{n=1}^{\infty} U_n : \Leftrightarrow x \in U_k$ für alle $k \in \mathbb{N}$.

ii) Für alle $n \in \mathbb{N}$ sei $U_n := \{x \in \mathbb{R} : -1/n < x < 1/n\}$. Auch hier sind die Mengen U_n offen und es gilt für festes $1 < N \in \mathbb{N}$

$$\bigcap_{n=1}^N U_n = (-1/N, 1/N) ,$$

der **endliche Durchschnitt** der offenen Mengen ist offen.

Der **abzählbare (unendliche) Durchschnitt** der offenen Mengen U_n ist jedoch **nicht offen**:

$$\bigcap_{n=1}^{\infty} U_n = \{0\} .$$

iii) Bei abgeschlossenen Mengen ist die Situation genau umgekehrt.

Wie in den Beispielen motiviert, kann bewiesen werden:

Satz 6.4.1

- i) Die **Vereinigung** eines **beliebigen Systems offener Mengen** ist offen.
- ii) Der **Durchschnitt** eines **endlichen Systems offener Mengen** ist offen.
- iii) Der **Durchschnitt** eines **beliebigen Systems abgeschlossener Mengen** ist abgeschlossen.
- iv) Die **Vereinigung** eines **endlichen Systems abgeschlossener Mengen** ist abgeschlossen.

Wozu Kompaktheit?

Mit Hilfe des folgenden Satzes wird später gezeigt werden, dass hinreichend “gutartige” Funktionen mit einem **kompakten Definitionsbereich** ihr Maximum und Minimum annehmen. Ohne die Voraussetzung der Kompaktheit ist das i.A. nicht richtig (vgl. Abbildung 6.17, vgl. auch Abbildung 3.1).

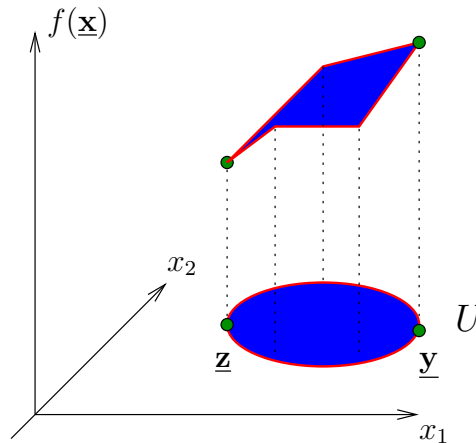


Abbildung 6.17: Ist der Definitionsbereich U von f kompakt, d.h. gehört die rote Randlinie zu U , so gibt es \underline{y} und $\underline{z} \in U$ mit $f(\underline{y}) = \max_{\underline{x} \in U} f(\underline{x})$ bzw. $f(\underline{z}) = \min_{\underline{x} \in U} f(\underline{x})$. Gehört die rote Randlinie nicht zu U , so gilt dies nicht mehr.

Satz 6.4.2

Es sei U eine Teilmenge des \mathbb{R}^n . Dann sind die folgenden Aussagen äquivalent.

- i) U ist kompakt.*
- ii) Jede unendliche Teilmenge von U besitzt einen Häufungspunkt in U .*
- iii) Jede Folge mit Werten in U besitzt eine konvergente Teilfolge mit Grenzwert in U (Folgenkompaktheit).*

Bemerkungen.

- i)* Ebenfalls äquivalent zur Kompaktheit ist: Aus jeder Überdeckung von U mit offenen Mengen lassen sich endlich viele dieser Mengen so auswählen, dass U von ihnen überdeckt wird (“ U hat die Heine-Borel-Eigenschaft”⁷).
- ii)* Der Beweis von Satz 6.4.2 findet sich in jedem Lehrbuch zur Analysis.

⁷E. Heine, 1821-1881; Berlin, Bonn, Halle. E. Borel, 1871-1956; Paris.

- iii) Der Satz **gilt nicht mehr**, wenn statt des \mathbb{R}^n Räume **unendlicher Dimension** (etwa Funktionenräume) betrachtet werden.

Kapitel 7

Die komplexen Zahlen \mathbb{C}

7.1 Einführung der komplexen Zahlen (der Körper der komplexen Zahlen; Fundamentalsatz der Algebra)

- i) Das Studium von Gleichungen 2^{ten} und 3^{ten} Grades führt im 16^{ten} Jahrhundert zur Einführung der komplexen Zahlen (Cardano¹, Bombelli², ..., Euler führt 1777 das Symbol i für die **imaginäre Einheit** ein).
- ii) Was ist eine Lösung der Gleichung $x^2 = -1$?
- iii) Standardanwendung in der Elektrotechnik: Man betrachte einen Wechselstromkreis mit Spule und Ohmschen Widerstand in Reihenschaltung:

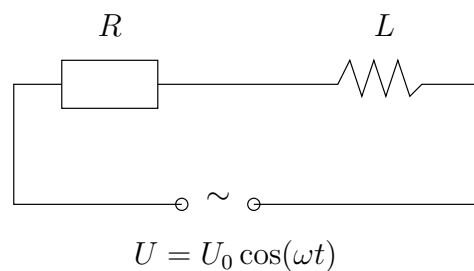


Abbildung 7.1: Reihenschaltung mit Spule und Ohmschen Widerstand.

Ein **komplexer Ansatz** führt auf

$$I(t) = I_0 \cos(\omega t - \alpha)$$

¹G. Cardano, 1501-1576; Pavia, Padua, Mailand, Rom.

²R. Bombelli, 1526-1572; Bologna, Rom.

(I : Stromstärke, $I_0 = \text{const}$, ω : Kreisfrequenz, α : Phasenverschiebung). Der Wechselstromwiderstand (**Impedanz**) ist

$$R^* = R + i\omega L .$$

Mit Eulers Notation ist im Folgenden (zunächst nur **symbolisch**) i eine Lösung der Gleichung $x^2 + 1 = 0$.

Genauer. Betrachtet sei die Menge \mathbb{R}^2 der Paare (a, b) reeller Zahlen.

i) Mit der **Vektoraddition**

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

wird der \mathbb{R}^2 eine **kommutative Gruppe**.

ii) Eine **Multiplikation** ist definiert mittels

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1) .$$

iii) Diese Multiplikation ist **assoziativ** und **kommutativ**.

iv) Es existiert ein **neutrales Element**, nämlich $(1, 0)$:

$$(a, b) \cdot (1, 0) = (a, b) .$$

v) Ist $(a, b) \neq (0, 0)$, so ist $\left(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\right)$ das eindeutig bestimmte **multiplikative inverse Element**:

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}\right) = (1, 0) .$$

Zusammen mit dem **Distributivgesetz** folgt

Definition 7.1.1

*Der \mathbb{R}^2 mit der oben definierten Addition und Multiplikation ist ein **Körper**. Er heißt der Körper der **komplexen Zahlen** und wird mit \mathbb{C} bezeichnet.*

Bemerkungen.

i) Mittels der Abbildung $\varphi: \mathbb{R} \rightarrow \mathbb{C}$,

$$\varphi(a) := (a, 0)$$

können **reelle Zahlen mit komplexen Zahlen der Form $(a, 0)$ identifiziert werden**. Dann schreibt man statt $(a, 0)$ auch einfach a .

ii) Jede komplexe Zahl lässt sich darstellen in der Form

$$(a, b) = (a, 0) + \underbrace{(b, 0) \cdot (0, 1)}_{(0, b)} = a + b(0, 1) ,$$

mit der Abkürzung **$(0, 1) =: i$** (imaginäre Einheit) gilt also

$$(a, b) = a + ib .$$

iii) Wegen

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1$$

ist mit \mathbb{C} ein Körper konstruiert, der den Körper der reellen Zahlen \mathbb{R} enthält und **in dem die Gleichung $x^2 + 1 = 0$ lösbar ist**, nämlich mit i und $-i$. Zu beachten ist, dass i^2 bzgl. der **komplexen Multiplikation** definiert ist.

iv) Formal wird mit komplexen Zahlen (unter Berücksichtigung von $i^2 = -1$) ebenso gerechnet wie mit reellen.

v) Es ist beispielsweise

$$\frac{1+i}{2-i} = \frac{(1+i)(2+i)}{(2-i)(2+i)} = \frac{1}{5}(2+3i+i^2) = \frac{1}{5} + \frac{3}{5}i .$$

Einerseits können die komplexen Zahlen **mit dem \mathbb{R}^2 identifiziert werden** (vgl. Paragraph 7.4). Damit sind konvergente Folgen, offene und abgeschlossene Mengen etc. definiert. **Zusätzlich** ist eine Multiplikation erklärt ist, die \mathbb{C} zu einem Körper macht.

Bezeichnungen, erste Eigenschaften.

- i) Komplexe Zahlen werden häufig mit z oder w bezeichnet.
- ii) In der Regel wird z dargestellt als (s.o.) $z = x + iy$, $x, y \in \mathbb{R}$. Dann heißt $x =: \operatorname{Re} z$ der **Realteil** von z , $y =: \operatorname{Im} z$ der **Imaginärteil**.
- iii) $\bar{z} := x - iy$ heißt die zu $z = x + iy$ **konjugiert komplexe Zahl**.
- iv) $|z| := \sqrt{x^2 + y^2}$ heißt der **Betrag** oder die **Norm** der komplexen Zahl $z = x + iy$.
- v) $\frac{1}{z} := z^{-1} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}$.
- vi) Es gilt:

$$\begin{aligned} \bar{\bar{z}} &= z, & \overline{z+w} &= \bar{z} + \bar{w}, & \overline{z \cdot w} &= \bar{z} \cdot \bar{w}; \\ \operatorname{Re} z &= \frac{1}{2}(z + \bar{z}), & \operatorname{Im} z &= \frac{1}{2i}(z - \bar{z}); \\ z \in \mathbb{R} &\Leftrightarrow z = \bar{z}; \\ |z| &= \sqrt{z\bar{z}}, & \text{d.h. } z\bar{z} &= x^2 + y^2; \\ |z| &\geq 0, & |z| = 0 &\Leftrightarrow z = 0; \\ |z_1 z_2| &= |z_1| |z_2|; \\ |z_1 + z_2| &\leq |z_1| + |z_2| \quad (\text{Dreiecksungleichung}). \end{aligned}$$

Zurück zur Frage nach der **Lösbarkeit algebraischer Gleichungen** beliebigen Grades.

Die Gleichung $x^2 = -1$ hat wie oben gesehen eine Lösung in \mathbb{C} . Dies gilt auch für algebraische Gleichungen beliebigen Grades:

Satz 7.1.1 (Fundamentalsatz der Algebra)

Es sei $n \in \mathbb{N}$ (insbesondere $n \geq 1$) und

$$p_n(z) := a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

ein Polynom n^{ten} Grades ($a_n \neq 0$) mit komplexen Koeffizienten $a_i \in \mathbb{C}$, $i = 1, \dots, n$. Dann gibt es **mindestens eine Nullstelle** $\lambda \in \mathbb{C}$ von p_n , d.h. $p_n(\lambda) = 0$.

Bemerkung. Der Körper \mathbb{C} heißt **algebraisch abgeschlossen**.

Nach Satz 7.1.1 kann p_n geschrieben werden als $p_{n-1}(z)(z - \lambda)$ und sukzessive erhält man die Zerlegung von p_n in **Linearfaktoren**:

$$p_n(z) = a_n(z - \lambda_1)^{k_1}(z - \lambda_2)^{k_2} \dots (z - \lambda_r)^{k_r} .$$

Dabei ist $\mathbb{N} \ni r \leq n$, $k_1 + k_2 + \dots + k_r = n$ und für $i = 1, \dots, r$ sind die λ_i die Nullstellen von $p_n(z)$ mit der **Vielfachheit** k_i .

Werden die Nullstellen **mit ihrer Vielfachheit gezählt, so hat $p_n(z)$ genau n Nullstellen (nicht notwendig verschieden) in \mathbb{C} .**

Beispiele.

i) Es ist (ausmultiplizieren, ein geometrisches Argument folgt später)

$$z^4 - 1 = (z - 1) \cdot (z + 1) \cdot (z - i) \cdot (z + i) .$$

ii) Anhand von Beispiel *i)* erkennt man, dass zusammen mit der Nullstelle $\lambda = i$ auch die konjugiert komplexe Zahl $\bar{\lambda} = -i$ Nullstelle ist (für reelle Nullstellen ist diese Eigenschaft trivial).

Allgemein gilt für ein Polynom mit **reellen Koeffizienten**: Ist

$$p_n(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0$$

und gilt **$a_i \in \mathbb{R}$** für $i = 1, \dots, n$, so folgt aus $\overline{z\bar{w}} = \bar{z}w$ und $\overline{z + w} = \bar{z} + \bar{w}$

$$0 = \overline{p_n(\lambda)} = a_n\bar{\lambda}^n + a_{n-1}\bar{\lambda}^{n-1} + \dots + a_1\bar{\lambda} + a_0 = p_n(\bar{\lambda}) ,$$

demnach ist mit λ auch $\bar{\lambda}$ eine Nullstelle von p_n .

7.2 Potenzreihen im Komplexen (Identitätssatz; Exponentialfunktion; trigonometrische Funktionen; Hyperbelfunktionen; Eulerische Formeln)

Grob gesprochen kann **bis auf eine Ausnahme** im Körper der komplexen Zahlen genauso gerechnet werden wie im Reellen. Die Ausnahme ist:

In \mathbb{C} gibt es **keine Anordnungsrelation** “ $<$ ” im Sinne der Axiomatik aus Paragraph 2.3.

Allerdings ist **der Betrag einer komplexen Zahl reell** und die Beträge von komplexen Zahlen können mit “<” verglichen werden.

Mit anderen Worten: $|z| < |w|$ ist auch im Komplexen definiert, wohingegen $z < w$ im Komplexen nicht erklärt ist (vgl. Übungen).

Ersetzt man beispielsweise in Kapitel 4 die reellen Zahlen \mathbb{R} durch \mathbb{C} , so ist Betragsfunktion im Reellen durch die im Komplexen zu ersetzen, und alle Definitionen und Sätze bleiben gleich, sofern nicht komplexe Zahlen durch “<” miteinander verglichen werden (Übungsaufgabe).

Beispiele.

- i) Definition 4.1.2 überträgt sich wörtlich zu der einer konvergenten komplexen Zahlenfolge.
- ii) Die (komplexe) geometrische Reihe $\sum_{n=0}^{\infty} z^n$ ist konvergent für $|z| < 1$, es ist dann

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}.$$

Für $|z| \geq 1$ divergiert die Reihe (dabei müssen auch im Komplexen die Punkte mit $|z| = 1$ genauer untersucht werden).

- iii) Das Konvergenzkriterium von Leibniz (Satz 4.2.2) ist nicht auf den komplexen Fall übertragbar.

Ebenfalls völlig analog zur reellen Situation ist eine **komplexe Potenzreihe** eine (unendliche Reihe) der Form

$$P(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Die $a_n \in \mathbb{C}$ sind die **Koeffizienten**, z_0 heißt **Entwicklungspunkt**.

Es gilt jetzt (vgl. Satz 5.2.1 und Abbildung 7.2)

Satz 7.2.1

Eine Potenzreihe

$$P(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

konvergiert entweder (absolut und lokal gleichmäßig) auf ganz \mathbb{C} oder es existiert ein $\rho \in [0, \infty)$, sodass $P(z)$ auf

$$B_\rho(z_0) = \{z \in \mathbb{C} : |z - z_0| < \rho\}$$

*(absolut und lokal gleichmäßig) **konvergiert**, aber **divergiert** auf*

$$\{z \in \mathbb{C} : |z - z_0| > \rho\}.$$

Bemerkung. Wieder heißt ρ der **Konvergenzradius**, $B_\rho(z_0)$ heißt der **Konvergenzkreis**. Die Bemerkungen aus Kapitel 5.2 übertragen sich, insbesondere gilt die Formel von Cauchy-Hadamard

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

Wie ist \exp im Komplexen definiert?

Die Exponentialfunktion ist im Reellen als Potenzreihe definiert. Die **komplexe Potenzreihe** $\sum_{n=0}^{\infty} z^n/n!$ stimmt für $z \in \mathbb{R}$ mit der reellen Exponentialfunktion überein. Gibt es noch andere komplexe Potenzreihen mit dieser Eigenschaft?

Der folgende Satz verneint diese Frage.

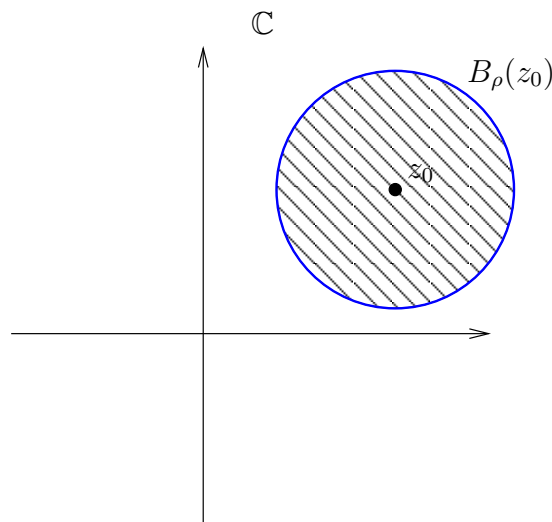


Abbildung 7.2: Zum Konvergenzverhalten einer komplexen Potenzreihe: In der schraffierten offenen Kreisscheibe liegt Konvergenz vor, der (blau dargestellte) Rand ist genau zu analysieren, außerhalb der abgeschlossenen Kreisscheibe divergiert die Reihe.

Satz 7.2.2 (*Identitätssatz für Potenzreihen*)

Für zwei Potenzreihen mit Konvergenzradien $r_1, r_2 > 0$,

$$f(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad |z - z_0| < r_1,$$

$$g(z) := \sum_{n=0}^{\infty} b_n (z - z_0)^n, \quad |z - z_0| < r_2,$$

gilt: Gibt es unendlich viele $z_k, k \in \mathbb{N}$, und ein z_0 mit

- i) $z_k \neq z_0$ für alle k ,
- ii) $\lim_{k \rightarrow \infty} z_k = z_0$,
- iii) $f(z_k) = g(z_k)$ für alle k ,

so gilt $a_n = b_n$ für alle $n = 1, 2, \dots$

Korollar 7.2.1

Zwei komplexe Potenzreihen, die auf \mathbb{R} übereinstimmen, sind *notwendigerweise schon gleich*.

Insbesondere gibt es **nur eine Möglichkeit**, die reelle Exponentialfunktion zu einer Potenzreihe $\exp: \mathbb{C} \rightarrow \mathbb{C}$ fortzusetzen.

Definition 7.2.1

Als Funktionen $\mathbb{C} \rightarrow \mathbb{C}$ sind die *Exponentialfunktion*, die *trigonometrischen Funktionen Sinus und Kosinus* sowie die *Hyperbelfunktionen Sinus und Kosinus Hyperbolicus* definiert als

$$\begin{aligned} \exp(z) &:= \sum_{n=0}^{\infty} \frac{1}{n!} z^n ; \\ \sin(z) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} z^{2n+1} ; \\ \cos(z) &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} z^{2n} ; \\ \sinh(z) &:= \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} z^{2n+1} ; \\ \cosh(z) &:= \sum_{n=0}^{\infty} \frac{1}{(2n)!} z^{2n} . \end{aligned}$$

Bemerkungen.

- i)* In Kapitel 6.2 wurde der **Kosinus zweier Vektoren** im \mathbb{R}^n definiert, und es wurde gezeigt, dass die Definition mit der elementargeometrischen konsistent ist. Ebenso muss hier gezeigt werden (siehe

Paragraph 7.3), dass die Definition der trigonometrischen Funktionen als Potenzreihen mit der elementargeometrischen (und damit auch mit der Definition des Kosinus aus Kapitel 6.2) verträglich ist.

ii) Der komplexe Logarithmus ist nicht so direkt zu definieren wie im reellen Fall. Dies ist ein Problem aus der Funktionentheorie.

iii) Insbesondere ist $\cos(0) = 1$, $\sin(0) = 0$ und

$$\cos(-z) = \cos(z) \ , \quad \sin(-z) = -\sin(z) \ .$$

iv) Die reellen Hyperbelfunktionen wurden bereits in den Übungen diskutiert.

In der Reihendarstellung des Sinus bzw. des Kosinus kann $(-1)^n$ durch $(i^2)^n$ ersetzt werden, d.h.

$$\begin{aligned} i \sin(z) &= i \sum_{n=0}^{\infty} \frac{i^{2n}}{(2n+1)!} z^{2n+1} = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} (iz)^{2n+1} \ , \\ \cos(z) &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} (iz)^{2n} \ , \end{aligned}$$

und es gilt für alle $z \in \mathbb{C}$

$$\exp(iz) = \cos(z) + i \sin(z) \ . \tag{*}$$

Ersetzt man weiter in (*) z durch $-z$ und beachtet, dass per definitionem $\cos(-z) = \cos(z)$, $\sin(-z) = -\sin(z)$, so folgt für alle $z \in \mathbb{C}$

$$\exp(-iz) = \cos(z) - i \sin(z) \ .$$

Zusammen ergeben sich die [Eulerschen Formeln](#)

$$\begin{aligned} \cos(z) &= \frac{1}{2} \left(\exp(iz) + \exp(-iz) \right) \ , \\ \sin(z) &= \frac{1}{2i} \left(\exp(iz) - \exp(-iz) \right) \end{aligned}$$

(da die Funktionalgleichung der Exponentialfunktion auch im Komplexen gilt, ist $\sin^2(z) + \cos^2(z) = 1$ eine unmittelbare Folgerung).

Für $z = x + iy$, $x, y \in \mathbb{R}$, lautet (*)

$$\exp(z) = \exp(x) \exp(iy) = \exp(x)(\cos(y) + i \sin(y)) .$$

Insbesondere gilt für alle $t \in \mathbb{R}$ (man wähle $x = 0$ und schreibe $y = t$)

$$e^{it} := \exp(it) = \cos(t) + i \sin(t) . \quad (**)$$

Mit Hilfe der Gleichung (**) sollen nun die reellen trigonometrischen Funktionen näher studiert werden.

7.3 Die reellen trigonometrischen Funktion (Konsistenz der Definitionen; Eigenschaften)

Wie oben bereits angemerkt, ist zunächst zu untersuchen, ob die Definition der trigonometrischen Funktionen als Potenzreihen mit den bereits bekannten verträglich ist. Zu diesem Zweck sei für ein beliebiges **reelles** t die **komplexe Zahl** e^{it} betrachtet. Identifiziert man \mathbb{C} mit dem \mathbb{R}^2 (wie in Kapitel 7.1 angedeutet), so kann e^{it} als ein Vektor

$$\underline{\mathbf{v}} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \operatorname{Re} e^{it} \\ \operatorname{Im} e^{it} \end{pmatrix} \in \mathbb{R}^2$$

angesehen werden (vgl. Abbildung 7.3).

Schließt dieser Vektor mit der x -Achse den Winkel α ein, so liefert die elementargeometrische Definition von Sinus und Kosinus (für einen Moment mit \sin_{el} , \cos_{el} bezeichnet) über den Einheitskreis

$$\underline{\mathbf{v}} = \cos_{el}(\alpha) \underline{\mathbf{e}}^{(1)} + \sin_{el}(\alpha) \underline{\mathbf{e}}^{(2)} ,$$

wobei $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 bezeichne. Hierbei ist zu beachten, dass aus der Definition der Exponentialfunktion und den Eulerschen Formeln insbesondere $|\underline{\mathbf{v}}| = |e^{it}| = 1$ folgt.

Definiert man den Kosinus und den Sinus hingegen mit der Reihendarstellung aus Definition 7.2.1, so liefert (**) aus Paragraph 7.2

$$\underline{\mathbf{v}} = \cos(t) \underline{\mathbf{e}}^{(1)} + \sin(t) \underline{\mathbf{e}}^{(2)} ,$$

es ist festzuhalten

$$\cos_{el}(\alpha) = \cos(t) \quad \text{und} \quad \sin_{el}(\alpha) = \sin(t) .$$

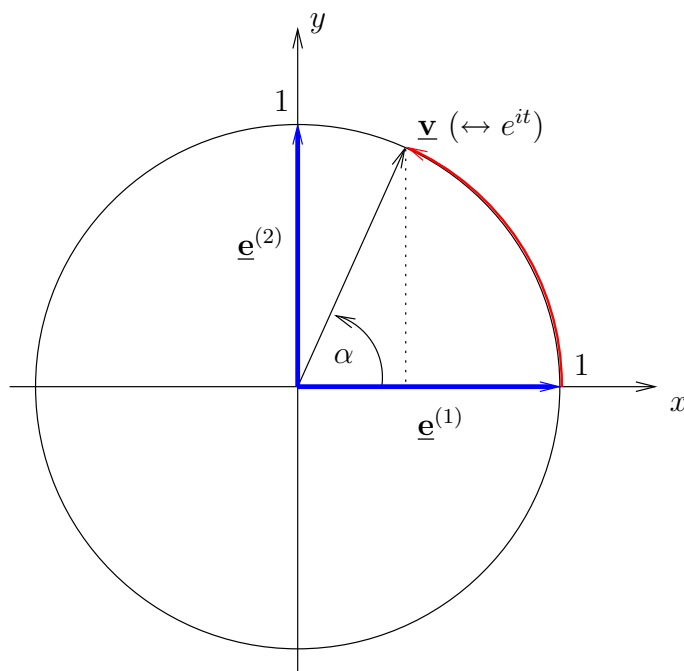


Abbildung 7.3: Zur Darstellung von e^{it} auf dem Einheitskreis.

In diesem Sinne sind die beiden Definitionen miteinander verträglich. Während α jedoch den Winkel in einer 360-Grad Skala misst, ist t das **Bogenmaß** (die Länge) des in Abbildung 7.3 rot skizzierten Kreisbogens, der im **mathematisch positiven Sinne** (vgl. Abbildung 7.3) durchlaufen wird (t geht aus α durch Multiplikation mit $2\pi/360$ hervor).

Im Folgenden werden **Winkel** (wenn nicht anders vereinbart) **im Bogenmaß gemessen**, sodass mit der Reihendarstellung aus Definition 7.2.1 gearbeitet werden kann.

Eigenschaften der reellen trigonometrischen Funktionen.

- i) In den Abbildungen 7.4 und 7.5 sind die Graphen der Funktionen \cos , \sin , \cosh , \sinh dargestellt.

In der Abbildung 7.6 sind die **reelle Tangensfunktion** ($\tan(x) := \frac{\sin(x)}{\cos(x)}$) und die **reelle Kotangensfunktion** ($\cot(x) := \frac{\cos(x)}{\sin(x)}$) dargestellt (vgl. Übungen zu den Umkehrfunktionen).

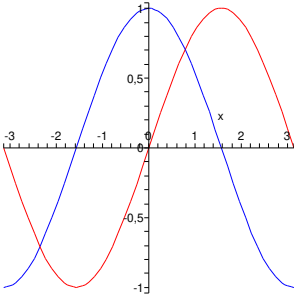


Abbildung 7.4: Der Kosinus und der Sinus.

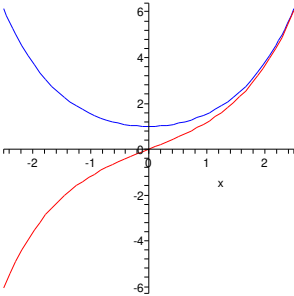


Abbildung 7.5: Die hyperbolische Kosinusfunktion bzw. Sinusfunktion.

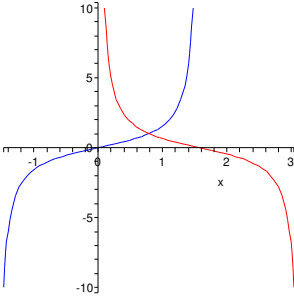


Abbildung 7.6: Der Tangens und der Kotangens.

ii) Es gelten die **Additionstheoreme**

$$\begin{aligned}\cos(x \pm y) &= \cos(x) \cos(y) \mp \sin(x) \sin(y) , \\ \sin(x \pm y) &= \cos(x) \sin(y) \pm \sin(x) \cos(y) , \\ \cos(2x) &= 2 \cos^2(x) - 1 , \\ \sin(2x) &= 2 \cos(x) \sin(x) .\end{aligned}$$

iii) Die transzendente Zahl π **ist definiert als** die eindeutig bestimmte reelle Zahl mit

$$\cos(\pi/2) = 0 \quad \text{und} \quad 0 < \frac{\pi}{2} < 2 .$$

iv) Der in Definition 7.2.1 definierte Kosinus (bzw. der Sinus) ist **2π -periodisch**, d.h. es gilt für alle $x \in \mathbb{R}$

$$\cos(x + 2\pi) = \cos(x) \quad \text{und} \quad \sin(x + 2\pi) = \sin(x) .$$

7.4 Die Gaußsche Zahlenebene (Polarkoordinaten; Einheitswurzeln)

Bei der Einführung der komplexen Zahlen in Kapitel 7.1 sind diese mit Paaren reeller Zahlen identifiziert worden, geometrisch kann eine komplexe Zahl damit als Vektor in der Ebene \mathbb{R}^2 dargestellt werden und man spricht von der **Gaußschen³Zahlenebene** (oder von der **komplexen Ebene**) (vgl. Abbildung 7.7).

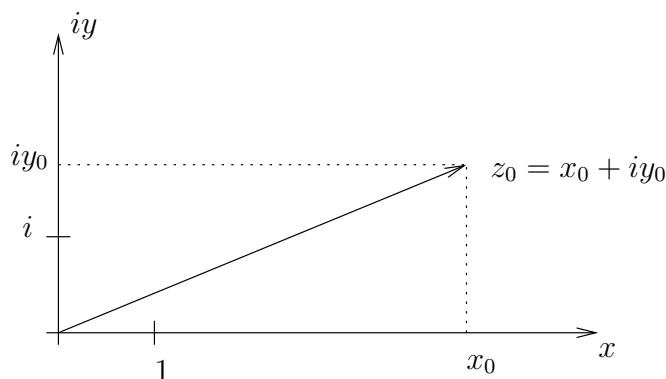


Abbildung 7.7: Die Gaußsche Zahlenebene.

³C.F. Gauß, 1777-1855; Göttingen.

In der Gaußschen Zahlenebene können viele Operationen (wie etwa die Addition und die Konjugation komplexer Zahlen unmittelbar veranschaulicht werden (vgl. Abbildung 7.8 und Abbildung 7.9)).

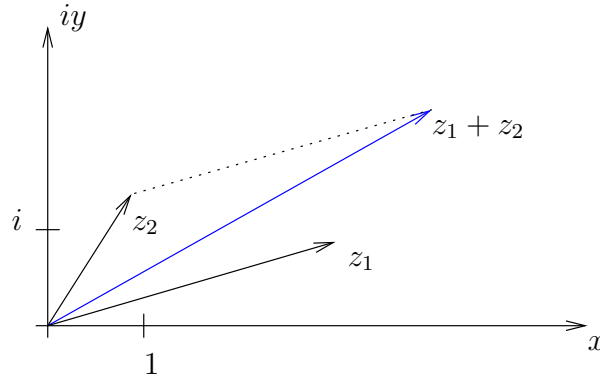


Abbildung 7.8: Zur Addition komplexer Zahlen, zur Dreiecksungleichung.

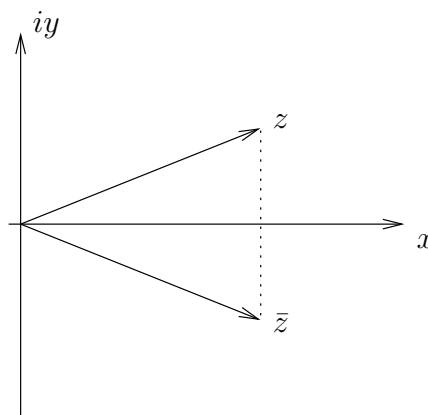


Abbildung 7.9: Zur Konjugation komplexer Zahlen.

Wesentlich ist aber die **Darstellung in Polarkoordinaten**: Aus der Diskussion der letzten beiden Abschnitte folgt, dass jede komplexe Zahl z geschrieben werden kann als (vgl. Abbildung 7.10)

$$z = re^{i\varphi} = r(\cos(\varphi) + i\sin(\varphi)), \quad 0 \leq r, \quad \varphi \in \mathbb{R}.$$

Hier bei ist $r = |z|$ und $\varphi =: \arg z$ heißt **ein** (man beachte die Periodizität der reellen trigonometrischen Funktionen) **Argument** von z .

Mit dieser Darstellung kann auch die Multiplikation komplexer Zahlen in der Gaußschen Zahlenebene geometrisch interpretiert werden. Ist $w = |w|e^{i\varphi}$ und $z = |z|e^{i\psi}$, so gilt

$$wz = |w||z|e^{i(\varphi+\psi)} = |w||z|(\cos(\varphi + \psi) + i\sin(\varphi + \psi)) :$$

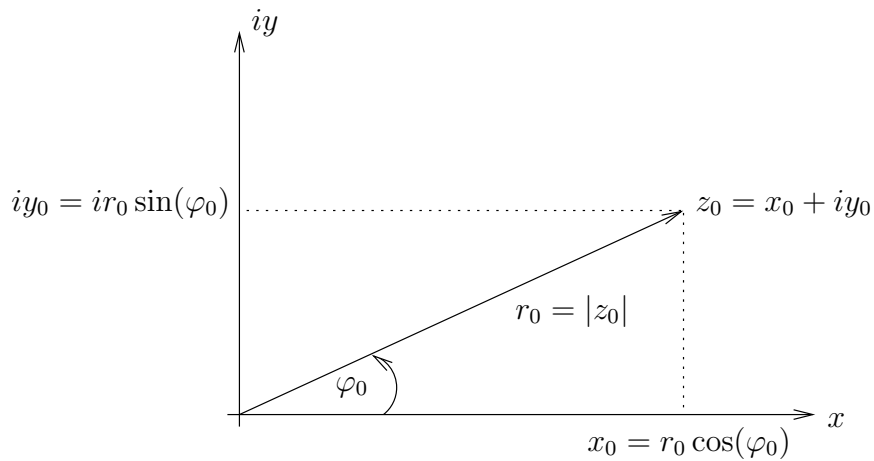


Abbildung 7.10: Komplexe Zahlen in Polarkoordinaten.

Es werden die **Beträge multipliziert** und die **Argumente addiert** (vgl. Abbildung 7.11).

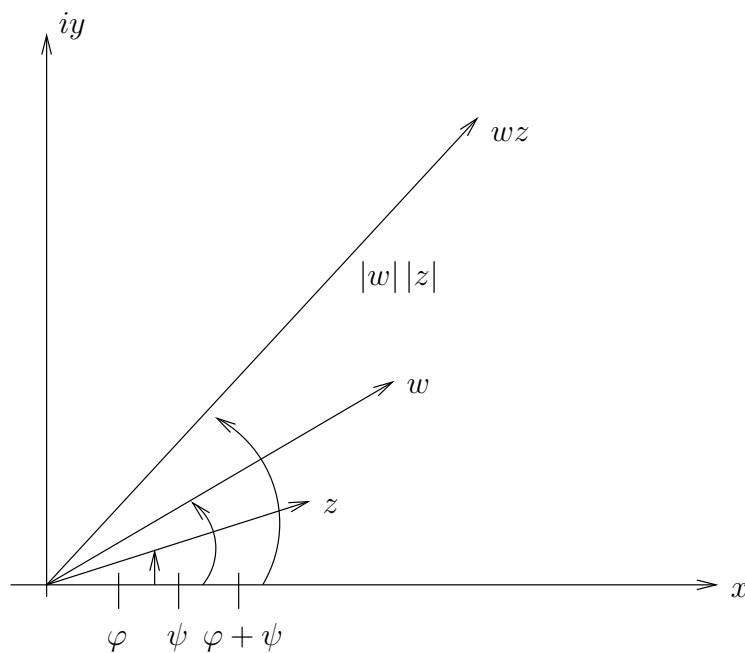


Abbildung 7.11: Zur Multiplikation in der Gaußschen Zahlenebene.

Daraus folgt unmittelbar: Es gibt genau n verschiedene komplexe Zahlen z_0, \dots, z_{n-1} , die der Gleichung $z^n = 1$ genügen. Sie heißen **Einheitswurzeln**:

$$z_k = e^{i\frac{2\pi k}{n}}, \quad k = 0, 1, \dots, n-1.$$

Teil II
Lineare Algebra (Teil 1)

Kapitel 8

Matrizen und lineare Gleichungssysteme

8.1 Matrizenkalkül (der Vektorraum $M(n, m, \mathbb{K})$; lineare Abbildungen; lineare Gleichungssysteme; Matrizenprodukt; Transponierte; Konjugierte; Adjungierte; hermitesches Skalarprodukt im \mathbb{C}^n)

Bevor typische Problemstellungen der **linearen Algebra** (wie die **Lösung linearer Gleichungssysteme** oder die Diskussion **linearer Abbildungen**) angegangen werden können, muss zunächst als wesentlicher Bestandteil dieser Fragestellungen der Begriff einer **Matrix** eingeführt werden, anschließend müssen die wichtigsten Operationen mit Matrizen (**Matrizenkalkül**) diskutiert werden. Im Folgenden ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition 8.1.1

*Es seien $n, m \in \mathbb{N}$. Ein $n \times m$ Koeffizientenschema (d.h. eine **Tabelle aus n Zeilen und m Spalten**) der Form*

$$(a_{ij})_{i=1, \dots, n}^{j=1, \dots, m} := A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix},$$

*$a_{ij} \in \mathbb{K}$ für $1 \leq i \leq n$, $1 \leq j \leq m$, heißt eine **$n \times m$ Matrix**.*

Bemerkungen.

- i) **Merke:** Der Eintrag a_{ij} steht in der i^{ten} Zeile und in der j^{ten} Spalte.
- ii) Die Menge der $n \times m$ -Matrizen über \mathbb{K} wird häufig mit $M(n, m, \mathbb{K})$ bezeichnet (weitere Bezeichnungen: $\mathbb{K}^{(n,m)}$, $\mathbb{K}^{n \times m}$).
- iii) Können bzgl. der Zeilenzahl und der Spaltenzahl keine Missverständnisse auftreten, so schreibt man oft einfach (a_{ij}) für die Matrix.
- iv) Die Elemente des \mathbb{R}^n wurden in Kapitel 6.1 als Spaltenvektoren eingeführt. Diese Vektoren können als $n \times 1$ Matrizen über \mathbb{R} aufgefasst werden:

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in M(n, 1, \mathbb{R}).$$

Wie der \mathbb{R}^n oder der¹ \mathbb{C}^n ist auch die Menge $M(n, m, \mathbb{K})$ ein **Vektorraum**. Dabei sind die Addition und die Multiplikation mit einem Skalar komponentenweise definiert: Für zwei Matrizen $(a_{ij}), (b_{ij}) \in M(n, m, \mathbb{K})$ und für $\lambda \in \mathbb{K}$ gilt per definitionem

$$\begin{aligned} \lambda(a_{ij}) &= \lambda \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = (\lambda a_{ij}) \\ &= \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix} \end{aligned}$$

¹ $\mathbb{C}^n := \underbrace{\mathbb{C} \times \mathbb{C} \cdots \times \mathbb{C}}_n$; n -Tupel komplexer Zahlen.

sowie

$$\begin{aligned}
 (a_{ij}) + (b_{ij}) &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} \\
 &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{pmatrix} \\
 &= (a_{ij}) + (b_{ij}).
 \end{aligned}$$

Klarerweise ist die **Nullmatrix** (0) die Matrix, deren Eintragungen nur aus 0 bestehen, die zu (a_{ij}) **inverse Matrix** (bzgl. der Addition) ist die Matrix $(-a_{ij})$.

Ebenso wie eine reelle Zahl a mittels $f: x \mapsto ax$, $\mathbb{R} \rightarrow \mathbb{R}$, als **lineare Abbildung** interpretiert werden kann, werden später Matrizen geeignete Darstellungen für lineare Abbildungen $\mathbb{R}^m \rightarrow \mathbb{R}^n$ liefern. Dazu benötigt man eine **Matrizenmultiplikation**, die einem Vektor aus dem \mathbb{R}^m einen Vektor im \mathbb{R}^n zuordnet. Damit die “Dimensionen” zusammenpassen, bietet es sich an, eine $n \times m$ Matrix auf einen Vektor des \mathbb{R}^m anzuwenden.

Dies geschieht wie folgt: Es sei $A = (a_{ij}) \in M(n, m, \mathbb{K})$ und $\underline{x} \in \mathbb{K}^m$. Dann ist

$$\underline{A}\underline{x} := \underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}}_{\in M(n,m,\mathbb{K})} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}}_{\in \mathbb{K}^m} := \underbrace{\begin{pmatrix} \sum_{k=1}^m a_{1k}x_k \\ \sum_{k=1}^m a_{2k}x_k \\ \vdots \\ \sum_{k=1}^m a_{nk}x_k \end{pmatrix}}_{\in \mathbb{K}^n}.$$

Es muss dabei die **Spaltenzahl der Matrix A gleich m sein**, damit die Matrix mit einem Vektor aus dem \mathbb{K}^m multipliziert werden kann. Die

Zeilenzahl muss gleich n sein, damit ein Vektor aus dem \mathbb{K}^n entsteht.

Merkregel.

Das Produkt der i^{ten} Zeile der Matrix A mit dem Spaltenvektor $\underline{\mathbf{x}}$ (gemäß $\sum_{k=1}^m a_{ik}x_k$) ergibt den i^{ten} Eintrag des resultierenden Spaltenvektors.

Beispiel. Es sei $\mathbb{K} = \mathbb{R}$, $n = 3$, $m = 4$ und

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix} \in M(3, 4, \mathbb{R}), \quad \underline{\mathbf{x}} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^4.$$

Dann ist

$$\begin{aligned} A\underline{\mathbf{x}} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \\ 2 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 + 4 + 3 + 0 \\ 0 + 2 + 0 + 0 \\ 2 + 4 + 1 + 0 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 7 \end{pmatrix} \in \mathbb{R}^3. \end{aligned}$$

Bemerkung. Ist $A \in M(n, m, \mathbb{K})$ fixiert, so verifiziert man unmittelbar, dass die Funktion $T: \mathbb{K}^m \rightarrow \mathbb{K}^n$, $\underline{\mathbf{x}} \mapsto A\underline{\mathbf{x}}$ eine **lineare Abbildung** ist, d.h. für $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{K}^m$ und für $\lambda \in \mathbb{K}$ gilt

$$\begin{aligned} A(\underline{\mathbf{x}} + \underline{\mathbf{y}}) &= A\underline{\mathbf{x}} + A\underline{\mathbf{y}}, \\ A(\lambda\underline{\mathbf{x}}) &= \lambda(A\underline{\mathbf{x}}). \end{aligned}$$

Zudem ist direkt zu erkennen (B ebenfalls aus $M(n, m, \mathbb{K})$)

$$(A + B)\underline{\mathbf{x}} = A\underline{\mathbf{x}} + B\underline{\mathbf{x}}.$$

Wichtiges Beispiel: Lineare Gleichungssysteme. Für $1 \leq i \leq n$ und $1 \leq j \leq m$ seien a_{ij} und $b_i \in \mathbb{R}$ fixiert. Man betrachte dann ein lineares System aus n Gleichungen in m Unbekannten x_1, \dots, x_m . Dies sieht wie folgt aus:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= b_n. \end{aligned}$$

Werden die x_j und die b_i zu Spaltenvektoren zusammengefasst,

$$\underline{\mathbf{x}} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m, \quad \underline{\mathbf{b}} := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n,$$

und bezeichnet A die Matrix (a_{ij}) , so kann das Gleichungssystem äquivalent geschrieben werden als

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

Mit Hilfe der Multiplikation einer Matrix mit einem Spaltenvektor kann auch definiert werden, was das Produkt zweier Matrizen (geeigneter "Dimensionen") sein soll. Ist l eine weitere natürliche Zahl und ist $B \in M(m, l, \mathbb{K})$, so kann B geschrieben werden als Tupel von **Spaltenvektoren**

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1l} \\ b_{21} & b_{22} & \dots & b_{2l} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \dots & b_{ml} \end{pmatrix} = (\underline{\mathbf{b}}^{(1)} \ \underline{\mathbf{b}}^{(2)} \ \dots \ \underline{\mathbf{b}}^{(l)}),$$

wobei für alle $i = 1, \dots, l$ gesetzt ist:

$$\underline{\mathbf{b}}^{(i)} = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{mi} \end{pmatrix} \in \mathbb{K}^m .$$

Mit dieser Schreibweise kann in natürlicher Weise definiert werden ($A \in M(n, m, \mathbb{K}), B \in M(m, l, \mathbb{K})$)

$$AB := (\underline{A}\underline{\mathbf{b}}^{(1)} \quad \underline{A}\underline{\mathbf{b}}^{(2)} \quad \dots \quad \underline{A}\underline{\mathbf{b}}^{(l)}) \in M(n, l, \mathbb{K}) .$$

Definition 8.1.2

*Es seien $n, m, l \in \mathbb{N}$ und $A \in M(n, m, \mathbb{K}), B \in M(m, l, \mathbb{K})$. Dann liefert das **Matrizenprodukt** AB per definitionem eine Matrix*

$$C = (c_{ij})_{i=1, \dots, n}^{j=1, \dots, l} \in M(n, l, \mathbb{K}) ,$$

deren Eintrag in der i^{ten} Zeile ($i = 1, \dots, n$) und der j^{ten} Spalte ($j = 1, \dots, l$) gegeben ist durch

$$c_{ij} := \sum_{k=1}^m a_{ik} b_{kj} .$$

Merkregel.

Das Produkt der i^{ten} Zeile der Matrix A mit der j^{ten} Spalte der Matrix B (gemäß $\sum_{k=1}^m a_{ik} b_{kj}$) ergibt den Eintrag in der i^{ten} Zeile und der j^{ten} Spalte der resultierenden Matrix C .

Schema.

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{im} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1l} \\ b_{21} & \dots & b_{2j} & \dots & b_{2l} \\ \vdots & & \vdots & & \vdots \\ b_{m1} & \dots & b_{mj} & \dots & b_{ml} \end{pmatrix} \\
 = \begin{pmatrix} c_{11} & \dots & c_{1j} & \dots & c_{1l} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \dots & c_{ij} & \dots & c_{il} \\ \vdots & & \vdots & & \vdots \\ c_{n1} & \dots & c_{nj} & \dots & c_{nl} \end{pmatrix}$$

Die **Spaltenzahl der Matrix A muss gleich der Zeilenzahl der Matrix B sein**, damit das Schema (und damit die Matrizenmultiplikation) funktionieren kann.

Beispiele.

i) Es seien

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{pmatrix} \in M(2, 3, \mathbb{R}), \\
 B = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 3 & -1 & -1 & 0 \end{pmatrix} \in M(3, 4, \mathbb{R}).$$

Dann ist

$$AB = \begin{pmatrix} 5 & 0 & 1 & 2 \\ 7 & 0 & 2 & 1 \end{pmatrix} \in M(2, 4, \mathbb{R}).$$

ii) Es bezeichne I_m die **quadratische $m \times m$ Einheitsmatrix**,

$$I_m := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix},$$

d.h. die Eintragungen a_{ii} , $i = 1, \dots, m$, auf der **Hauptdiagonalen** sind 1, alle anderen Eintragungen sind 0. Für $A \in M(n, m, \mathbb{K})$ ist $AI_m \in M(n, m, \mathbb{K})$ und es gilt

$$AI_m = A.$$

Für $A \in M(n, m, \mathbb{K})$ ist $I_n A \in M(n, m, \mathbb{K})$ und es gilt

$$I_n A = A.$$

Rechenregeln. Es seien A, B, C Matrizen über \mathbb{K} und $\lambda \in \mathbb{K}$. Dann gilt:

- i) $(A + B)C = AC + BC$ (Distributivgesetz);
- ii) $A(B + C) = AB + AC$ (Distributivgesetz);
- iii) $A(BC) = (AB)C$ (Assoziativgesetz);
- iv) $A(\lambda B) = (\lambda A)B = \lambda(AB)$.

Übung. Wie müssen in den einzelnen Regeln jeweils die Zeilen- und Spaltenzahlen gewählt werden, damit die Aussagen sinnvoll sind?

Vorsicht. Es übertragen sich nicht alle Regeln der Multiplikation etwa reeller Zahlen.

Beispiel. Es sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in M(2, 2, \mathbb{R}), \quad B = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} \in M(2, 2, \mathbb{R}).$$

Dann gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

es ist aber

$$BA = \begin{pmatrix} 10 & 20 \\ -5 & -10 \end{pmatrix}.$$

Mit anderen Worten:

- i)* Aus $AB = (0)$ kann im Allgemeinen nicht gefolgert werden, dass A oder B eine Nullmatrix ist.
- ii)* Auch im Falle **quadratischer Matrizen** ($n = m$), bei denen sowohl das Produkt AB als auch das Produkt BA definiert ist, kann **kein Kommutativgesetz** gelten.

Eine wichtige Operation mit Matrizen ist die **Transposition**. Dabei werden die **Zeilen und Spalten einer Matrix vertauscht**, d.h.: Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \in M(n, m, \mathbb{K}),$$

so heißt

$$A^T := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{pmatrix} \in M(m, n, \mathbb{K}),$$

die **zu A transponierte Matrix**.

Beispiel. Es sei $A \in M(3, 2, \mathbb{R})$,

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} .$$

Dann ist

$$A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M(2, 3, \mathbb{R}) .$$

Die Transposition wurde bereits in einem Spezialfall in Kapitel 6.1 verwendet, als n -Tupel reeller Zahlen (Zeilenvektoren) in der Form von Spaltenvektoren geschrieben wurden: Ist X die $1 \times n$ Matrix $(x_1 \ x_2 \ \dots \ x_n) \in M(1, n, \mathbb{R})$, so ist

$$X^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \underline{\mathbf{x}} \in \mathbb{R}^n (= M(n, 1, \mathbb{R})) .$$

Rechenregeln.

i) Man erkennt sofort für $A, B \in M(n, m, \mathbb{K})$ und $\lambda \in \mathbb{K}$

(a) $(A + B)^T = A^T + B^T$;

(b) $(\lambda A)^T = \lambda A^T$;

(c) $(A^T)^T = A$.

ii) Für $A \in M(n, m, \mathbb{K})$ und $B \in M(m, l, \mathbb{K})$ ist das Matrizenprodukt $AB \in M(n, l, \mathbb{K})$ definiert. Wegen $B^T \in M(l, m, \mathbb{K})$ und $A^T \in M(m, n, \mathbb{K})$ ist aber ebenso die Bildung $B^T A^T \in M(l, n, \mathbb{K})$ erlaubt und es gilt (Übung)

$$(AB)^T = B^T A^T .$$

Ist $\mathbb{K} = \mathbb{C}$, so ist auch die **konjugierte Matrix** von Bedeutung: Zu $A \in M(n, m, \mathbb{C})$ setzt man dazu $\bar{A} := (\bar{a}_{ij})$, bei der konjugierten Matrix werden also alle Eintragungen der Matrix A durch die konjugiert

komplexen Zahlen ersetzt. Es definiert

$$A^* := \bar{A}^T,$$

die so genannte **adjungierte Matrix**. Es folgen die gleichen Rechenregeln wie für transponierte Matrizen, wobei lediglich auf der rechten Seite von i), (b) λ durch $\bar{\lambda}$ zu ersetzen ist.

Beispiel. (**Hermitesches Skalarprodukt**²) Ist $\mathbb{K} = \mathbb{R}$, so kann das Euklidische Skalarprodukt im \mathbb{R}^n geschrieben werden als ($\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^n$)

$$\langle \underline{\mathbf{x}}, \underline{\mathbf{y}} \rangle = \underline{\mathbf{y}}^T \underline{\mathbf{x}} = \begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^n x_k y_k.$$

Ist $\mathbb{K} = \mathbb{C}$, so erinnere man sich daran, dass der Betrag einer komplexen Zahl z als $\sqrt{z\bar{z}}$ definiert ist. Bei der Definition eines geeigneten Skalarproduktes im \mathbb{C}^n sollte entsprechend die Transponierte durch die Adjungierte ersetzt werden: Durch

$$\langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle := \underline{\mathbf{w}}^* \underline{\mathbf{z}} = \sum_{k=1}^n z_k \bar{w}_k$$

für alle

$$\underline{\mathbf{z}} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n, \quad \underline{\mathbf{w}} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$$

ist im \mathbb{C}^n ein hermitesches Skalarprodukt definiert. Man beachte, dass in diesem Fall gilt

$$\langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle = \overline{\langle \underline{\mathbf{w}}, \underline{\mathbf{z}} \rangle}, \quad \langle \underline{\mathbf{z}}, \lambda \underline{\mathbf{w}} \rangle = \bar{\lambda} \langle \underline{\mathbf{z}}, \underline{\mathbf{w}} \rangle.$$

Sprechweisen. Zum Abschluss dieses Kapitels seien noch einige Sprechweisen (Definitionen) für **quadratische Matrizen** zusammengefasst:

²C. Hermite, 1822-1901; Paris.

- i) Eine Matrix $A \in M(n, n, \mathbb{R})$ heißt
- (a) **symmetrisch**, wenn $A = A^T$;
 - (b) **orthogonal**, wenn $AA^T = A^T A = I_n$;
 - (c) **positiv definit**, wenn A symmetrisch ist und für alle $\underline{\mathbf{x}} \in \mathbb{R}^n$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, gilt $\underline{\mathbf{x}}^T A \underline{\mathbf{x}} > 0$.
- ii) Eine Matrix $A \in M(n, n, \mathbb{C})$ heißt
- (a) **Hermiteisch**, wenn $A = A^*$;
 - (b) **unitär**, wenn $AA^* = A^* A = I_n$;
 - (c) **positiv definit**, wenn A Hermiteisch ist und für alle $\underline{\mathbf{z}} \in \mathbb{C}^n$, $\underline{\mathbf{z}} \neq \underline{\mathbf{0}}$, gilt $\underline{\mathbf{z}}^T A \underline{\mathbf{z}} > 0$.

8.2 Zur Lösbarkeit linearer Gleichungssysteme (homogene und inhomogene Systeme; Kern einer Matrix; allgemeine Lösung; Superposition von Lösungen; erweiterte Matrix; Spalten- und Zeilenrang; Rang einer Matrix; Dimensionsformel)

Im Folgenden werden lineare Gleichungssysteme mit reellen (der Einfachheit halber) Koeffizienten betrachtet. Wie im Beispiel aus Paragraph 8.1 ist die Situation: Für $1 \leq i \leq n$ und $1 \leq j \leq m$ seien a_{ij} und $b_i \in \mathbb{R}$ fixiert. Zu diesen Daten wird ein **lineares System aus n Gleichungen in m Unbekannten** x_1, \dots, x_m betrachtet:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \cdots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 \cdots + a_{2m}x_m &= b_2 \\ &\vdots \quad \vdots \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 \cdots + a_{nm}x_m &= b_n . \end{aligned}$$

Die äquivalente Schreibweise in der Form

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

wurde bereits diskutiert.

Wann ist ein solches System überhaupt lösbar?

Dazu ist zunächst zu unterscheiden, ob das System **homogen**, d.h. von der Form

$$A\underline{\mathbf{x}} = \underline{\mathbf{0}}$$

ist, oder ob es ein **inhomogenes System** der Gestalt

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}, \quad \underline{\mathbf{b}} \neq \underline{\mathbf{0}},$$

ist. Das homogene System hat immer die **triviale Lösung** $\underline{\mathbf{x}} = \underline{\mathbf{0}}$, hier ist die Suche nach **nicht-trivialen** Lösungen entscheidend. Den Zusammenhang mit dem inhomogenen Fall liefert der

Satz 8.2.1

*i) Die Menge der Lösungen des homogenen Systems $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ ist ein **Unterraum des \mathbb{R}^m** , genannt der **Kern der Matrix A** , **kern A** . Entweder ist $\text{kern } A = \{\underline{\mathbf{0}}\}$ oder es ist $\dim(\text{kern } A) = k \in \mathbb{N}$, $k \leq m$. Dann existiert eine **Basis $(\underline{\mathbf{x}}^{(1)}, \underline{\mathbf{x}}^{(2)}, \dots, \underline{\mathbf{x}}^{(k)})$** von kern A und die Menge der Lösungen des homogenen Systems ist gegeben durch*

$$\underline{\mathbf{x}}_h = \sum_{j=1}^k \lambda_j \underline{\mathbf{x}}^{(j)}, \quad \lambda_i \in \mathbb{R}, \quad i = 1, \dots, k. \quad (*)$$

Man nennt () auch die **allgemeine Lösung des homogenen Systems**.*

*ii) Ist $\underline{\mathbf{x}}_s$ (irgend-) eine Lösung des inhomogenen Gleichungssystems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ($\underline{\mathbf{x}}_s$ heißt dann **spezielle Lösung**), so ist die **allgemeine Lösung des inhomogenen Systems** gegeben durch*

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}_s + \underline{\mathbf{x}}_h. \quad (**)$$

Bemerkungen.

- i) “Allgemeine Lösung” bedeutet: Für beliebige λ_i ist durch (*) (bzw. (**)) im inhomogenen Fall) eine Lösung gegeben, und jede Lösung ist von dieser Form.
- ii) Kennt man also alle Lösungen des homogenen Systems und nur eine Lösung des inhomogenen System, so kennt man auch schon alle Lösungen des inhomogenen Systems.
- iii) Die Idee des Satzes ist: Die Summe von zwei Lösungen des homogenen Systems muss ebenfalls eine Lösung des homogenen Systems sein (**Superpositionsprinzip**). Die Summe einer Lösung des inhomogenen Systems und einer Lösung des homogenen Systems ist eine Lösung des inhomogenen Systems, die Summe zweier Lösungen des inhomogenen Systems ist aber nicht mehr Lösung des inhomogenen Systems (nachrechnen!).

Mit Satz 8.2.1 ist natürlich die Frage nach der Lösbarkeit eines Gleichungssystems noch nicht beantwortet. Hier spielt der **Rang einer Matrix** die entscheidende Rolle. Die Idee ist: Man schreibe die Matrix $A \in M(n, m, \mathbb{R})$ in der Form

$$A = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(m)}) , \quad \underline{\mathbf{a}}^{(i)} \in \mathbb{R}^n , \quad i = 1, \dots, m .$$

Weiterhin sei $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(m)})$ die kanonische Basis des \mathbb{R}^m . Dann gilt für alle $i = 1, 2, \dots, m$

$$A\underline{\mathbf{e}}^{(i)} = \underline{\mathbf{a}}^{(i)} .$$

Nun kann jeder Vektor $\underline{\mathbf{x}} \in \mathbb{R}^m$ geschrieben werden als

$$\underline{\mathbf{x}} = \sum_{i=1}^m x_i \underline{\mathbf{e}}^{(i)} , \quad x_i \in \mathbb{R} , \quad i = 1, \dots, m ,$$

d.h.

$$A\underline{\mathbf{x}} = \sum_{i=1}^m x_i A\underline{\mathbf{e}}^{(i)} = \sum_{i=1}^m x_i \underline{\mathbf{a}}^{(i)} .$$

Damit das System

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

lösbar sein kann, muss es also Koeffizienten x_i , $i = 1, 2, \dots, m$, geben (die Unbekannten), sodass

$$\underline{\mathbf{b}} = \sum_{i=1}^m x_i \underline{\mathbf{a}}^{(i)}$$

gilt. Mit anderen Worten: $\underline{\mathbf{b}}$ ist eine Linearkombination der Spaltenvektoren $\underline{\mathbf{a}}^{(i)}$ der Matrix A , $\underline{\mathbf{b}} \in \text{Spann}(\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)})$.

Diese Bedingung kann über den Rang einer Matrix formuliert werden.

Definition 8.2.1

Der Spaltenrang einer Matrix A ist die maximale Anzahl linear unabhängiger Spaltenvektoren von A , $\text{rg } A$.

Mit der Bezeichnung (erweiterte Matrix, $A \in M(n, m, \mathbb{R})$, $\underline{\mathbf{b}} \in \mathbb{R}^n$)

$$(A|\underline{\mathbf{b}}) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array} \right)$$

folgt

Satz 8.2.2

Das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ist genau dann lösbar, wenn gilt

$$\text{rg } A = \text{rg } (A|\underline{\mathbf{b}}) .$$

Beispiel. Es sei $n = m = 2$ und

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} .$$

Dann ist $\text{rg } A = 1$, ebenso ist

$$\text{rg} \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 1 & 2 & 3 \end{array} \right) = 1,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \underline{\mathbf{x}} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

ist lösbar (wie lauten die Lösungen?).

Dahingegen ist

$$\text{rg} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 1 & 2 & 2 \end{array} \right) = 2,$$

das System

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \underline{\mathbf{x}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

ist nicht lösbar (ausprobieren!).

Völlig analog zur Definition 8.2.1 ist

Definition 8.2.2

Der Zeilenrang einer Matrix A ist die maximale Anzahl linear unabhängiger Zeilenvektoren von A .

Bemerkung. Dabei wird $A \in M(n, m, \mathbb{R})$ geschrieben als (vertikales) Tupel von **Zeilenvektoren**

$$A = \begin{pmatrix} \underline{\mathbf{a}}^{(1)T} \\ \underline{\mathbf{a}}^{(2)T} \\ \vdots \\ \underline{\mathbf{a}}^{(n)T} \end{pmatrix}, \quad \underline{\mathbf{a}}^{(i)T} = (a_{i1} \ a_{i2} \ \dots \ a_{im}), \quad i = 1, \dots, n.$$

Es stellt sich die natürliche Frage:

Was hat der Spaltenrang mit dem Zeilenrang zu tun?

Dazu sei beispielhaft eine Matrix $A \in M(2, 3, \mathbb{R})$ betrachtet,

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} .$$

Der Spaltenrang muss in diesem Beispiel kleiner oder gleich zwei sein (im \mathbb{R}^2 gibt es höchstens zwei linear unabhängige Vektoren), er sei etwa gleich 1. Dann kann die Matrix geschrieben werden als (der zweite und der dritte Spaltenvektor sind linear abhängig vom ersten)

$$A = \begin{pmatrix} a_{11} & \alpha a_{11} & \beta a_{11} \\ a_{21} & \alpha a_{21} & \beta a_{21} \end{pmatrix}, \quad \alpha, \beta \in \mathbb{R} .$$

Damit sind die Zeilenvektoren (die Transponierten)

$$a_{11} \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}, \quad a_{21} \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix},$$

linear abhängig und der Zeilenrang ist ebenfalls gleich 1.

Satz 8.2.3

*Der Zeilenrang einer Matrix ist gleich dem Spaltenrang der Matrix, man spricht einfach vom **Rang** der Matrix. Insbesondere folgt für $A \in M(n, m, \mathbb{R})$:*

$$\text{rg } A \leq \min\{n, m\} .$$

Mit diesen Vorbereitungen kann im nächsten Paragraphen ein Verfahren entwickelt werden, das konstruktiv auf die Lösungen linearer Gleichungssysteme führt, aus dem aber auch folgt:

Satz 8.2.4

Es sei $A \in M(n, m, \mathbb{R})$. Dann sind die folgenden Aussagen richtig:

*i) Es gilt die **Dimensionsformel***

$$\dim(\text{kern } A) + \text{rg } A = m .$$

*ii) Ist das homogene Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{0}}$ **unterbestimmt**, d.h. ist $n < m$, so hat es stets nicht-triviale Lösungen.*

*iii) Ist $n = m$, so ist das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ für jede rechte Seite $\underline{\mathbf{b}}$ **genau dann eindeutig lösbar**, wenn $\text{rg } A = n$.*

Bemerkungen.

- i) Zur Erinnerung: Es ist n die Anzahl der Gleichungen, m die Anzahl der Unbekannten.*
- ii) Im Fall $n > m$ (mehr Gleichungen als Unbekannte) kann im Allgemeinen nicht mehr die Lösbarkeit eines Gleichungssystems erwartet werden. Dennoch ist dieser Fall auch gerade in den Anwendungen von Interesse. Ein Beispiel wird in Kapitel 8.4 skizziert.*

8.3 Das Gaußsche Eliminationsverfahren (elementare Zeilen- und Spaltenumformungen; Trapezschemata zur Rangbestimmung; Gaußsches Eliminationsverfahren zur Lösung linearer Gleichungssysteme; Pivotelement)

Im Verlauf dieses Paragraphen ist es wesentlich zu verstehen, welche **elementaren Umformungen** den **Rang einer Matrix unverändert lassen** bzw. ob diese die Lösungsmenge eines linearen Gleichungssystems beeinflussen. Wieder wird exemplarisch der Fall $\mathbb{K} = \mathbb{R}$ betrachtet.

In diesem Abschnitt sei **stets** $n \leq m$ (siehe obige Bemerkung zum Fall $n > m$). Für die Überlegungen zum Rang einer Matrix kann das ohnehin o.E. angenommen werden, da $\text{rg } A = \text{rg } A^T$ (Zeilenrang = Spaltenrang).

Definition 8.3.1

Es sei $A \in M(n, m, \mathbb{R})$. Eine **elementare Zeilenumformung** ist einer der drei folgenden Vorgänge:

- i) Vertauschung zweier beliebiger Zeilen;
- ii) Multiplikation einer Zeile mit $0 \neq \lambda \in \mathbb{R}$;
- iii) Addition eines beliebigen Vielfachen einer Zeile zu einer anderen Zeile.

Elementare Spaltenumformungen sind analog definiert.

Zum Verständnis dieser elementaren Umformungen wird die Matrix

$$A = \begin{pmatrix} 2 & 4 & 1 \\ 1 & 2 & 2 \end{pmatrix} \in M(2, 3, \mathbb{R})$$

mit $\text{rg } A = 2$ sowie das homogene Gleichungssystem

$$\underline{A}\underline{x} = \underline{\mathbf{0}} \quad \Leftrightarrow \quad \begin{aligned} 2x_1 + 4x_2 + 1x_3 &= 0 \\ 1x_1 + 2x_2 + 2x_3 &= 0 \end{aligned}$$

betrachtet (nach dem noch zu beweisenden Satz 8.2.4 existieren nicht-triviale Lösungen).

Man erkennt sofort, dass elementare Umformungen den Rang von A unverändert lassen, beispielsweise ist die lineare Unabhängigkeit der beiden Zeilenvektoren

$$\begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

für alle $\lambda \in \mathbb{R}$ äquivalent zur linearen Unabhängigkeit der Vektoren

$$\begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 + \lambda 2 \\ 2 + \lambda 4 \\ 2 + \lambda 1 \end{pmatrix}.$$

Es ist nämlich ($\alpha, \beta \in \mathbb{R}$)

$$\begin{aligned} \alpha \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 + \lambda 2 \\ 2 + \lambda 4 \\ 2 + \lambda 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow (\alpha + \beta \lambda) \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \end{aligned}$$

woraus sich die Behauptung ergibt (Übung).

Für das lineare Gleichungssystem bedeuten die elementaren Zeilenumformungen:

- i) Vertauschung der Reihenfolge der Gleichungen;
- ii) Multiplikation einer Gleichung mit $0 \neq \lambda \in \mathbb{R}$;
- iii) Addition des Vielfachen einer Gleichung zu einer anderen Gleichung;

sie ändern die Lösungsmenge nicht.

Vorsicht ist bei Spaltenumformungen gegeben: Vertauscht man etwa die erste und die zweite Spalte von A , so ergibt sich

$$\begin{aligned} 4x_1 + 2x_2 + 1x_3 &= 0 \\ 2x_1 + 1x_2 + 2x_3 &= 0, \end{aligned}$$

die ursprüngliche Lösungsmenge erhält man zurück, wenn die **Variablen x_1 und x_2 vertauscht werden**.

Auch die beiden anderen elementaren Spaltenumformungen sind mit Vorsicht zu genießen. (Welchen Operationen entsprechen diese?)

Um den Rang einer Matrix zu bestimmen, wird sie mit elementaren Umformungen in **Trapezform** übergeführt. Der zugrunde liegende Algorithmus wird nach der Formulierung des folgenden Satzes anhand von linearen Gleichungssystemen diskutiert.

Satz 8.3.1

*Elementare Zeilen- und Spaltenumformungen **erhalten den Rang einer Matrix**. Jede Matrix $(0) \neq A \in M(n, m, \mathbb{R})$ lässt sich durch solche Umformungen auf die **Trapezform***

$$\left(\begin{array}{cccc|ccc} t_{11} & t_{12} & \dots & t_{1(r-1)} & t_{1r} & t_{1(r+1)} & \dots & t_{1m} \\ 0 & t_{22} & \dots & t_{2(r-1)} & t_{2r} & t_{2(r+1)} & \dots & t_{2m} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & t_{rr} & t_{r(r+1)} & \dots & t_{rm} \\ \hline 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{array} \right)$$

bringen, wobei $r \in \mathbb{N}$ und $t_{jj} \neq 0$ für $j = 1, \dots, r$. Insbesondere ist $r \leq \min\{n, m\}$ der Rang der Matrix A .

Bemerkung. Dabei kann evtl. der untere bzw. der rechte Teil wegfallen.

Der Beweis von Satz 8.3.1 verläuft völlig analog zum **Gaußschen Eliminationsverfahren zur Lösung linearer Gleichungssysteme**, das nun

vorgelegt wird. Dazu sei ein Gleichungssystem

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}}$$

in der Form

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array} \right)$$

untersucht. Als begleitendes Beispiel zur Erläuterung sei betrachtet

$$\left(\begin{array}{cccc|c} 0 & 0 & 1 & 2 & 2 \\ 1 & 2 & 3 & 0 & 3 \\ 2 & 4 & 4 & -2 & 4 \\ 1 & 2 & 4 & 0 & \lambda \end{array} \right),$$

wobei $\lambda \in \mathbb{R}$ eine beliebige reelle Zahl bezeichne.

I. Der Eliminationsprozess.

i) Schritt 1.

- (a) Ist $a_{11} = 0$, so überprüfe man, ob in der ersten Spalte ein Element a_{k1} existiert mit $a_{k1} \neq 0$. Man vertausche dann die erste und die k^{te} Zeile und erhält

$$(\tilde{A}|\tilde{b}) \quad \text{mit } \tilde{a}_{11} \neq 0. \quad (*)$$

Im Beispiel ergibt sich durch Vertauschen der ersten und zweiten Zeile

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 3 \\ 0 & 0 & 1 & 2 & 2 \\ 2 & 4 & 4 & -2 & 4 \\ 1 & 2 & 4 & 0 & \lambda \end{array} \right).$$

- (b) Existiert kein solches Element in der ersten Spalte, so ist die “Restmatrix” $(a_{ij})_{i=1, \dots, n}^{j=2, \dots, m}$ nach einem Element $a_{kl} \neq 0$ zu durchsuchen. Ist dieses gefunden, so vertausche man die erste und die l^{te} Spalte, die dabei vorgenommene **Vertauschung der Variablen** ist zu berücksichtigen. Anschließend kann (a) durchgeführt werden. Wieder ergibt sich ein System der Form (*).

Dieser Schritt entfällt in obigem Beispiel.

- (c) Sind alle verbleibenden Eintragungen Null, so ist der Algorithmus beendet.

Auch dies trifft im Beispiel nicht zu.

- (d) Ist jetzt $\tilde{a}_{11} \neq 0$ (\tilde{a}_{11} heißt dann **Pivotelement** des ersten Schrittes), so werden für $i = 2, \dots, n$ die folgenden elementaren Zeilenoperationen durchgeführt:

Das $\left(-\frac{\tilde{a}_{i1}}{\tilde{a}_{11}}\right)$ -fache der ersten Zeile wird zur Zeile mit der Nummer i addiert.

Man erhält ein System der Form ($a_{11}^{(1)} \neq 0$)

$$(A^{(1)} | \underline{\mathbf{b}}^{(1)}) = \left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1m}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2m}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nm}^{(1)} & b_n^{(1)} \end{array} \right).$$

Im Beispiel bleibt die zweite Zeile unverändert, von der dritten Zeile wird das zweifache der ersten Zeile abgezogen, von der vierten Zeile wird die erste abgezogen. Dies ergibt

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 3 \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & -2 & -2 & -2 \\ 0 & 0 & 1 & 0 & \lambda - 3 \end{array} \right).$$

ii) *Schritt 2.* Mit den gleichen Operationen wie im ersten Schritt transformiere man das System zunächst auf die Form (vgl. (a), (b), (c))

$$(\tilde{A}^{(1)}|\tilde{b}^{(1)}),$$

wobei $\tilde{a}_{22}^{(1)} \neq 0$ sicherzustellen ist ($\tilde{a}_{22}^{(1)}$ ist das Pivotelement des zweiten Schrittes) und die **erste Spalte unverändert bleibt**.

Im Beispiel gilt $a_{22}^{(1)} = a_{32}^{(1)} = a_{42}^{(1)} = 0$, es wird nach (b) etwa die zweite mit der dritten Spalte vertauscht. Das Ergebnis ist

$$\left(\begin{array}{cccc|c} 1 & 3 & 2 & 0 & 3 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & -2 & 0 & -2 & -2 \\ 0 & 1 & 0 & 0 & \lambda - 3 \end{array} \right).$$

Anschließend (vgl. (d)) addiert man das $-\tilde{a}_{i2}^{(1)}/\tilde{a}_{22}^{(1)}$ -fache der zweiten Zeile zur Zeile mit der Nummer i , $i = 3, \dots, n$, und erhält ein System der Gestalt

$$(A^{(2)}|\underline{\mathbf{b}}^{(2)}) = \left(\begin{array}{cccc|c} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \dots & a_{1m}^{(2)} & b_1^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \dots & a_{2m}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nm}^{(2)} & b_n^{(2)} \end{array} \right)$$

mit $a_{11}^{(2)} \neq 0$, $a_{22}^{(2)} \neq 0$.

Im Beispiel ist es erforderlich, das zweifache der zweiten Zeile zur dritten Zeile zu addieren, von der vierten Zeile wird die zweite abgezogen. Man gelangt im zweiten Schritt zu

$$\left(\begin{array}{cccc|c} 1 & 3 & 2 & 0 & 3 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -2 & \lambda - 5 \end{array} \right).$$

iii) *Letzter Schritt.* Der Eliminationsprozess wird solange fortgesetzt, bis man auf die Form

$$\left(\begin{array}{cccccc|c} a_{11}^{(r)} & a_{12}^{(r)} & \dots & a_{1r}^{(r)} & \dots & a_{1m}^{(r)} & b_1^{(r)} \\ 0 & a_{22}^{(r)} & \dots & a_{2r}^{(r)} & \dots & a_{2m}^{(r)} & b_2^{(r)} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & a_{rr}^{(r)} & \dots & a_{rm}^{(r)} & b_r^{(r)} \\ 0 & 0 & \dots & 0 & \dots & 0 & b_{r+1}^{(r)} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 & b_n^{(r)} \end{array} \right) \quad (**)$$

geführt wird. Dabei ist $r \in \mathbb{N}$, $r \leq \min\{n, m\}$, und für $k = 1, \dots, r$ ist $a_{kk} \neq 0$.

Im Beispiel ist die dritte mit der vierten Spalte zu vertauschen, das Resultat ist

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 3 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & -2 & 0 & \lambda - 5 \end{array} \right) .$$

Als finale Operation ist zur vierten Zeile die dritte zu addieren, es ergibt sich

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & 2 & 3 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & \lambda - 3 \end{array} \right) .$$

Bemerkung. Lässt man die rechte Seite \mathbf{b} bei den obigen Betrachtungen weg, so folgt Satz 8.3.1, insbesondere ist $\operatorname{rg} A = r$.

Im Beispiel ist $\operatorname{rg} A = 3$.

II. Zur Lösbarkeit.

Das System (***) lautet ausführlich geschrieben

$$\begin{aligned}
 a_{11}^{(r)} x_1^{(r)} + a_{12}^{(r)} x_2^{(r)} + \cdots + a_{1r}^{(r)} x_r^{(r)} + \cdots + a_{1m}^{(r)} x_m^{(r)} &= b_1^{(r)} \\
 a_{22}^{(r)} x_2^{(r)} + \cdots + a_{2r}^{(r)} x_r^{(r)} + \cdots + a_{2m}^{(r)} x_m^{(r)} &= b_2^{(r)} \\
 &\vdots \\
 &\vdots \\
 a_{rr}^{(r)} x_r^{(r)} + \cdots + a_{rm}^{(r)} x_m^{(r)} &= b_r^{(r)} \quad (***) \\
 0 &= b_{r+1}^{(r)} \\
 &\vdots \\
 0 &= b_n^{(r)}.
 \end{aligned}$$

Das System ist also **genau dann lösbar, wenn:** $b_{r+1}^{(r)} = b_{r+2}^{(r)} = \cdots = b_n^{(r)} = 0$.

Das obige Beispiel lautet

$$\begin{aligned}
 1x_1^{(3)} + 3x_2^{(3)} + 0x_3^{(3)} + 2x_4^{(3)} &= 3 \\
 1x_2^{(3)} + 2x_3^{(3)} + 0x_4^{(3)} &= 2 \\
 2x_3^{(3)} + 0x_4^{(3)} &= 2 \\
 0x_4^{(3)} &= \lambda - 3.
 \end{aligned}$$

Es ist demnach genau dann lösbar, wenn $\lambda = 3$. In Übereinstimmung mit den Betrachtungen aus Kapitel 8.2 entspricht dies exakt dem Fall, dass \underline{b} linear abhängig von den Spaltenvektoren der Matrix A ist.

III. Die expliziten Lösungen.

Es sei jetzt angenommen, dass das System lösbar ist, d.h. wie bereits gesagt gelte $b_{r+1}^{(r)} = b_{r+2}^{(r)} = \cdots = b_n^{(r)} = 0$.

- i)* Eine **spezielle Lösung des inhomogenen Systems** ist beispielsweise durch die Wahl $x_{r+1}^{(r)} = x_{r+2}^{(r)} = \cdots = x_m^{(r)} = 0$ bestimmt. Dann besagt (***) : $x_r^{(r)} = b_r^{(r)} / a_{rr}^{(r)}$ (zur Erinnerung: $a_{rr}^{(r)} \neq 0$). Das Ergebnis wird in die vorhergehende Gleichung eingesetzt, man erhält $x_{r-1}^{(r)}$

usw. Insgesamt ergibt **Rückwärtseinsetzen** eine spezielle Lösung von $(***)$. Die Komponenten lauten für $i = r, r - 1, \dots, 1$:

$$x_i^{(r)} = \frac{\left[b_i - \sum_{k=i+1}^r a_{ik}^{(r)} x_k^{(r)} \right]}{a_{ii}^{(r)}},$$

wieder ist die Bedingung $a_{kk}^{(r)} \neq 0$ für $k \leq r$ wesentlich.

Eine spezielle Lösung des ursprünglichen Systems wird gewonnen, indem die Variablenvertauschungen (durch Spaltenvertauschungen) rückgängig gemacht werden.

Im Beispiel sieht dies wie folgt aus ($\lambda = 3$): Für die spezielle Lösung nach obiger Methode ist $x_4^{(3)} = 0$, die dritte Gleichung zeigt zudem $x_3^{(3)} = 1$, die zweite liefert $x_2^{(3)} = 0$. Schließlich sieht man anhand der ersten Gleichung $x_1^{(3)} = 3$. Die Spaltenvertauschungen während des Eliminationsprozesses kann man sich in einer Tabelle in Erinnerung rufen

| | | | |
|-----------------|-----------------|-----------------|-----------------|
| $x_1^{(3)} = 3$ | $x_2^{(3)} = 0$ | $x_3^{(3)} = 1$ | $x_4^{(3)} = 0$ |
| $x_1^{(2)} = 3$ | $x_2^{(2)} = 0$ | $x_3^{(2)} = 0$ | $x_4^{(2)} = 1$ |
| $x_1^{(1)} = 3$ | $x_2^{(1)} = 0$ | $x_3^{(1)} = 0$ | $x_4^{(1)} = 1$ |
| $x_1 = 3$ | $x_2 = 0$ | $x_3 = 0$ | $x_4 = 1$. |

Insgesamt ist eine spezielle Lösung des ursprünglichen Systems gefunden (Probe!):

$$\underline{\mathbf{x}}_s = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

- ii) Die **allgemeine Lösung des homogenen Systems** wird wie folgt gewonnen: Das homogene System transformiert sich ebenfalls auf die Form $(***)$, wobei die rechte Seite, also der Vektor $\underline{\mathbf{b}}^{(r)}$, durch $\underline{\mathbf{0}}$ zu ersetzen ist. Ein Element aus kern A findet man, indem $x_{r+1}^{(r)} = 1$ und $x_{r+2}^{(r)} = x_{r+3}^{(r)} = \dots = x_m^{(r)} = 0$ gesetzt

wird und wie oben rückwärts eingesetzt wird. Eine weitere, linear unabhängige Lösung ergibt sich aus der Wahl $x_{r+2}^{(r)} = 1$ und $x_{r+1}^{(r)} = x_{r+3}^{(r)} = \dots = x_m^{(r)} = 0$. Insgesamt sieht man, dass $\dim(\text{kern } A) = m - r$, die Dimensionsformel ist damit gezeigt.

Im Beispiel liefert das $x_4^{(3)} = 1$, d.h. $x_3 = 0$, woraus $x_2^{(3)} = 0$ und schließlich $x_1^{(3)} = -2$ folgt. Die Berücksichtigung der Spaltenvertauschungen führt damit auf die Lösung

$$\underline{\mathbf{x}} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

des homogenen Systems. Im Beispiel ist kern A eindimensional, die allgemeine Lösung des homogenen Systems ist

$$\underline{\mathbf{x}}_h = \alpha \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha \in \mathbb{R},$$

die allgemeine Lösung des inhomogenen Systems lautet

$$\underline{\mathbf{x}} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha \in \mathbb{R}.$$

Bemerkungen.

- i)* Ist $n = m$ und $\text{rg } A = n$, so sind keine Spaltenumformungen notwendig (warum?).
- ii)* Im Fall $n = m$ sieht man anhand von (**), dass die Lösbarkeit des Systems für alle rechten Seiten äquivalent ist zur Bedingung

$\text{rg } A = n$. Man erkennt auch, dass dann das homogene System nur die triviale Lösung hat, was wiederum die eindeutige Lösbarkeit für jede rechte Seite zeigt.

iii) Eine weitere Bemerkung zum Fall $n = m$, $\text{rg } A = n$: Hier ist es möglich, das System (allein durch elementare Zeilenumformungen) auf die Gestalt

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & 0 & \hat{b}_1 \\ 0 & 1 & \dots & 0 & 0 & \hat{b}_2 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \hat{b}_{n-1} \\ 0 & 0 & \dots & 0 & 1 & \hat{b}_n \end{array} \right)$$

zu bringen. Die Lösung ist somit $\underline{\mathbf{x}} = \underline{\hat{\mathbf{b}}}$, d.h. es gilt $A\underline{\hat{\mathbf{b}}} = \underline{\mathbf{b}}$. Zudem kann das Gleichungssystem für mehrere rechte Seiten simultan gelöst werden. Dazu betrachte man die erweiterte Matrix

$$(A | \underline{\mathbf{b}}^{(1)} \quad \underline{\mathbf{b}}^{(2)} \quad \dots \quad \underline{\mathbf{b}}^{(l)})$$

und bringe sie durch elementare Zeilenumformungen auf die Gestalt

$$\left(\begin{array}{cccc|cccc} 1 & 0 & \dots & 0 & 0 & \hat{b}_1^{(1)} & \hat{b}_1^{(2)} & \dots & \hat{b}_1^{(l)} \\ 0 & 1 & \dots & 0 & 0 & \hat{b}_2^{(1)} & \hat{b}_2^{(2)} & \dots & \hat{b}_2^{(l)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \hat{b}_{n-1}^{(1)} & \hat{b}_{n-1}^{(2)} & \dots & \hat{b}_{n-1}^{(l)} \\ 0 & 0 & \dots & 0 & 1 & \hat{b}_n^{(1)} & \hat{b}_n^{(2)} & \dots & \hat{b}_n^{(l)} \end{array} \right) .$$

Von besonderer Bedeutung ist dabei der Fall $l = n$ und

$$(\underline{\mathbf{b}}^{(1)} \quad \underline{\mathbf{b}}^{(2)} \quad \dots \quad \underline{\mathbf{b}}^{(n)}) = I_n .$$

Auf diese Art kann die zu A **inverse Matrix** bestimmt werden (vgl. Übungen und Paragraph 8.5).

iv) Bei der numerischen Behandlung von linearen Gleichungssystemen kann die Auswahl des Pivotelementes eine große Rolle spielen.

8.4 Überbestimmte lineare Gleichungssysteme: Eine Anwendung (Residuum; Methode der kleinsten Quadrate; Ausgleich nach Tschebyscheff; lineare Ausgleichsprobleme; Normalgleichung)

In diesem Paragraphen wird ein Anwendungsbeispiel vorgestellt, das auf ein **überbestimmtes** ($n > m$, mehr Gleichungen als Unbekannte) lineares Gleichungssystem führt.

Eine quadratische Gesetzmäßigkeit $f(x) = a_1 + a_2x + a_3x^2$ soll experimentell gefunden werden. Aufgrund von Messfehlern ist es sicherlich nicht angeraten, die gesuchten Parameter a_1 , a_2 und a_3 nur mit Hilfe von drei Messungen bestimmen zu wollen, man wird eher eine Vielzahl von Messungen heranziehen. Es sei hier ein Beispiel mit vier (der Einfachheit wegen) Messungen diskutiert (y_i bezeichne den Messwert an der Stelle x_i):

$$\begin{array}{c|cccc} x_i & 1 & 2 & 3 & 4 \\ \hline y_i & 3 & 1 & 1 & 2 \end{array}.$$

Die Aufgabe, nämlich die a_i , $i = 1, 2, 3$, zu bestimmen, kann als **lineares Gleichungssystem in den a_i** formuliert werden:

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} =: A\mathbf{a} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} =: \mathbf{y}.$$

Man erhält in diesem Beispiel vier Gleichungen in den drei Unbekannten a_1 , a_2 , a_3 , die Aufgabe ist also in der Regel nicht exakt lösbar, was ja auch gerade wegen evtl. Messfehler nicht erwartet werden kann. Um dennoch eine Parabel **möglichst gut an die Daten anzupassen**, wird das **Residuum** \mathbf{r} (sozusagen die “Abweichung von der Exaktheit”) betrachtet

$$\mathbf{r} := A\mathbf{a} - \mathbf{y}.$$

Dieses Residuum soll möglichst klein gemacht werden, wobei **“möglichst klein” auf unterschiedliche Arten gemessen werden kann**:

- i)* Versucht man etwa $(\sum_{i=1}^4 r_i^2)^{1/2} = \|\underline{\mathbf{r}}\|_2$ (also die Euklidische Norm) zu minimieren, so spricht man von der **Methode der kleinsten Quadrate**.
- ii)* Die Minimierung von $\max_{1 \leq i \leq 4} |r_i| = \|\underline{\mathbf{r}}\|_\infty$ heißt der **Ausgleich nach Tschebyscheff**³.

Allgemein lautet die Aufgabe,

$$\|\underline{\mathbf{r}}\| = \|A\underline{\mathbf{a}} - \underline{\mathbf{y}}\|$$

(bzgl. irgendeiner Norm) zu minimieren, ein **lineares Ausgleichsproblem**.

Im Folgenden wird der Fall $\|\cdot\| = \|\cdot\|_2$, die Methode der kleinsten Quadrate, studiert.

Dazu sei in Anlehnung an das Beispiel $A \in M(n, m, \mathbb{R})$, $n > m$, $\underline{\mathbf{y}} \in \mathbb{R}^n$ und für alle $\underline{\mathbf{a}} \in \mathbb{R}^m$ sei

$$\underline{\mathbf{r}}(\underline{\mathbf{a}}) = A\underline{\mathbf{a}} - \underline{\mathbf{y}}.$$

Es ist für $\underline{\mathbf{a}}, \underline{\mathbf{b}} \in \mathbb{R}^m$ ($\underline{\mathbf{c}} := \underline{\mathbf{b}} - \underline{\mathbf{a}}$)

$$\begin{aligned} \underline{\mathbf{r}}(\underline{\mathbf{b}}) &= \underline{\mathbf{r}}(\underline{\mathbf{a}} + \underline{\mathbf{c}}) \\ &= A(\underline{\mathbf{a}} + \underline{\mathbf{c}}) - \underline{\mathbf{y}} \\ &= \underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}}, \end{aligned}$$

und es ergibt sich

$$\begin{aligned} \|\underline{\mathbf{r}}(\underline{\mathbf{b}})\|^2 &= \underline{\mathbf{r}}(\underline{\mathbf{b}})^T \underline{\mathbf{r}}(\underline{\mathbf{b}}) \\ &= (\underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}})^T (\underline{\mathbf{r}}(\underline{\mathbf{a}}) + A\underline{\mathbf{c}}) \\ &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + \underline{\mathbf{r}}(\underline{\mathbf{a}})^T A\underline{\mathbf{c}} + (A\underline{\mathbf{c}})^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) + \|A\underline{\mathbf{c}}\|^2 \\ &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + \langle \underline{\mathbf{r}}(\underline{\mathbf{a}}), A\underline{\mathbf{c}} \rangle + \langle A\underline{\mathbf{c}}, \underline{\mathbf{r}}(\underline{\mathbf{a}}) \rangle + \|A\underline{\mathbf{c}}\|^2 \\ &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + 2\langle A\underline{\mathbf{c}}, \underline{\mathbf{r}}(\underline{\mathbf{a}}) \rangle + \|A\underline{\mathbf{c}}\|^2 \\ &= \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2 + 2\underline{\mathbf{c}}^T (A^T \underline{\mathbf{r}}(\underline{\mathbf{a}})) + \|A\underline{\mathbf{c}}\|^2. \end{aligned} \quad (*)$$

Ist nun $\text{rg } A = m$, was $\|A\underline{\mathbf{c}}\|^2 > 0$ für $\underline{\mathbf{c}} \neq \underline{\mathbf{0}}$ impliziert, und gilt $A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}}$, so folgt für alle $\underline{\mathbf{b}} \in \mathbb{R}^m$, $\underline{\mathbf{b}} \neq \underline{\mathbf{a}}$,

$$\|\underline{\mathbf{r}}(\underline{\mathbf{b}})\|^2 > \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|^2.$$

³P.L. Tschebyscheff, 1821-1894; Moskau, St. Petersburg

Mit anderen Worten: Ist

$$A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}}, \quad (**)$$

so minimiert $\underline{\mathbf{a}}$ das Residuum nach der Methode der kleinsten Quadrate. Umgekehrt folgt aus (*) auch, dass ein Minimierer (**) erfüllen muss (andernfalls wähle man $\underline{\mathbf{b}}$ mit $\underline{\mathbf{c}}^T(A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}))$ negativ und $\|\underline{\mathbf{c}}\|$ hinreichend klein, sodass $\underline{\mathbf{c}}^T(A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}))$ den quadratischen (in $\underline{\mathbf{c}}$) Term $\|A\underline{\mathbf{c}}\|^2$ betragsmäßig dominiert).

Schließlich wird die Gleichung (**) noch umformuliert,

$$\begin{aligned} A^T \underline{\mathbf{r}}(\underline{\mathbf{a}}) = \underline{\mathbf{0}} &\Leftrightarrow A^T(A\underline{\mathbf{a}} - \underline{\mathbf{y}}) = \underline{\mathbf{0}} \\ &\Leftrightarrow A^T A\underline{\mathbf{a}} = A^T \underline{\mathbf{y}}, \end{aligned}$$

und man definiert

Definition 8.4.1

Es sei $A \in M(n, m, \mathbb{R})$, $n > m$, $\text{rg } A = m$ und $\underline{\mathbf{y}} \in \mathbb{R}^n$. Dann heißt das *lineare Gleichungssystem*

$$A^T A\underline{\mathbf{a}} = A^T \underline{\mathbf{y}}$$

die *Normalgleichung* des linearen Ausgleichsproblems.

Die Matrix $A^T A \in M(m, m, \mathbb{R})$ ist quadratisch und anhand von $\underline{\mathbf{x}}^T A^T A \underline{\mathbf{x}} > 0$ für $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$ erkennt man $\text{rg } A^T A = m$ (warum?). Damit ist die Normalgleichung eindeutig lösbar, es gilt

Satz 8.4.1

Unter den Voraussetzungen aus Definition 8.4.1 hat das lineare Ausgleichsproblem nach der Methode der kleinsten Quadrate eine *eindeutige Lösung*, die sich aus der Normalgleichung berechnen lässt.

Bemerkungen. Zu numerischen Berechnungen ist die Normalgleichung oft mit Vorsicht zu betrachten, da sich Rundungsfehler sehr stark auf das Ergebnis auswirken können. Dies wird später nochmals aufgegriffen.

In obigem Beispiel kann aber exakt gerechnet werden, als Lösung des Ausgleichsproblems findet man

$$f(x) = 6.25 - 4.05x + 0.75x^2 ,$$

der Vergleich mit den Daten ist in Abbildung 8.1 dargestellt. Als Maß für die Güte der Anpassung kann $\|\underline{\mathbf{r}}(\underline{\mathbf{a}})\|$ herangezogen werden, im Beispiel ist

$$\underline{\mathbf{r}}(\underline{\mathbf{a}}) = \frac{1}{20} \begin{pmatrix} -1 \\ 3 \\ -3 \\ 1 \end{pmatrix} , \quad \text{d.h. } \|\underline{\mathbf{r}}(\underline{\mathbf{a}})\| = \frac{1}{\sqrt{20}} .$$

Bemerkung. Ebenso können natürlich andere Funktionen als eine Parabel an die Daten angepasst werden. Versucht man etwa eine Gerade nach der Methode der kleinsten Quadrate an die Daten anzupassen, so erhält man die in Abbildung 8.2 dargestellte Lösung (Übungsaufgabe).

8.5 Invertierbare Matrizen (inverse, reguläre, singuläre Matrizen; Dreiecksmatrizen; LR-Zerlegung; Cholesky-Zerlegung)

In diesem Abschnitt sei $A \in M(n, n, \mathbb{R})$ **stets eine quadratische $n \times n$ Matrix über \mathbb{R}** , wobei der komplexe Fall völlig analog zu behandeln ist. Es soll die Frage untersucht werden, wann eine inverse Matrix A^{-1} existiert, für die per definitionem gilt

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

Existiert eine solche Matrix, so ist notwendigerweise das lineare Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ für jede rechte Seite lösbar, es gilt nämlich in diesem Fall

$$A\underline{\mathbf{x}} = \underline{\mathbf{b}} \quad \Leftrightarrow \quad \underline{\mathbf{x}} = A^{-1}\underline{\mathbf{b}} .$$

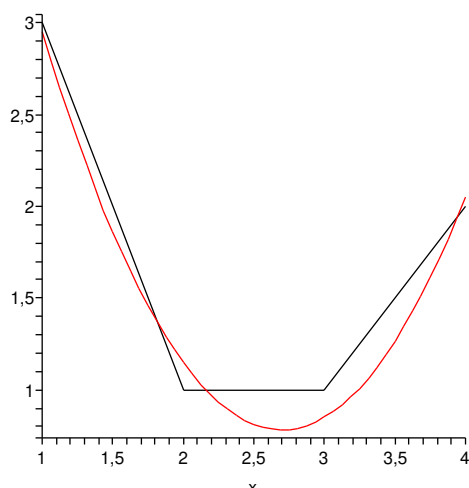


Abbildung 8.1: Die Ausgleichsparabel.

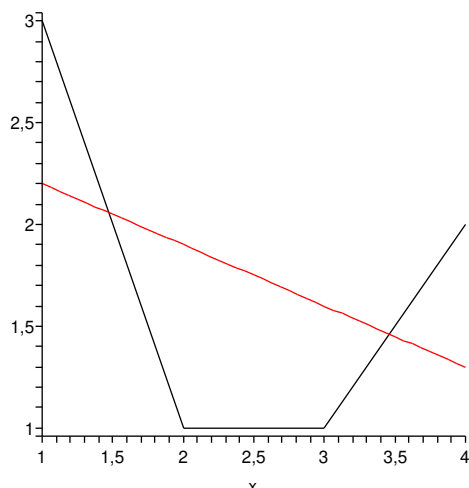


Abbildung 8.2: Die Ausgleichsgerade.

Insbesondere folgt aus der Existenz einer inversen Matrix $\operatorname{rg} A = n$.

Ist umgekehrt $\operatorname{rg} A = n$, so sind die n linearen Gleichungen

$$A\underline{\mathbf{x}} = \underline{\mathbf{e}}^{(i)}, \quad i = 1, 2, \dots, n,$$

$((\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \dots, \underline{\mathbf{e}}^{(n)})$ ist wie üblich die Standardbasis des \mathbb{R}^n) eindeutig lösbar, diese Lösungen seien mit $\tilde{\mathbf{a}}^{(i)}$, $i = 1, \dots, n$, bezeichnet. Es gilt dann

$$A \underbrace{(\tilde{\mathbf{a}}^{(1)} \quad \tilde{\mathbf{a}}^{(2)} \quad \dots \quad \tilde{\mathbf{a}}^{(n)})}_{=:\tilde{A}} = I_n .$$

Man sieht leicht ein, dass auch $\operatorname{rg} \tilde{A} = n$ gilt, es gibt demnach ebenso $\tilde{\tilde{A}}$ mit

$$\tilde{A}\tilde{\tilde{A}} = I_n .$$

Daraus folgt aber

$$\tilde{\tilde{A}}A = \tilde{\tilde{A}}AI_n = \tilde{\tilde{A}}A\tilde{A}\tilde{\tilde{A}} = \tilde{A} \underbrace{(A\tilde{A})}_{=I_n} \tilde{\tilde{A}} = \tilde{A}\tilde{\tilde{A}} = I_n ,$$

d.h.

$$\tilde{\tilde{A}}A = A\tilde{\tilde{A}} = I_n ,$$

und $\tilde{\tilde{A}} = A^{-1}$ ist die inverse Matrix, wobei die **Eindeutigkeit** der inversen Matrix ganz analog eingesehen werden kann.

Satz 8.5.1

*Eine Matrix $A \in M(n, n, \mathbb{R})$ besitzt genau dann eine inverse Matrix (**diese ist immer eindeutig bestimmt**), wenn A **maximalen Rang** hat, wenn also gilt $\operatorname{rg} A = n$. Dann heißt die Matrix **regulär** (oder nicht-singulär). Ist die Matrix nicht regulär, so heißt sie **singulär**.*

Bemerkung. Wie eine inverse Matrix mit Hilfe des Gaußschen Algorithmus bestimmt werden kann, wurde bereits am Ende des Paragraphen 8.3 angesprochen (vgl. Übungen).

Beispiel. Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2, 2, \mathbb{R}) .$$

Man verifiziert sofort

$$A \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = I_2$$

und die inverse Matrix existiert genau dann, wenn $ad - bc \neq 0$.

Diese Bedingung ist wiederum äquivalent dazu, dass die Vektoren

$$\begin{pmatrix} a \\ c \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} b \\ d \end{pmatrix}$$

linear unabhängig sind (Übung), was genau der Aussage des Satzes entspricht.

Einige Rechenregeln zu inversen Matrizen (die elementar nachzuprüfen sind, Übung) seien festgehalten in

Satz 8.5.2

Es seien $A, B \in M(n, n, \mathbb{R})$ invertierbar, d.h. es sei angenommen, dass die Inversen existieren. Dann gilt

- i) $(A^{-1})^{-1} = A$;*
- ii) $(A^T)^{-1} = (A^{-1})^T$;*
- iii) $(AB)^{-1} = B^{-1}A^{-1}$.*

Bemerkung. Wie bereits in Paragraph 8.1 diskutiert, ist in iii) sehr wohl auf die Reihenfolge zu achten, auch wenn die Matrix $A^{-1}B^{-1}$ gebildet werden kann

Ist A eine reguläre Matrix, so sind **Dreieckszerlegungen** der Matrix A von besonderem numerischen Interesse, insbesondere wenn man ein Gleichungssystem für verschiedene rechte Seiten zu lösen hat.

Unter einer **LR-Zerlegung** von A versteht man eine Zerlegung der Form

$$A = LR = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ l_{21} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ l_{n1} & l_{n2} & \dots & l_{n(n-1)} & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix}.$$

Wie eine solche Zerlegung analog zum Gaußschen Eliminationsverfahrens gewonnen werden kann und wie damit lineare Gleichungssysteme gelöst werden können, wird hier lediglich anhand eines Beispiels vorgestellt.

Beispiel. Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Wie üblich wird A mit Hilfe von elementaren Zeilenumformungen transformiert, man merkt sich jedoch die entsprechenden Vorfaktoren. D.h. von der zweiten Zeile wird dreimal die erste Zeile abgezogen, die **3 ist unten rot markiert**. Von der dritten Zeile wird die erste abgezogen, die **1 ist ebenfalls rot markiert**

$$\begin{pmatrix} 1 & 2 & 3 \\ \mathbf{3} & -4 & -8 \\ \mathbf{1} & -1 & -1 \end{pmatrix}.$$

Nun ist von der dritten Zeile ein viertel der zweiten Zeile abzuziehen, wobei die rot markierten Eintragungen unverändert bleiben, man merkt sich das viertel wiederum und gelangt zu

$$\begin{pmatrix} 1 & 2 & 3 \\ \mathbf{3} & -4 & -8 \\ \mathbf{1} & \mathbf{1/4} & \mathbf{1} \end{pmatrix}.$$

Die schwarzen Eintragungen bilden eine **rechte obere Dreiecksmatrix R** , die roten Eintragungen füllt man mit Einsen auf der Hauptdiagonalen auf. Es entsteht eine **linke untere Dreiecksmatrix L** mit der Eigenschaft

$$LR := \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 1 & 1/4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -4 & -8 \\ 0 & 0 & 1 \end{pmatrix} = A.$$

Mit der LR Zerlegung kann nun zunächst das Gleichungssystem

$$L\underline{\mathbf{y}} = \underline{\mathbf{b}}$$

durch **Vorwärtseinsetzen** gelöst werden.

Ist etwa $\underline{\mathbf{b}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, so ergibt das im obigen Beispiel

$$\begin{aligned} y_1 &= 1, \\ 3y_1 + y_2 &= 1, & \text{also } y_2 &= -2, \\ y_1 + \frac{1}{4}y_2 + y_3 &= 1, & \text{also } y_3 &= \frac{1}{2}. \end{aligned}$$

Anschließend ist durch **Rückwärtseinsetzen** das System

$$R\underline{\mathbf{x}} = \underline{\mathbf{y}}$$

zu lösen, insgesamt hat man

$$A\underline{\mathbf{x}} = LR\underline{\mathbf{x}} = L(R\underline{\mathbf{x}}) = L\underline{\mathbf{y}} = \underline{\mathbf{b}}.$$

Im Beispiel sieht das wie folgt aus:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ -4x_2 - 8x_3 &= -2 \\ x_3 &= \frac{1}{2}, \end{aligned}$$

und in der Tat ist

$$\underline{\mathbf{x}} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

die gesuchte Lösung des Systems $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ (Probe!).

Bemerkungen.

- i) Nicht jede reguläre Matrix besitzt eine solche Dreieckszerlegung, in der Regel wird man zunächst Zeilen vertauschen müssen, um auf der Hauptdiagonalen von R Eintragungen ungleich Null erhalten zu können.

Anstelle einer LR -Zerlegung von

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

wird etwa die LR -Zerlegung von

$$\tilde{A} = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

zu konstruieren sein.

- ii) Ist

$$D = \begin{pmatrix} r_{11} & 0 & \dots & 0 & 0 \\ 0 & r_{22} & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & r_{nn} \end{pmatrix},$$

$$\tilde{R} = \begin{pmatrix} 1 & \tilde{r}_{12} & \dots & \tilde{r}_{1(n-1)} & \tilde{r}_{1n} \\ 0 & 1 & \dots & \tilde{r}_{2(n-1)} & \tilde{r}_{2n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix},$$

wobei $\tilde{r}_{ij} = r_{ij}/r_{ii}$, $r_{ii} \neq 0$, $i, j = 1, \dots, n$, $i < j$, so ist $D\tilde{R} = R$.

Ist A zusätzlich als **symmetrisch** angenommen, d.h. $A = A^T$, so folgt

$$A = A^T = (LD\tilde{R})^T = \tilde{R}^T DL^T .$$

Nun ist aber \tilde{R}^T eine **normierte untere Dreiecksmatrix**, DL^T ist eine obere Dreiecksmatrix und aus der Eindeutigkeit der LR -Zerlegung folgt $\tilde{R}^T = L$ und $R = DL^T$.

Für symmetrische Matrizen bekommt die LR Zerlegung also eine besondere Gestalt,

$$A = LDL^T ,$$

(D wie oben **Diagonalmatrix**), die Zerlegung heißt **Cholesky-Zerlegung**.

8.6 Determinanten (Volumenfunktion; Transposition; Signum; Spat; Parallelotop; Regel von Sarrus; n -Linearform; alternierend; Multiplikationssatz; Streichungsmatrizen; Laplacescher Entwicklungssatz; Cramersche Regel)

Ebenso wie im vorherigen Paragraphen sind die Überlegungen hier **nur für quadratische Matrizen** zulässig, ein Analogon für $n \neq m$ gibt es nicht.

Zur Einführung sei zunächst der Fall $n = 2$ betrachtet, d.h. $A \in M(2, 2, \mathbb{R})$. Ist A regulär, so sind die beiden Spaltenvektoren

$$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

linear unabhängig. Das ist aber genau dann der Fall, wenn der **Flächeninhalt des von ihnen aufgespannten Parallelogramms nicht verschwindet**. Diese Zusammenhang soll jetzt näher analysiert werden, wobei als Erstes festzuhalten ist, welche Eigenschaften eine "natürliche" **Volumenfunktion** haben sollte.

Diese Volumenfunktion $\tau: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ soll zwei Spaltenvektoren ein **orientiertes Volumen** zuordnen, wobei die Vertauschung von zwei

Vektoren lediglich einen Vorzeichenwechsel zu bewirken hat, d.h. man fordert (vgl. Abbildung 8.3)

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = -\tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) . \quad (\text{V1})$$



Abbildung 8.3: $\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) > 0$ entspricht der auf der linken Seite angedeuteten Orientierung, $\tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = -\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) < 0$ entspricht der auf der rechten Seite angedeuteten Orientierung.

Als zweite ‘‘vernünftige’’ Bedingung ist sicherlich zu fordern, dass sich ein Volumen proportional zu einer Verlängerung einer Seite verhalten sollte, mit anderen Worten soll für alle $\lambda \in \mathbb{R}$

$$\tau(\lambda \underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) \quad (\text{V2})$$

erfüllt sein. Daraus folgt auch sofort

$$\tau(\underline{\mathbf{a}}^{(1)}, \lambda \underline{\mathbf{a}}^{(2)}) = -\tau(\lambda \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = -\lambda \tau(\underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(1)}) = \lambda \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) .$$

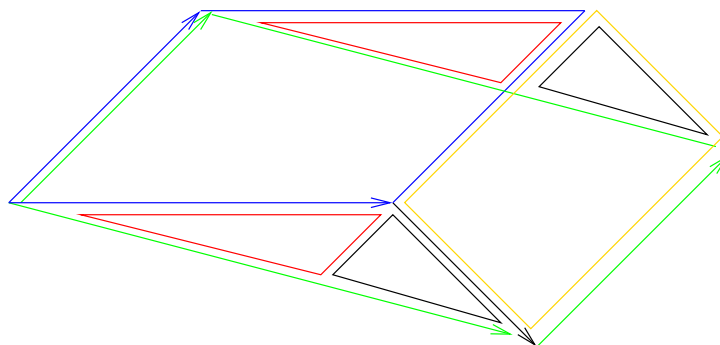
Als nächster Punkt ist zu klären, welches Ergebnis $\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ liefern sollte. In Abbildung 8.4 entspricht das dem Flächeninhalt des grün angedeuteten Parallelogramms. Der Flächeninhalt des blau angedeuteten Parallelogramms entspricht der Größe $\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)})$. Beachtet man nun, dass die rot bzw. schwarz angedeuteten Dreiecke jeweils gleich sind, so ist die Differenz der Flächeninhalt $\tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$ des golden skizzierten Parallelogramms.

Als dritte Forderung ist demnach zu stellen

$$\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)}) , \quad (\text{V3})$$

wobei sich als Folgerung wieder unmittelbar ergibt

$$\tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)} + \tilde{\underline{\mathbf{a}}}^{(2)}) = \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}) + \tau(\underline{\mathbf{a}}^{(1)}, \tilde{\underline{\mathbf{a}}}^{(2)}) .$$

Abbildung 8.4: Zur Berechnung von $\tau(\underline{\mathbf{a}}^{(1)} + \tilde{\underline{\mathbf{a}}}^{(1)}, \underline{\mathbf{a}}^{(2)})$.

Ist schließlich $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 , so normiert man

$$\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}) = 1, \quad (\text{V4})$$

genau wie es dem Flächeninhalt des Quadrates mit der Seitenlänge 1 entspricht.

Für beliebige Vektoren $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} \in \mathbb{R}^2$ folgt aus (V1)–(V4)

$$\begin{aligned} \tau\left(\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}\right) &= \tau(a_{11}\underline{\mathbf{e}}^{(1)} + a_{21}\underline{\mathbf{e}}^{(2)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &= a_{11}\tau(\underline{\mathbf{e}}^{(1)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &\quad + a_{21}\tau(\underline{\mathbf{e}}^{(2)}, a_{12}\underline{\mathbf{e}}^{(1)} + a_{22}\underline{\mathbf{e}}^{(2)}) \\ &= a_{11}a_{12}\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(1)}) + a_{11}a_{22}\tau(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}) \\ &\quad + a_{21}a_{12}\tau(\underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(1)}) + a_{21}a_{22}\tau(\underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(2)}) \\ &= a_{11}a_{22} - a_{12}a_{21}, \quad (*) \end{aligned}$$

insbesondere ist die Funktion τ **allein durch die Forderungen (V1)–(V4) eindeutig festgelegt**.

Das positive Vorzeichen des Terms $a_{11}a_{22}$ rührt daher, dass die blauen auf die roten Indizes mit der Permutation $(1, 2)$ abgebildet werden, es werden also keine Vertauschungen vorgenommen, $\underline{\mathbf{e}}^{(1)}$ und $\underline{\mathbf{e}}^{(2)}$ sind oben in der richtigen Reihenfolge aufgetreten.

Das negative Vorzeichen des Terms $a_{12}a_{21}$ kommt daher, dass die Permutation $(2, 1)$ durch eine Vertauschung (**Transposition**) aus der $(1, 2)$ entsteht, $\underline{\mathbf{e}}^{(1)}$ und $\underline{\mathbf{e}}^{(2)}$ sind sozusagen in der falschen Reihenfolge aufgetreten.

Wie hat eine analoge Konstruktion im Fall $n = 3$ auszusehen?

In drei Dimensionen gibt es nicht nur zwei Permutationen, es sind die sechs Permutationen (entsprechend der Reihenfolge der $\underline{\mathbf{e}}^{(i)}$, $i = 1, 2, 3$)

$$(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)$$

zu betrachten.

Dabei ist zu untersuchen, ob diese Permutationen durch eine gerade oder eine ungerade Anzahl von Vertauschungen von jeweils zwei Einträgen aus der $(1, 2, 3)$ entsteht. Man definiert das **Vorzeichen** (oder **Signum**) einer Permutation σ als 1 im Falle einer geraden Anzahl und als -1 im Falle einer ungeraden Anzahl (auch wenn verschiedene Vertauschungen zum Ziel führen, so ist das Signum eindeutig bestimmt). Notation: $\text{sgn}(\sigma)$.

Beispiel. Man betrachte $(3, 2, 1)$. Es wird etwa wie folgt vertauscht

$$(1, 2, 3) \rightarrow (1, 3, 2) \rightarrow (3, 1, 2) \rightarrow (3, 2, 1),$$

mit drei Vertauschungen ist das Ziel erreicht und das Signum ist in diesem Falle gleich -1 .

In obigen Permutationen ist das Signum (der Reihenfolge nach) $1, -1, -1, 1, 1, -1$, in Verallgemeinerung von (*) wird sich das orientierte Volumen eines **Spates** im \mathbb{R}^3 (vgl. Abbildung 8.5) (**Parallelotops** im \mathbb{R}^n), das von den Vektoren

$$\underline{\mathbf{a}}^{(1)} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \underline{\mathbf{a}}^{(2)} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \underline{\mathbf{a}}^{(3)} = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} \in \mathbb{R}^3$$

aufgespannt wird, berechnen zu

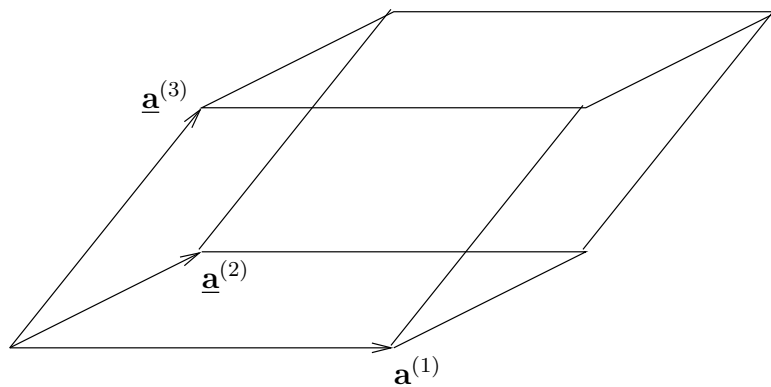


Abbildung 8.5: Der von $\underline{\mathbf{a}}^{(1)}$, $\underline{\mathbf{a}}^{(2)}$, $\underline{\mathbf{a}}^{(3)}$ aufgespannte Spat S , $S := \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = \alpha \underline{\mathbf{a}}^{(1)} + \beta \underline{\mathbf{a}}^{(2)} + \gamma \underline{\mathbf{a}}^{(3)}, \alpha, \beta, \gamma \in [0, 1]\}$.

$$\begin{aligned} \tau(\underline{\mathbf{a}}^{(1)}, \underline{\mathbf{a}}^{(2)}, \underline{\mathbf{a}}^{(3)}) &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ &\quad + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \quad (**) \end{aligned}$$

Die Spaltenvektoren fasst man nun in einer Matrix zusammen und definiert in Verallgemeinerung von (*) und (**)

Definition 8.6.1

Es sei $A = (a_{ij}) \in M(n, n, \mathbb{R})$. Dann heißt

$$\det A := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \sum_{\sigma} \operatorname{sgn}(\sigma) a_{1\sigma_1} \dots a_{n\sigma_n}$$

*die **Determinante** der Matrix A . Es ist dabei über alle unterschiedlichen Permutationen von $(1, 2, \dots, n)$ zu summieren.*

Bemerkungen.

- i) Die Determinante ist **nur für quadratische Matrizen** definiert.

- ii) Die Determinante ist genau die oben gesuchte Volumenfunktion.
- iii) Die Vorzeichen in (**) merkt man sich mit der **Regel von Sarrus**, die in Abbildung 8.6 veranschaulicht ist.

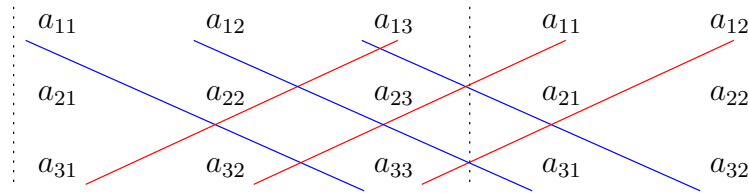


Abbildung 8.6: Zur Regel von Sarrus: **Positive** Vorzeichen hat man für die **blau unterstrichenen** Produkte, **negative** Vorzeichen für die **rot unterstrichenen**.

- iv) Aus der Definition folgt recht leicht $\det A = \det A^T$.
- v) Analog ist die Determinante für $A \in M(n, n, \mathbb{C})$ definiert. Ebenso wie in den vorherigen Paragraphen wird dies nicht weiter diskutiert, da dadurch keine zusätzlichen Einsichten gewonnen werden.
- vi) Nach den obigen Betrachtungen ist die Determinante vollständig charakterisiert durch
- $\det I_n = 1$;
 - die Determinante ist eine **n-Linearform**, d.h. in Abhängigkeit eines Zeilenvektors (Spaltenvektors) der Matrix ist (bei fixierten $(n-1)$ Zeilenvektoren (Spaltenvektoren)) die Determinante eine lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}$;
 - die Determinante ist **alternierend**, d.h. bei Vertauschung zweier Zeilen (Spalten) ändert sich das Vorzeichen.
- vii) Übung: Ändert sich die Determinante im Verlaufe des Gaußschen Eliminationsverfahrens?

Beispiele.

- i) Es sei $n = 3$ und

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Dann ist

$$\begin{aligned}\det A &= 1 \cdot 1 \cdot 1 + 2 \cdot 1 \cdot 0 + 3 \cdot 1 \cdot 1 - 3 \cdot 1 \cdot 0 - 1 \cdot 1 \cdot 1 - 2 \cdot 1 \cdot 1 \\ &= 1.\end{aligned}$$

ii) Die Linearität der Determinante in jeder Spalte bedeutet beispielsweise

$$\begin{aligned}\det \begin{pmatrix} \alpha x_1 + \beta y_1 & 1 & 3 \\ \alpha x_2 + \beta y_2 & 2 & 3 \\ \alpha x_3 + \beta y_3 & 1 & 1 \end{pmatrix} \\ = \alpha \det \begin{pmatrix} x_1 & 1 & 3 \\ x_2 & 2 & 3 \\ x_3 & 1 & 1 \end{pmatrix} + \beta \det \begin{pmatrix} y_1 & 1 & 3 \\ y_2 & 2 & 3 \\ y_3 & 1 & 1 \end{pmatrix}.\end{aligned}$$

Vorsicht. Die **Determinante ist selbst nicht linear**, d.h. im Allgemeinen gilt für $A, B \in M(n, n, \mathbb{R})$

$$\det(A + B) \neq \det A + \det B.$$

Ist $\lambda \in \mathbb{R}$, so gilt

$$\det(\lambda A) = \lambda^n \det A.$$

Als (**sehr wichtiges**) Korollar aus den zu Beginn dieses Kapitels angestellten Betrachtungen sei nun festgehalten:

Korollar 8.6.1

Es sei $A \in M(n, n, \mathbb{R})$. Dann sind die folgenden Aussagen äquivalent:

- i) Es ist $\det A \neq 0$.*
- ii) Die **Zeilenvektoren** von A sind **linear unabhängig**.*
- iii) Die **Spaltenvektoren** von A sind **linear unabhängig**.*
- iv) Die Matrix A ist **regulär**.*

Beispiel. Es folgt direkt, dass die Matrix A aus obigem Beispiel invertierbar ist (Übung: Man berechne die Inverse!).

In Paragraph 8.1 wurde zwar ein Beispiel gegeben, dass das Produkt zweier Matrizen durchaus auch dann die Nullmatrix ergeben kann, wenn keine der beiden Matrizen selbst eine Nullmatrix ist. Für reguläre Matrizen kann man eine solche Situation ausschließen, wie der [Determinantenmultiplikationssatz](#) zeigt.

Satz 8.6.1

Es seien $A, B \in M(n, n, \mathbb{R})$. dann gilt

$$\det (AB) = \det A \det B .$$

Sind also A und B regulär, d.h. $\det A, \det B \neq 0$, so folgt $\det (AB) \neq 0$, AB ist auch regulär. Ist umgekehrt AB regulär, so müssen auch A und B regulär sein.

Wegen

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

folgt sofort aus Satz 8.6.1

Korollar 8.6.2

*Ist $A \in M(n, n, \mathbb{R})$ eine **reguläre Matrix**, so gilt*

$$\det A^{-1} = \frac{1}{\det A} .$$

Die Definition 8.6.1 ist natürlich insbesondere für große n zur Berechnung der Determinante nicht geeignet. Für praktische Zwecke benötigt man so genannte [Streichungsmatrizen](#) A_{ij} von A , die für $i, j = 1, \dots, n$, aus A durch die Streichung der i^{ten} Zeile und der j^{ten} Spalte hervorgehen. Es handelt sich um $(n - 1) \times (n - 1)$ Matrizen

$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

Abbildung 8.7: Die i^{te} Zeile und die j^{te} Spalte werden aus A entfernt, um die Streichungsmatrix A_{ij} zu erhalten.

(vgl. Abbildung 8.7).

Mit Hilfe der Streichungsmatrizen kann formuliert werden:

Satz 8.6.2 (*Laplacescher⁴ Entwicklungssatz*)

Für $A \in M(n, n, \mathbb{R})$ und $1 \leq i, j \leq n$ bezeichne $A_{ij} \in M(n-1, n-1, \mathbb{R})$ die entsprechende Streichungsmatrix. Dann gilt (*Entwicklung nach der i^{ten} Zeile*)

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

sowie (*Entwicklung nach der j^{ten} Spalte*)

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

Beispiele.

i) Es sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix} \in M(3, 3, \mathbb{R}) .$$

⁴P.M. Marquis de Laplace, 1749-1827; Paris.

Entwicklung nach der ersten Zeile liefert

$$\begin{aligned} \det A &= 1 \begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} \\ &= 0 - 4 + 6 = 2. \end{aligned}$$

Entwicklung nach der zweiten Spalte liefert

$$\begin{aligned} \det A &= -2 \begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} + 2 \begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} - 1 \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} \\ &= -4 + 2 + 4 = 2. \end{aligned}$$

ii) Es sei

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \in M(4, 4, \mathbb{R}).$$

Bei der Entwicklung von $\det A$ wird man **die Nullen in der ersten Spalte** ausnutzen, d.h. nach der ersten Spalte entwickeln. Dies ergibt (vgl. Beispiel *i*)

$$\det A = 1 \begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 2 \\ 0 & 1 & 1 \end{vmatrix} = 2.$$

Zum Abschluss des Paragraphen sei noch eine geschlossene Formel zur Lösung eines linearen Gleichungssystems angegeben.

Satz 8.6.3 (*Cramersche Regel*)

Es sei $A = (\underline{\mathbf{a}}^{(1)} \ \underline{\mathbf{a}}^{(2)} \ \dots \ \underline{\mathbf{a}}^{(n)}) \in M(n, n, \mathbb{R})$ regulär und $\underline{\mathbf{b}} \in \mathbb{R}^n$, das Gleichungssystem $A\underline{\mathbf{x}} = \underline{\mathbf{b}}$ ist also insbesondere eindeutig lösbar. Für alle $j = 1, \dots, n$ ist die j^{te} Komponente des Lösungsvektors gegeben durch

$$x_j = \frac{\det(\underline{\mathbf{a}}^{(1)} \ \dots \ \underline{\mathbf{a}}^{(j-1)} \ \underline{\mathbf{b}} \ \underline{\mathbf{a}}^{(j+1)} \ \underline{\mathbf{a}}^{(n)})}{\det A}.$$

Bemerkung. Da die Cramersche Regel hauptsächlich für theoretische Untersuchungen von Belang ist (die tatsächliche Berechnung ist meist zu aufwendig), wird hier auf eine weitere Diskussion verzichtet.

8.7 Vektorprodukt, Spatprodukt, Ebenen und Geraden im \mathbb{R}^3 (Orientierung; rechte-Hand-Regel; Normalenvektor; Richtungsvektor; Hessesche Normalform; Abstand zwischen zwei Mengen)

In diesem Paragraphen wird ausschließlich der \mathbb{R}^3 betrachtet (oder der \mathbb{R}^2 als Unterraum des \mathbb{R}^3). **Ein Analogon zum Vektorprodukt in höheren Dimensionen gibt es nicht.** Es bezeichne weiter $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}, \underline{\mathbf{e}}^{(3)})$ stets die kanonische Basis des \mathbb{R}^3 .

Definition 8.7.1

Das **Vektorprodukt** \times im \mathbb{R}^3 ist die Abbildung, die allen $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^3$ einen weiteren **Vektor im \mathbb{R}^3 zuordnet,**

$$\underline{\mathbf{x}} \times \underline{\mathbf{y}} := \underline{\mathbf{e}}^{(1)} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - \underline{\mathbf{e}}^{(2)} \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + \underline{\mathbf{e}}^{(3)} \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}.$$

Merkregel. **Formal** kann das Vektorprodukt geschrieben werden als

$$\begin{vmatrix} \underline{\mathbf{e}}^{(1)} & \underline{\mathbf{e}}^{(2)} & \underline{\mathbf{e}}^{(3)} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix},$$

wobei die Determinante als **Entwicklung nach der ersten Zeile** zu verstehen ist.

Interpretation. Zur Interpretation nehme man o.E. an, dass die zwei Vektoren $\underline{\mathbf{x}}$ und $\underline{\mathbf{y}}$ in der $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ Ebene liegen, d.h. es gelte

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix}.$$

Dann ist

$$\underline{\mathbf{x}} \times \underline{\mathbf{y}} = \underline{\mathbf{e}}^{(3)} \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}.$$

Zum Einen liefert also $\|\underline{\mathbf{x}} \times \underline{\mathbf{y}}\|$ nach Paragraph 8.6 genau den **Flächeninhalt** (größer oder gleich Null, d.h. nicht orientiert) **des von $\underline{\mathbf{x}}$ und $\underline{\mathbf{y}}$ aufgespannten Parallelogramms**.

Andererseits steht $\underline{\mathbf{x}} \times \underline{\mathbf{y}}$ **senkrecht auf $\underline{\mathbf{x}}$ und $\underline{\mathbf{y}}$** .

Schließlich ist $(\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{x}} \times \underline{\mathbf{y}})$ **positiv orientiert** (falls nur $\underline{\mathbf{x}}$ und $\underline{\mathbf{y}}$ linear unabhängig sind, im Falle der linearen Abhängigkeit ist $\underline{\mathbf{x}} \times \underline{\mathbf{y}} = \underline{\mathbf{0}}$), das bedeutet:

Definition 8.7.2

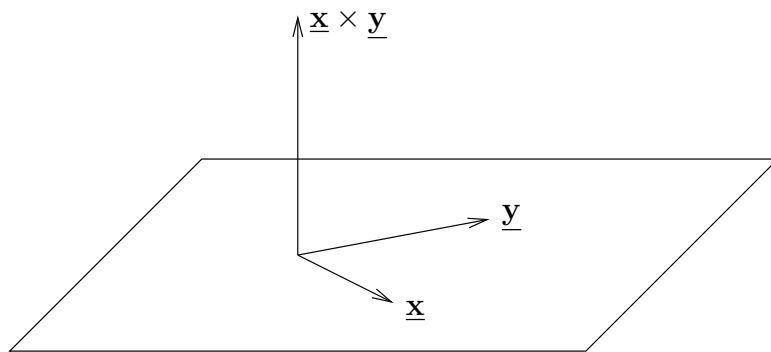
*Es sei $(\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}})$ eine Basis des \mathbb{R}^3 . Dann heißt $(\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}})$ **positiv orientiert**, falls*

$$\det(\underline{\mathbf{x}} \ \underline{\mathbf{y}} \ \underline{\mathbf{z}}) > 0,$$

*andernfalls heißt $(\underline{\mathbf{x}}, \underline{\mathbf{y}}, \underline{\mathbf{z}})$ **negativ orientiert**.*

Bemerkungen und Beispiele.

- i)* Es kommt **wesentlich auf die Reihenfolge der Vektoren an**.
- ii)* $(\underline{e}^{(1)}, \underline{e}^{(2)}, \underline{e}^{(3)})$ ist positiv orientiert.
- iii)* $(\underline{e}^{(2)}, \underline{e}^{(1)}, \underline{e}^{(3)})$ ist dahingegen negativ orientiert.
- iv)* Die positive Orientierung (allgemein und insbesondere) von $(\underline{x}, \underline{y}, \underline{x} \times \underline{y})$ (im Falle der linearen Unabhängigkeit von \underline{x} und \underline{y}) kann man sich mit Hilfe der **rechte-Hand-Regel** merken: Zeigt der Daumen der rechten Hand in Richtung von \underline{x} , der Zeigefinger in Richtung von \underline{y} , so zeigt der Mittelfinger senkrecht dazu in Richtung von $\underline{x} \times \underline{y}$ (vgl. Abbildung 8.8).

Abbildung 8.8: Zum Vektorprodukt im \mathbb{R}^3 .

- v)* Als Anwendungen seien hier die Kraft auf einen stromdurchflossenen Leiter im Magnetfeld, das Drehmoment ... erwähnt.
- vi)* Eine weitere wichtige Anwendung (vgl. die nachfolgende Diskussion von Ebenen im \mathbb{R}^3) ist es, zwei linear unabhängige Vektoren \underline{x} und \underline{y} durch den dazu senkrechten Vektor $\underline{x} \times \underline{y}$ zu einer Basis des \mathbb{R}^3 zu ergänzen.
- vii)* In Paragraph 8.6 wurde gezeigt, dass das orientierte Volumen des von \underline{x} , \underline{y} und \underline{z} im \mathbb{R}^3 aufgespannten Spates gegeben ist durch

$$\det(\underline{x} \ \underline{y} \ \underline{z}) =: [\underline{x}, \underline{y}, \underline{z}].$$

Diese Größe heißt auch das **Spatprodukt** der Vektoren $\underline{x}, \underline{y}, \underline{z} \in \mathbb{R}^3$, das mit Hilfe des Vektorproduktes geschrieben werden kann als

$$[\underline{x}, \underline{y}, \underline{z}] = \langle \underline{x} \times \underline{y}, \underline{z} \rangle.$$

Aus der Definition des Vektorproduktes lässt sich leicht ableiten:

Eigenschaften des Vektorproduktes. Für $\underline{x}, \underline{y}, \underline{z} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$ gilt:

$$i) \quad \underline{x} \times \underline{y} = -(\underline{y} \times \underline{x});$$

$$ii) \quad (\lambda \underline{x}) \times \underline{y} = \lambda(\underline{x} \times \underline{y});$$

$$iii) \quad \underline{x} \times (\underline{y} + \underline{z}) = \underline{x} \times \underline{y} + \underline{x} \times \underline{z}.$$

Vorsicht. Das Vektorprodukt ist **im Allgemeinen nicht assoziativ**, d.h.

$$\underline{x} \times (\underline{y} \times \underline{z}) \neq (\underline{x} \times \underline{y}) \times \underline{z} \quad \text{im Allgemeinen.}$$

Dies sieht man etwa über den **Entwicklungssatz**

$$\underline{x} \times (\underline{y} \times \underline{z}) = \langle \underline{x}, \underline{z} \rangle \underline{y} - \langle \underline{x}, \underline{y} \rangle \underline{z}$$

ein.

Mit Hilfe des Vektorproduktes sollen nun kurz **Ebenen im \mathbb{R}^3** diskutiert werden.

Eine Ebene **durch den Ursprung** kann einerseits interpretiert werden als eine Menge der Form

$$\{\underline{N}\}^\perp, \quad \underline{N} \neq \underline{0},$$

also als die Menge aller Punkte $\underline{x} \in \mathbb{R}^3$, die senkrecht auf dem **Normalenvektor** \underline{N} stehen (vgl. Abbildung 6.7).

Andererseits (vgl. Abbildung 6.4) kann sie aufgefasst werden als die Menge

$$\text{Spann}(\underline{v}, \underline{w}), \quad \underline{v}, \underline{w} \text{ linear unabhängig,}$$

d.h. als die Menge, die von den **Richtungsvektoren** \underline{v} und \underline{w} aufgespannt wird.

Beide Interpretationen sind äquivalent. Ergänzt man etwa \underline{N} zu einer Orthogonalbasis des \mathbb{R}^3 , so sind die weiteren Basisvektoren linear unabhängige Richtungsvektoren der Ebene.

Sind umgekehrt $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ Richtungsvektoren in einer Ebene, so ist

$$\underline{\mathbf{N}} := \underline{\mathbf{v}} \times \underline{\mathbf{w}}$$

ein Normalenvektor. Natürlich ist jedes Vielfache $\lambda \underline{\mathbf{N}}$, $\lambda \neq 0$, ebenso ein Normalenvektor an die Ebene.

Eine Ebene, die **nicht notwendig durch den Ursprung** verläuft, ist etwa durch drei **nicht kollineare** Punkte (liegen nicht auf einer Geraden) $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$ auf der Ebene bestimmt. Mit $\underline{\mathbf{v}} = \underline{\mathbf{b}} - \underline{\mathbf{a}}$, $\underline{\mathbf{w}} = \underline{\mathbf{c}} - \underline{\mathbf{a}}$ (diese Vektoren sind linear unabhängig, da drei nicht kollineare Punkte gewählt waren) ist diese Ebene die Menge gegeben durch die **Parameterdarstellung** (vgl. Abbildung 8.9)

$$E = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = \underline{\mathbf{a}} + \lambda \underline{\mathbf{v}} + \mu \underline{\mathbf{w}}, \lambda, \mu \in \mathbb{R} \} .$$

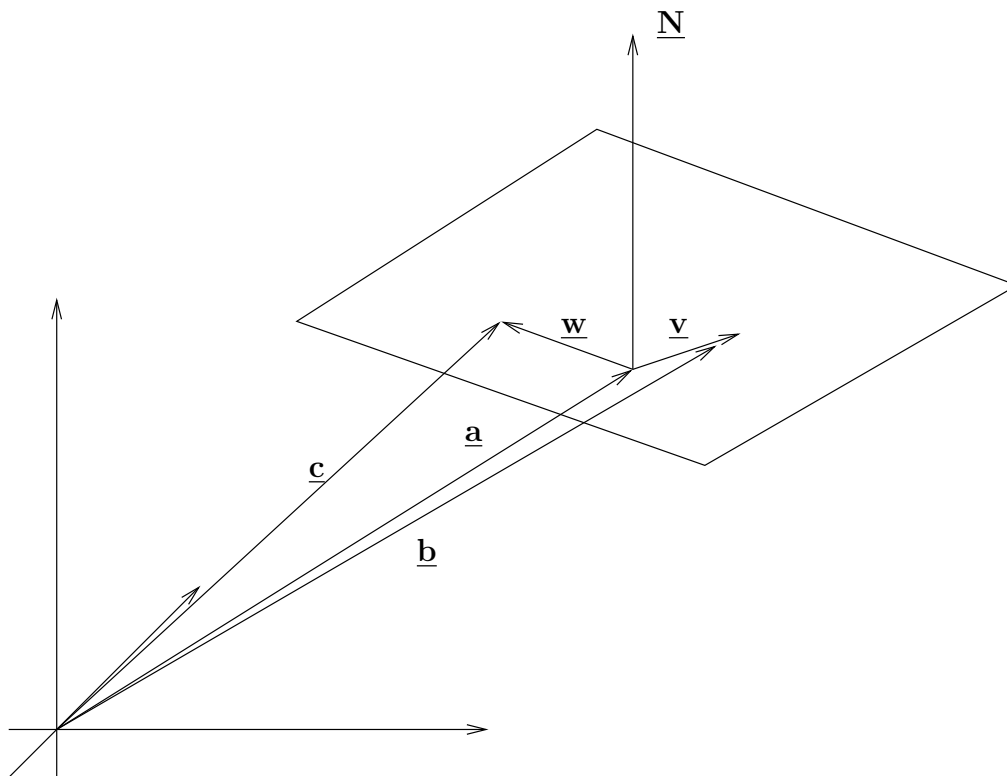


Abbildung 8.9: Die Ebene durch die Punkte $\underline{\mathbf{a}}$, $\underline{\mathbf{b}}$, $\underline{\mathbf{c}}$.

Wieder ist mit $\underline{\mathbf{v}} \times \underline{\mathbf{w}}$ ein Normalenvektor an die Ebene bestimmt. Dieser wird auf die Länge 1 normiert und mit einem geeigneten Vor-

zeichen versehen, d.h. man setzt

$$\underline{\mathbf{N}} = \pm \frac{\underline{\mathbf{v}} \times \underline{\mathbf{w}}}{\|\underline{\mathbf{v}} \times \underline{\mathbf{w}}\|},$$

insbesondere gilt $\|\underline{\mathbf{N}}\| = 1$.

War die Ebene durch den Ursprung noch durch $\langle \underline{\mathbf{N}}, \underline{\mathbf{x}} \rangle = 0$ gekennzeichnet, so bewirkt die Verschiebung nun

$$E = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \langle \underline{\mathbf{N}}, \underline{\mathbf{x}} \rangle - p = 0\},$$

wobei das Vorzeichen von N so gewählt wird, dass $p \geq 0$. Dann spricht man von der **Hesseschen⁵ Normalform** der Ebene.

Bemerkungen.

- i) Die beiden Vektoren $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ bilden eine Basis des Kerns von $\underline{\mathbf{N}}^T \in M(1, 3, \mathbb{R})$, $\underline{\mathbf{a}}$ ist eine spezielle Lösung des Gleichungssystems $\underline{\mathbf{N}}^T \underline{\mathbf{x}} = p$ (vgl. Satz 8.2.1).
- ii) Der Abstand d eines beliebigen Punktes $\underline{\mathbf{x}}^{(0)} \in \mathbb{R}^3$ zu einer Ebene errechnet sich aus der Hesseschen Normalform zu

$$d = |\langle \underline{\mathbf{N}}, \underline{\mathbf{x}}^{(0)} \rangle - p|.$$

Dabei ist der Abstand zwischen zwei Mengen $A, B \in \mathbb{R}^N$ definiert als $\inf_{\underline{\mathbf{x}} \in A} (\inf_{\underline{\mathbf{y}} \in B} \|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|)$.

Beispiel. Eine Ebene sei durch die Punkte

$$\underline{\mathbf{a}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{b}} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{c}} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

festgelegt. Dann ist

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \underline{\mathbf{N}}$$

⁵L.O. Hesse, 1811-1874; Königsberg, Halle, Heidelberg, München.

normal zu den Richtungsvektoren $\underline{\mathbf{b}} - \underline{\mathbf{a}}$, $\underline{\mathbf{c}} - \underline{\mathbf{a}}$ und bereits auf die Länge 1 normiert. Wegen $\langle \underline{\mathbf{N}}, \underline{\mathbf{a}} \rangle = 1 = p$ ist dabei auch das richtige Vorzeichen gewählt, die Normalform lautet

$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \underline{\mathbf{x}} \right\rangle - 1 = 0$$

(Probe!).

Schneiden sich zwei ungleiche Ebenen im \mathbb{R}^3 , so erhält man eine Gerade im \mathbb{R}^3 , die auch interpretiert werden kann als die allgemeine Lösung des Gleichungssystems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \end{aligned}$$

das den maximalen Rang zwei habe.

Geraden im \mathbb{R}^2 können ebenso auf eine Hessesche Normalform gebracht werden, d.h. man findet $\underline{\mathbf{n}} \in \mathbb{R}^2$ mit $\|\underline{\mathbf{n}}\| = 1$ und ein $p \geq 0$, sodass die Gerade in der Form

$$\langle \underline{\mathbf{n}}, \underline{\mathbf{x}} \rangle - p = 0$$

geschrieben werden kann. Beispiele werden in den Übungen diskutiert.

Kapitel 9

Lineare Abbildungen

Konkrete Aufgabenstellungen und Modellbildungen in den Anwendungen sind in der Regel von nichtlinearer Struktur und dementsprechend schwer lösbar. Um dennoch Vorhersagen treffen zu können, wird das Problem meist **linearisiert** (für kleine Auslenkungen setzt man $\sin(x) \approx x$ etc.), die mathematische Behandlung wird wesentlich einfacher und innerhalb gewisser Bereiche kann der entstandene Fehler abgeschätzt werden.

Aber auch Drehungen und weitere wichtige Operationen können als lineare Abbildungen interpretiert werden: Eine Rotation im \mathbb{R}^2 wird z.B. beschrieben durch die lineare Abbildung

$$\mathbb{R}^2 \ni \underline{\mathbf{x}} \mapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \underline{\mathbf{x}},$$

wobei der Drehwinkel φ fixiert ist.

Allgemeiner gilt: Ist $\underline{\mathbf{x}} \in \mathbb{R}^m$ (Spaltenvektor bzgl. der kanonischen Basis) und $A \in M(n, m, \mathbb{R})$, so ist die Funktion

$$\underline{\mathbf{x}} \mapsto \underline{\mathbf{y}} = A\underline{\mathbf{x}}$$

eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^n .

Lineare Abbildungen vertragen sich mit linearen Räumen, also Vektorräumen, die in Kapitel 6.1 behandelt wurden. Was das genau bedeutet, wird in Definition 9.1.1 deutlich.

9.1 Definition und erste Eigenschaften (Kern und Bild einer linearen Abbildung; Rang; Rangformel; Injektivität und Surjektivität einer linearen Abbildung)

Definition 9.1.1

Es seien V und W zwei Vektorräume über einem Körper \mathbb{K} . Eine Abbildung

$$L : V \rightarrow W$$

*heißt **lineare Abbildung (Homomorphismus)**, falls für alle $\underline{\mathbf{v}}, \underline{\mathbf{w}} \in V$ und für alle $\lambda \in \mathbb{K}$ gilt:*

$$\begin{aligned} L(\underline{\mathbf{v}} + \underline{\mathbf{w}}) &= L(\underline{\mathbf{v}}) + L(\underline{\mathbf{w}}) , \\ L(\lambda \underline{\mathbf{v}}) &= \lambda L(\underline{\mathbf{v}}) . \end{aligned}$$

Bemerkung. Zu beachten ist, dass auf der linken Seite mit “+” die Addition in V , auf der rechten Seite die in W gemeint ist (Multiplikation mit Skalar analog).

Beispiele.

i) Das einfachste Beispiel ist die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = ax , \quad a \in \mathbb{R} \text{ fixiert .}$$

Zu beachten ist aber, dass mit $b \neq 0$

$$f(x) = ax + b$$

keine lineare Abbildung im Sinne von Definition 9.1.1 ist¹.

ii) Drehungen und Spiegelungen sind lineare Abbildungen (Übungsaufgabe).

¹Die Gerade $f(x) = ax + b$ wird auch als **affin lineare Funktion** bezeichnet.

Im Folgenden werden lineare Abbildungen von einem Vektorraum V mit $\dim V = m$ in einen Vektorraum W mit $\dim W = n$ studiert. In der Regel wird o.E. die Situation $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^m$, $W = \mathbb{R}^n$ betrachtet mit der Schreibweise $M(n, m, \mathbb{R}) = M(n, m)$. Streckenweise werden aber auch andere Beispiele diskutiert.

Aus der Definition linearer Abbildungen folgen unmittelbar die ersten beiden Aussagen des folgenden Satzes (vgl. Übungen).

Satz 9.1.1

Es sei $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung.

i) Dann ist der **Kern** der Abbildung L ,

$$\text{kern } L := L^{-1}(\{\underline{\mathbf{0}}\}) := \{\underline{\mathbf{x}} \in \mathbb{R}^m : L(\underline{\mathbf{x}}) = \underline{\mathbf{0}}\},$$

ein **Unterraum** des \mathbb{R}^m .

ii) Ebenso ist das **Bild** der Abbildung L ,

$$\text{bild } L := \{\underline{\mathbf{y}} \in \mathbb{R}^n : \underline{\mathbf{y}} = L(\underline{\mathbf{x}}) \text{ für ein } \underline{\mathbf{x}} \in \mathbb{R}^m\},$$

ein **Unterraum** des \mathbb{R}^n .

iii) Als **Rang** der Abbildung L definiert man

$$\text{rg } L := \dim \text{bild } L.$$

Damit folgt der **Rangsatz** für lineare Abbildungen

$$m = \dim \text{kern } L + \text{rg } L.$$

Bemerkungen und Beispiele.

i) Man betrachte die lineare Abbildung $L: \mathbb{R}^2 \rightarrow \mathbb{R}$, mit $L(\underline{\mathbf{x}}) = x_1 - x_2$. Dann ist $\underline{\mathbf{x}} \in \text{kern } L$ genau dann, wenn gilt $x_1 = x_2$, d.h.

$$\text{kern } L = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

In dem Beispiel ist $\text{Bild } L = \mathbb{R}$. Dies sieht man direkt oder über den Rangsatz ein.

ii) Der Zusammenhang zwischen dem Kern einer linearen Abbildung und dem Kern einer Matrix (Satz 8.2.1) ist evident, der Zusammenhang zwischen dem Rang einer linearen Abbildung und dem Rang einer Matrix sowie der Rangsatz werden in Kürze deutlich.

iii) Um den Rangsatz einzusehen, überlege man sich die folgenden Fragen. Es sei dazu L eine lineare Abbildung vom \mathbb{R}^m in den \mathbb{R}^n .

(a) Ist es möglich, dass zwei linear abhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear unabhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden?

(b) Ist es möglich, dass zwei linear unabhängige Vektoren $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in \mathbb{R}^m$ auf zwei linear abhängige Vektoren $\underline{\mathbf{u}}, \underline{\mathbf{v}} \in \mathbb{R}^n$ abgebildet werden? Wenn ja, hat das Konsequenzen für Kern L ?

iv) Eine Folgerung aus dem Rangsatz ist: Es sei $n = m$ und $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ sei eine lineare Abbildung. Dann ist L **injektiv** (d.h. $\dim \text{Kern } L = 0$) **genau dann, wenn L surjektiv ist** (d.h. $\text{rg } L = n$). Diese Folgerung und weitere Bemerkungen zur Injektivität bzw. Surjektivität linearer Abbildungen werden in den Übungen diskutiert.

Im Fall einer linearen Abbildung $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = ax$, $a \in \mathbb{R}$ fixiert, ist die Abbildung allein durch den Parameter a charakterisiert.

Wodurch ist eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ bestimmt?

Mit dieser und ähnlichen Fragen beschäftigt sich der nächste Abschnitt.

9.2 Matrixdarstellung (Koordinaten; darstellende Matrix; Basiswechsel; Koordinatentransformation; Transformation der darstellenden Matrix; Dualraum; lineares Funktional; duale Basis; Bidualraum; Tensor; kovariant; kontravariant; Tensorprodukt)

Betrachtet sei eine Basis $(v^{(1)}, \dots, v^{(m)})$ des \mathbb{R}^m und eine Basis $(w^{(1)}, \dots, w^{(n)})$ des \mathbb{R}^n . Damit ist jedes $\underline{\mathbf{x}} \in \mathbb{R}^m$ darstellbar als

$$\underline{\mathbf{x}} = \sum_{j=1}^m \alpha_j \underline{\mathbf{v}}^{(j)}, \quad \alpha_j \in \mathbb{R},$$

und für eine beliebige lineare Abbildung L folgt nach Definition 9.1.1

$$L(\underline{\mathbf{x}}) = L\left(\sum_{j=1}^m \alpha_j \underline{\mathbf{v}}^{(j)}\right) = \sum_{j=1}^m \alpha_j L(\underline{\mathbf{v}}^{(j)}).$$

Demnach ist $L(\underline{\mathbf{x}})$ für alle $\underline{\mathbf{x}} \in \mathbb{R}^m$ festgelegt, wenn nur die Bilder $L(\underline{\mathbf{v}}^{(j)})$, $j = 1, \dots, m$, irgendeiner Basis des \mathbb{R}^m bekannt sind.

Weiter kann man für jedes fixierte $1 \leq j \leq m$ schreiben:

$$L(\underline{\mathbf{v}}^{(j)}) = \sum_{i=1}^n a_{ij} \underline{\mathbf{w}}^{(i)}, \quad a_{ij} \in \mathbb{R}.$$

Dabei hängen die a_{ij} von den speziell gewählten Basen ab.

Beispiele.

i) Es sei $(\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)})$ die kanonische Basis des \mathbb{R}^2 , $(\underline{\mathbf{f}}^{(1)}, \underline{\mathbf{f}}^{(2)}, \underline{\mathbf{f}}^{(3)})$ sei die kanonische Basis des \mathbb{R}^3 . Weiterhin sei $L: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ eine lineare Abbildung und es gelte (wie bereits gesagt, ist durch diese Wahl die lineare Abbildung eindeutig bestimmt)

$$L(\underline{\mathbf{e}}^{(1)}) = \underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(3)} =: \underline{\mathbf{y}}, \quad L(\underline{\mathbf{e}}^{(2)}) = \underline{\mathbf{f}}^{(1)} + 2\underline{\mathbf{f}}^{(2)} =: \underline{\mathbf{z}}.$$

Mit obiger Notation ist

$$a_{11} = 1, \quad a_{21} = 0, \quad a_{31} = 1, \quad a_{12} = 1, \quad a_{22} = 2, \quad a_{32} = 0.$$

Bezüglich der kanonischen Basis des \mathbb{R}^2 wird ein beliebiges $\underline{\mathbf{x}} \in \mathbb{R}^2$ geschrieben als

$$\underline{\mathbf{x}} = x_1 \underline{\mathbf{e}}^{(1)} + x_2 \underline{\mathbf{e}}^{(2)} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

die Vektoren $\underline{\mathbf{y}}, \underline{\mathbf{z}} \in \mathbb{R}^3$ lauten (Koordinaten bzgl. der kanonischen Basis des \mathbb{R}^3):

$$\underline{\mathbf{y}} = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \underline{\mathbf{z}} = \begin{pmatrix} a_{21} \\ a_{22} \\ a_{32} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Bildet man aus diesen Spaltenvektoren eine Matrix A , so folgt für alle $\underline{\mathbf{x}} \in \mathbb{R}^2$

$$L(\underline{\mathbf{x}}) = x_1 L(\underline{\mathbf{e}}^{(1)}) + x_2 L(\underline{\mathbf{e}}^{(2)}) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Die lineare Abbildung L kann mit einer Matrix A identifiziert werden. Die Bilder der Basis des \mathbb{R}^2 werden als Spaltenvektoren im \mathbb{R}^3 geschrieben, und aus diesen Spaltenvektoren setzt sich die Matrix A zusammen.

ii) Man betrachte die gleiche lineare Abbildung wie im ersten Beispiel, als Basis des \mathbb{R}^2 diene aber jetzt $\mathcal{G} := (\underline{\mathbf{g}}^{(1)}, \underline{\mathbf{g}}^{(2)})$,

$$\underline{\mathbf{g}}^{(1)} = \underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}, \quad \underline{\mathbf{g}}^{(2)} = \underline{\mathbf{e}}^{(1)} - \underline{\mathbf{e}}^{(2)},$$

als Basis des \mathbb{R}^3 werde $\mathcal{H} = (\underline{\mathbf{h}}^{(1)}, \underline{\mathbf{h}}^{(2)}, \underline{\mathbf{h}}^{(3)})$,

$$\underline{\mathbf{h}}^{(1)} = \underline{\mathbf{f}}^{(1)} + \underline{\mathbf{f}}^{(3)}, \quad \underline{\mathbf{h}}^{(2)} = \underline{\mathbf{f}}^{(2)} + \underline{\mathbf{f}}^{(3)}, \quad \underline{\mathbf{h}}^{(3)} = \underline{\mathbf{f}}^{(3)},$$

gewählt. Man berechnet

$$\begin{aligned} L(\underline{\mathbf{g}}^{(1)}) &= L(\underline{\mathbf{e}}^{(1)} + \underline{\mathbf{e}}^{(2)}) = L(\underline{\mathbf{e}}^{(1)}) + L(\underline{\mathbf{e}}^{(2)}) \\ &= 2\underline{\mathbf{f}}^{(1)} + 2\underline{\mathbf{f}}^{(2)} + \underline{\mathbf{f}}^{(3)} =: \underline{\mathbf{v}} \end{aligned}$$

und analog

$$L(\underline{\mathbf{g}}^{(2)}) = \underline{\mathbf{f}}^{(3)} - 2\underline{\mathbf{f}}^{(2)} =: \underline{\mathbf{w}}.$$

Nun sollen $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ bzgl. der neuen Basis \mathcal{H} dargestellt werden. Dazu beachtet man

$$\begin{aligned}\underline{\mathbf{f}}^{(1)} &= \underline{\mathbf{h}}^{(1)} - \underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{f}}^{(2)} &= \underline{\mathbf{h}}^{(2)} - \underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{f}}^{(3)} &= \underline{\mathbf{h}}^{(3)}.\end{aligned}$$

Eingesetzt ergibt dies

$$\begin{aligned}\underline{\mathbf{v}} &= 2\underline{\mathbf{h}}^{(1)} + 2\underline{\mathbf{h}}^{(2)} - 3\underline{\mathbf{h}}^{(3)}, \\ \underline{\mathbf{w}} &= 3\underline{\mathbf{h}}^{(3)} - 2\underline{\mathbf{h}}^{(2)}.\end{aligned}$$

Schließlich wird für beliebiges $\underline{\mathbf{x}} \in \mathbb{R}^2$ bzw. für beliebiges $\underline{\mathbf{u}} \in \mathbb{R}^3$ die **Koordinatendarstellung bzgl. der Basis \mathcal{G} bzw. bzgl. der Basis \mathcal{H}** eingeführt, d.h.

$$\begin{aligned}\underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{g}}^{(1)} + \alpha_2 \underline{\mathbf{g}}^{(2)} \cong \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}}, \\ \underline{\mathbf{u}} &= \beta_1 \underline{\mathbf{h}}^{(1)} + \beta_2 \underline{\mathbf{h}}^{(2)} + \beta_3 \underline{\mathbf{h}}^{(3)} \cong \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}_{\mathcal{H}}.\end{aligned}$$

Insbesondere gilt (wenn ein Vektor mit seiner Koordinatendarstellung identifiziert wird)

$$\underline{\mathbf{g}}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\mathcal{G}}, \quad \underline{\mathbf{g}}^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_{\mathcal{G}}$$

sowie

$$\underline{\mathbf{v}} = \begin{pmatrix} 2 \\ 2 \\ -3 \end{pmatrix}_{\mathcal{H}}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ -2 \\ 3 \end{pmatrix}_{\mathcal{H}}.$$

Die lineare Abbildung wird auf diese Weise ebenso **repräsentiert** durch die Vorschrift

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}} \mapsto \begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_{\mathcal{G}}.$$

Das allgemeine Schema und die Interpretation der obigen Vorschrift lautet (man mache im Beispiel die Probe mit $\underline{\mathbf{g}}^{(1)}$, $\underline{\mathbf{g}}^{(2)}$ und bedenke, dass die lineare Abbildung durch die Bilder dieser Basisvektoren festgelegt ist):

- i) Es sei \mathcal{G} eine Basis des \mathbb{R}^m , \mathcal{H} eine Basis \mathbb{R}^n und $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ sei eine lineare Abbildung.
- ii) Man berechne die Bilder (unter L) der Basisvektoren von \mathcal{G} .
- iii) Von diesen Bildern wird dann die Koordinatendarstellung bzgl. \mathcal{H} berechnet.
- iv) Die **Matrixdarstellung** $A_{\mathcal{H}}^{\mathcal{G}} \in M(n, m)$ der linearen Abbildung L bzgl. der Basen \mathcal{G} und \mathcal{H} erhält man, indem man aus diesen Spaltenvektoren eine Matrix bildet. Diese **darstellende Matrix** enthält mit anderen Worten die **Koordinaten der Bilder der Basisvektoren** als Spalten.
- v) Ein beliebiges $\underline{\mathbf{x}} \in \mathbb{R}^m$ wird **mit seiner Koordinatendarstellung**

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{G}} \quad \text{bzgl. der Basis } \mathcal{G} \text{ identifiziert.}$$

- vi) Dann liefert $A_{\mathcal{H}}^{\mathcal{G}} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{G}}$ die **Koordinatendarstellung von $L(\underline{\mathbf{x}})$** bzgl. der Basis \mathcal{H} .

Zu gegebenen Basen kann eine lineare Abbildung mit einer Matrix identifiziert werden.

Beim Übergang von Beispiel i) zu Beispiel ii) wurde ein **Basiswechsel** vollzogen, der zu einer **Koordinatentransformation** führte. Dieser Vorgang soll nun präzisiert werden.

Es sei dazu $\mathcal{A} = (\underline{\mathbf{a}}^{(1)}, \dots, \underline{\mathbf{a}}^{(m)})$ eine Basis des \mathbb{R}^m , $\mathcal{B} = (\underline{\mathbf{b}}^{(1)}, \dots, \underline{\mathbf{b}}^{(m)})$ sei eine weitere Basis des \mathbb{R}^m . Die “neuen Basisvektoren” aus \mathcal{B} seien in Abhängigkeit der “alten” gegeben, d.h. mit Koeffizienten $\gamma_{ij} \in \mathbb{R}$, $1 \leq i, j \leq m$, gelte

$$\begin{aligned} \underline{\mathbf{b}}^{(1)} &= \gamma_{11}\underline{\mathbf{a}}^{(1)} + \dots + \gamma_{1m}\underline{\mathbf{a}}^{(m)}, \\ &\vdots \\ \underline{\mathbf{b}}^{(m)} &= \gamma_{m1}\underline{\mathbf{a}}^{(1)} + \dots + \gamma_{mm}\underline{\mathbf{a}}^{(m)}. \end{aligned}$$

Aus diesen Koeffizienten wird die **transponierte Matrix** gebildet, d.h. die Matrix

$$S := (\gamma_{ij})^T = \begin{pmatrix} \gamma_{11} & \dots & \gamma_{m1} \\ \vdots & & \vdots \\ \gamma_{1m} & \dots & \gamma_{mm} \end{pmatrix} \in M(m, m).$$

Hat nun $\underline{\mathbf{x}} \in \mathbb{R}^m$ bzgl. \mathcal{A} bzw. \mathcal{B} die Koordinatendarstellungen

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}},$$

so gilt

$$\begin{aligned} \underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{a}}^{(1)} + \dots + \alpha_m \underline{\mathbf{a}}^{(m)} = \beta_1 \underline{\mathbf{b}}^{(1)} + \dots + \beta_m \underline{\mathbf{b}}^{(m)} \\ &= \sum_{j=1}^m \beta_j \left(\sum_{i=1}^m \gamma_{ji} \underline{\mathbf{a}}^{(i)} \right) \\ &= \sum_{i=1}^m \left(\sum_{j=1}^m \gamma_{ji} \beta_j \right) \underline{\mathbf{a}}^{(i)}, \end{aligned}$$

mit anderen Worten gilt für die neuen Koordinaten

$$\alpha_i = \sum_{j=1}^m \gamma_{ji} \beta_j .$$

Dies liefert die Transformation der Koordinaten bzgl. \mathcal{A} bzw. \mathcal{B}

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}} = S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}} = S^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}},$$

wobei die Invertierbarkeit von S unmittelbar daraus folgt, dass sowohl \mathcal{A} als auch \mathcal{B} Basen des \mathbb{R}^m sind.

Wichtige Beobachtung. Entsteht die neue **Basis** aus der alten beispielsweise durch eine Rotation, so gehen die neuen **Koordinaten** genau durch eine Rotation in die entgegengesetzte Richtung aus den alten hervor (sozusagen zur Kompensation, Übung!). Entsprechendes sieht man auch für die Multiplikation mit einem Skalar.

Basis- und Koordinatenänderungen verhalten sich konträr.

Als natürliche Frage schließt sich an:

Wie transformiert sich die darstellende Matrix einer linearen Abbildung beim Basiswechsel?

Es sei dazu $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung, $\mathcal{A} = (\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(m)})$ sei eine Basis des \mathbb{R}^m , $\mathcal{U} = (\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(n)})$ sei eine Basis des \mathbb{R}^n und $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$ sei die darstellende Matrix der linearen Abbildung L bzgl. dieser Basen.

Man betrachte weiterhin “neue Basen” $\mathcal{B} = (\mathbf{b}^{(1)}, \dots, \mathbf{b}^{(m)})$ des \mathbb{R}^m bzw. $\mathcal{V} = (\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(n)})$ des \mathbb{R}^n . Gesucht ist die Matrixdarstellung $A_{\mathcal{V}}^{\mathcal{B}}$ von L bzgl. dieser Basen.

Wie oben sei der Basiswechsel von \mathcal{A} nach \mathcal{B} durch die Matrix $S \in M(m, m)$ beschrieben, der Basiswechsel von \mathcal{U} nach \mathcal{V} werde analog von einer Matrix $T \in M(n, n)$ erzeugt.

Ist $L(\underline{\mathbf{x}}) = \underline{\mathbf{y}}$ und

$$\begin{aligned}\underline{\mathbf{x}} &= \alpha_1 \underline{\mathbf{a}}^{(1)} + \cdots + \alpha_m \underline{\mathbf{a}}^{(m)} \\ &= \beta_1 \underline{\mathbf{b}}^{(1)} + \cdots + \beta_m \underline{\mathbf{b}}^{(m)}, \\ \underline{\mathbf{y}} &= \mu_1 \underline{\mathbf{u}}^{(1)} + \cdots + \mu_n \underline{\mathbf{u}}^{(n)} \\ &= \nu_1 \underline{\mathbf{v}}^{(1)} + \cdots + \nu_n \underline{\mathbf{v}}^{(n)},\end{aligned}$$

so gilt mit der darstellenden Matrix $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$

$$\begin{aligned}\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{U}} &= A_{\mathcal{U}}^{\mathcal{A}} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}_{\mathcal{A}} \\ &= A_{\mathcal{U}}^{\mathcal{A}} S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}},\end{aligned}$$

wobei die zweite Gleichheit aus der bereits gezeigten Koordinatentransformation folgt. Eine Transformation der linken Seite ergibt nun

$$\begin{aligned}\begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix}_{\mathcal{V}} &= T^{-1} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{U}} \\ &= T^{-1} A_{\mathcal{U}}^{\mathcal{A}} S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}} = A_{\mathcal{V}}^{\mathcal{B}} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}_{\mathcal{B}}.\end{aligned}$$

Zusammenfassend ist folgender Satz gezeigt:

Satz 9.2.1

i) *Koordinatentransformation:* Es seien \mathcal{A}, \mathcal{B} Basen des \mathbb{R}^m , die Matrix $S \in M(m, m)$ sei wie oben gegeben. Ist

$$\underline{\mathbf{x}} = \alpha_1 \underline{\mathbf{a}}^{(1)} + \cdots + \alpha_m \underline{\mathbf{a}}^{(m)} = \beta_1 \underline{\mathbf{b}}^{(1)} + \cdots + \beta_m \underline{\mathbf{b}}^{(m)},$$

so gilt

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = S \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}, \quad \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = S^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

ii) Sind weiterhin \mathcal{U}, \mathcal{V} Basen des \mathbb{R}^n , ist die Transformation dieser Basen durch die Matrix $T \in M(n, n)$ gegeben und ist $L: \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung mit *Matrixdarstellung* $A_{\mathcal{U}}^{\mathcal{A}} \in M(n, m)$ bzgl. der *Basen* \mathcal{A}, \mathcal{U} , so lautet die *Matrixdarstellung von* L *bzgl. der Basen* \mathcal{B} *und* \mathcal{V}

$$A_{\mathcal{V}}^{\mathcal{B}} = T^{-1} A_{\mathcal{U}}^{\mathcal{A}} S.$$

Bemerkung. Da die Spaltenvektoren aus den Koordinaten selbst als Vektoren aus dem \mathbb{R}^m bzw. \mathbb{R}^n aufgefasst werden, werden im obigen Satz die Indizes zur Basisbenennung weggelassen.

Beispiele.

i) Die Situation in den einführenden Beispielen i), ii) zu Beginn des Kapitels sieht wie folgt aus: Der Basiswechsel von i) nach ii) wird

beschrieben durch

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix},$$

die Inverse von T berechnet sich zu

$$T^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

Es folgt wie behauptet

$$\begin{aligned} A_{\mathcal{H}}^{\mathcal{G}} &= T^{-1} \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} S \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 2 & -2 \\ -3 & 3 \end{pmatrix}. \end{aligned}$$

ii) Man betrachte eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}$, d.h. die darstellenden Matrizen bzgl. einer Basis \mathcal{A} bzw. einer Basis \mathcal{B} des \mathbb{R}^m sind Zeilenvektoren. Ist

$$\left(\xi_1 \quad \dots \quad \xi_m \right)$$

die darstellende Matrix bzgl. der Basis \mathcal{A} , so ist nach Satz 9.2.1

$$\begin{pmatrix} \psi_1 & \dots & \psi_m \end{pmatrix} = \begin{pmatrix} \xi_1 & \dots & \xi_m \end{pmatrix} \begin{pmatrix} \gamma_{11} & \dots & \gamma_{m1} \\ \vdots & & \vdots \\ \gamma_{1m} & \dots & \gamma_{mm} \end{pmatrix}$$

die darstellende Matrix bzgl. der Basis \mathcal{B} , wobei der Matrixwechsel wie oben durch die Koeffizienten γ_{ij} , $1 \leq i, j \leq m$, gegeben sei. Somit gilt für alle $i = 1, \dots, n$:

$$\psi_i = \sum_{j=1}^m \gamma_{ij} \xi_j,$$

die Zeilenvektoren (Koordinaten der linearen Abbildung) transformieren sich bei einem Basiswechsel genau wie die Basis, man spricht von einem **kovarianten Tensor² erster Stufe**.

Die eben angesprochenen linearen Abbildungen eines Vektorraums in den zugrunde liegenden Körper sind so genannte **lineare Funktionale**.

Definition 9.2.1

*Ist V ein \mathbb{K} -Vektorraum, so ist die Menge aller linearen Abbildungen $\varphi: V \rightarrow \mathbb{K}$ selbst wieder ein \mathbb{K} -Vektorraum und heißt der zu V **duale Vektorraum** oder **Dualraum** von V . Bezeichnung: V^* . Jedes $\varphi \in V^*$ heißt **Linearform** oder **lineares Funktional** auf V .*

Beispiel. Ist $V = \mathbb{R}^m$ und sind $a_1, \dots, a_m \in \mathbb{R}$, so entsteht eine Linearform $\varphi \in (\mathbb{R}^m)^*$ aus der Vorschrift

$$\varphi: \mathbb{R}^m \rightarrow \mathbb{R}, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto a_1 x_1 + \dots + a_m x_m,$$

²Der Tensorbegriff wird in Definition 9.2.2 präzisiert.

wobei auch der Zusammenhang mit linearen Gleichungssystemen beachtet werden sollte.

Ist eine Basis \mathcal{A} des \mathbb{R}^m fixiert, so können (wie bereits oben angesprochen) lineare Funktionale **mit Zeilenvektoren identifiziert** werden. Folglich hat der Dualraum $(\mathbb{R}^m)^*$ des \mathbb{R}^m die **gleiche Dimension** m wie der \mathbb{R}^m .

Mit der kanonischen Basis des \mathbb{R}^m existiert auch eine kanonische Basis des $(\mathbb{R}^m)^*$ (die **duale Basis**). Sie ist gegeben durch

$$\begin{aligned}(\underline{\mathbf{e}}^{(1)})^T &= \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \\(\underline{\mathbf{e}}^{(2)})^T &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \end{pmatrix}, \\&\vdots \\(\underline{\mathbf{e}}^{(m)})^T &= \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix},\end{aligned}$$

und jede Linearform auf \mathbb{R}^m (d.h. jeder Zeilenvektor – bei fixierter Basis) kann klarerweise als Linearkombination der $(\underline{\mathbf{e}}^{(i)})^T$, $1 \leq i \leq m$, geschrieben werden.

Man beachte, dass für alle $1 \leq i, j \leq m$ gilt

$$(\underline{\mathbf{e}}^{(j)})^T(\underline{\mathbf{e}}^{(i)}) = \delta_{ij}.$$

Mit der Schreibweise einer Linearform als Zeilenvektor ist die **Darstellung der linearen Abbildung bzgl. einer fixierten Basis** gemeint, legt man eine andere Basis zugrunde, so erhält man einen anderen Zeilenvektor als Repräsentant der Linearform. Der Transformationsmechanismus wurde bereits diskutiert. Mit anderen Worten handelt es sich hier um die **Koordinaten einer linearen Abbildung bzgl. einer fixierten Basis**.

Bemerkung. Von besonderem Interesse sind auch lineare Funktionale auf Funktionenräumen, was aber an dieser Stelle nicht vertieft werden

soll.

Der Dualraum eines Vektorraums ist wie in Definition 9.2.1 angedeutet selbst wieder ein Vektorraum. Folglich können analog zu obigen Bemerkungen auch lineare Funktionale $\zeta: V^* \rightarrow \mathbb{K}$ betrachtet werden.

Wie hat man sich das vorzustellen?

Es sei etwa die kanonische Basis des \mathbb{R}^m (und damit die duale Basis) fixiert. Lineare Abbildungen vom $(\mathbb{R}^m)^*$ in die reellen Zahlen (lineare Funktionale auf $(\mathbb{R}^m)^*$) ordnen jedem Zeilenvektor (einer linearen Abbildung vom \mathbb{R}^m nach \mathbb{R}) eine reelle Zahl zu. Diese Vorschrift kann mit einem **fixierten Spaltenvektor** identifiziert werden.

Ist nämlich

$$\underline{\mathbf{v}} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in \mathbb{R}^m \quad (\text{fixiert}),$$

so ist $\zeta: (\mathbb{R}^m)^* \rightarrow \mathbb{R}$,

$$\zeta\left(\begin{pmatrix} w_1 & \dots & w_m \end{pmatrix}\right) := \begin{pmatrix} w_1 & \dots & w_m \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \in \mathbb{R}$$

eine lineare Abbildung, ζ wird mit einem Element des \mathbb{R}^m identifiziert (mit $\underline{\mathbf{v}}$).

Ist V ein endlichdimensionaler Vektorraum, so kann der **Bidualraum** $(V^*)^*$ mit V selbst **identifiziert werden**.

Lineare Funktionale auf dem $(\mathbb{R}^m)^*$ verhalten sich also wie der \mathbb{R}^m selbst. Es sei an dieser Stelle an die Koordinatentransformation aus Satz 9.2.1,

$$\beta_i = \sum_{j=1}^m \rho_{ij} \alpha_j, \quad (\rho_{ij}) = ((\gamma_{ij})^T)^{-1}, \quad i = 1, \dots, m,$$

erinnert, man spricht von einem **kontravarianten Tensor erster Stufe**.

Warum wird ein Tensor als (multi-) lineares Funktional definiert?

Wird beispielsweise ein Spannungstensor einfach als 3×3 -Matrix eingeführt, so ist er bzgl. einer gegebenen Basis des \mathbb{R}^3 definiert. Bezieht man sich auf eine andere Basis, so ändert sich die Matrixdarstellung.

Eine physikalische Größe sollte aber **unabhängig von der zugrunde liegenden Basis sein**, und deshalb wird von einer Matrixdarstellung zu Multilinearformen übergegangen.

Erinnerung. Der Begriff **Bilinearform** ($n = 2$) bzw. allgemein **n -Linearform** (**Multilinearform**) wurde im Paragraphen 8.6 erklärt. Es handelt sich um eine Abbildung M von n Vektorräumen in den zugrunde liegenden Körper, $M: V_1 \times \cdots \times V_n \rightarrow \mathbb{K}$, die **linear in jedem Argument ist**.

Mit diesen Vorbereitungen kann nun definiert werden, was unter einem **Tensor** zu verstehen ist.

Definition 9.2.2

Gegeben sei ein endlichdimensionaler \mathbb{K} -Vektorraum V mit Dualraum V^* . Die Elemente der Menge $\mathcal{T}_q^p(V)$ der **Multilinearformen** M ,

$$M : \underbrace{V \times V \times \cdots \times V}_{q \text{ mal}} \times \underbrace{V^* \times V^* \times \cdots \times V^*}_{p \text{ mal}} \rightarrow \mathbb{K},$$

heißen **q -fach kovariante** und **p -fach kontravariante** Tensoren über V . Dabei setzt man $\mathcal{T}_0^0 = \mathbb{K}$.

Bemerkung. Die Begriffe kovariant und kontravariant wurden bereits ausführlich diskutiert.

Tensoren erster Stufe sind oben angesprochen, es soll nun gezeigt werden, wie man mit Hilfe des **Tensorproduktes** Tensoren höherer Stufe produzieren kann.

Dazu bezeichnen U und V zwei Vektorräume über einem Körper \mathbb{K} . Jedes feste $u \in U$ kann als Linearform auf U^* interpretiert werden (wie das bereits bei Spaltenvektoren geschehen ist). Die Vorschrift lautet:

$$u(u^*) := u^*(u) \quad \text{für alle } u^* \in U^* .$$

Man beachte, dass die rechte Seite wohl definiert ist. Analoges gilt natürlich auch für jedes feste $v \in V$.

Schließlich wird $u \otimes v: U^* \times V^* \rightarrow \mathbb{K}$ definiert als

$$(u \otimes v)(u^*, v^*) = u(u^*)v(v^*) \quad \text{für alle } u^* \in U^* , v^* \in V^* .$$

Die Gesamtheit aller endlichen Summen von Termen dieser Form heißt das **Tensorprodukt** $U \otimes V$, $U \otimes V$ besteht aus **Bilinearformen auf** $U^* \times V^*$.

Bemerkung. Analog wird $U_1 \otimes U_2 \otimes \cdots \otimes U_r$ definiert.

Beispiel. Es sei $U = \mathbb{R}^m$, $V = (\mathbb{R}^m)^*$ und die kanonische Basis des \mathbb{R}^m sei als Grundlage fixiert. Dann kann $\underline{\mathbf{u}} \in \mathbb{R}^m$ als Spaltenvektor geschrieben werden, $\underline{\mathbf{v}}^T \in (\mathbb{R}^m)^*$ wird als Zeilenvektor interpretiert. Es ist

$$\underline{\mathbf{u}}(\underline{\mathbf{x}}^T) := \begin{pmatrix} x_1 & \dots & x_m \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \quad \text{für alle } \underline{\mathbf{x}}^T \in (\mathbb{R}^m)^* .$$

Analog ist

$$\underline{\mathbf{v}}^T(\underline{\mathbf{y}}) := \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad \text{für alle } \underline{\mathbf{y}} \in ((\mathbb{R}^m)^*)^* = \mathbb{R}^m .$$

Es ist also

$$\begin{aligned}
 \underline{\mathbf{u}} \otimes \underline{\mathbf{v}}^T(\underline{\mathbf{x}}^T, \underline{\mathbf{y}}) &= \begin{pmatrix} x_1 & \dots & x_m \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \\
 &= \begin{pmatrix} x_1 & \dots & x_m \end{pmatrix} \left[\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} \right] \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \\
 &= \begin{pmatrix} x_1 & \dots & x_m \end{pmatrix} \begin{pmatrix} u_1 v_1 & \dots & u_1 v_m \\ \vdots & & \vdots \\ u_m v_1 & \dots & u_m v_m \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.
 \end{aligned}$$

Auf diese Art wird der Tensor $\underline{\mathbf{u}} \otimes \underline{\mathbf{v}}^T$ zweiter Stufe für eine fixierte Basis als Matrix dargestellt. (Übung: Wie wird der Tensor im Sinne von Definition 9.2.2 bezeichnet?)

An dieser Stelle soll die (teils heuristische) Diskussion von Tensoren abgeschlossen werden.

Teil III

Analysis von Funktionen einer Veränderlichen

Kapitel 10

Stetige Funktionen

Im dritten Teil der Vorlesung geht es (wie der Titel schon sagt) um die Analysis (die [Infinitesimalrechnung](#), die [Differential- und Integralrechnung](#)) von Funktionen einer Veränderlichen. Da jedoch im Paragraphen 10.1 die Diskussion von Funktionen mehrerer Veränderlichen ohne großen Mehraufwand in die Betrachtungen eingeschlossen werden kann, wird dieser Fall mit abgedeckt.

Teil III schließt sich in zweierlei Hinsicht an das Kapitel 9 über lineare Abbildungen an. Zu Einem sind lineare Abbildungen besonders einfache Beispiele stetiger oder differenzierbarer Funktionen. Zum Anderen ist die Ableitung oder das Differential einer Funktion als lineare Abbildung zu interpretieren, die die Funktion in geeigneter Weise approximiert.

10.1 Stetige Funktionen in mehreren Veränderlichen (Lipschitz-stetige Funktionen; Grenzwert von Funktionen; Definition der Stetigkeit; Rechenregeln für stetige Funktionen; Stetigkeit und Kompaktheit)

Unter einer reellwertigen stetigen Funktion abhängig von einer Variablen stellt man sich üblicherweise eine Funktion vor, deren Graph “ohne Absetzen des Stiftes” gezeichnet werden kann, so wie z.B. in Abbildung 10.1 angedeutet.

Auch die Betragsfunktion erfüllt dieses heuristische Kriterium und ist eine stetige Funktion (vgl. Abbildung 10.2).

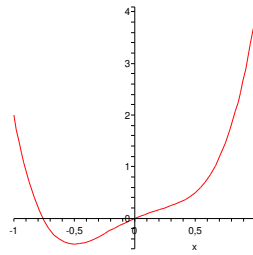


Abbildung 10.1: Der Graph einer stetigen Funktion.

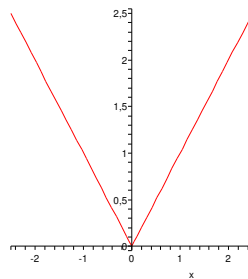


Abbildung 10.2: Der Graph der Betragsfunktion.

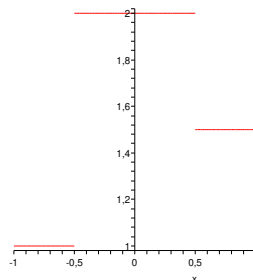
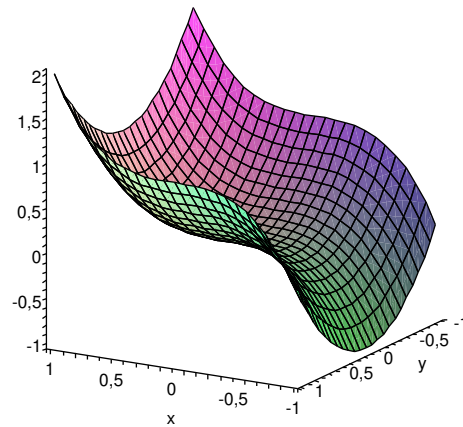
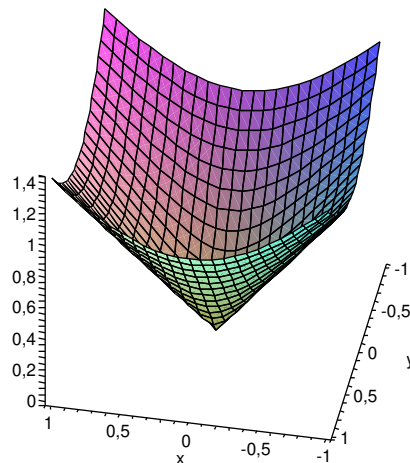


Abbildung 10.3: Der Graph einer unstetigen Funktion mit Sprungstellen.

Eine typische unstetige Funktion hat [Sprungstellen](#), so wie es in [Abbildung 10.3](#) angedeutet ist.

Analog sieht die Situation bei Funktionen $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ aus. Hier ist jedoch zu beachten, dass sich eine Funktion in einer Richtung stetig, in einer anderen Richtung aber unstetig verhalten kann. In den [Abbildungen 10.4](#), [10.5](#) und [10.6](#) sind Graphen von reellwertigen Funktionen zweier Veränderlicher dargestellt. Im Falle $n \geq 3$ versagt natürlich die zeichnerische Interpretation.


 Abbildung 10.4: Der Graph einer stetigen Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$.

 Abbildung 10.5: Der Graph der Betragsfunktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$.

Die begrenzte Tragweite des heuristischen Stetigkeitsbegriffes erkennt man beispielsweise anhand der Funktion $f(x) = x \cos(1/x)$. Wie soll hier die Stetigkeit verifiziert bzw. falsifiziert (vgl. die Abbildungen 10.7 und 10.8) werden?

Bemerkungen.

- i) **Vorurteil:** Alle in der Natur bzw. in den Anwendungen vorkommenden Funktionen sind sowieso “gutartig”. Funktionen wie in den

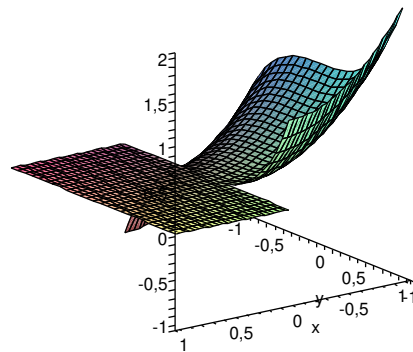


Abbildung 10.6: Der Graph einer unstetigen Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ (stetig in einer Richtung).

Abbildungen 10.7 und 10.8 angedeutet sind reine mathematische Konstrukte und physikalisch irrelevant.

- ii)* Die in *i)* formulierte Aussage ist **falsch!** In der Natur treten durchaus “singuläre Phänomene” auf. Beispielsweise Riss- oder Lochbildungen in Werkstoffen unter Belastung oder bei Erwärmung
- iii)* Gerade die Untersuchung solcher **Singularitäten** ist auch eine wesentliche Aufgabe in den Materialwissenschaften.
- iv)* Wie Abbildung 10.9 zeigt, sind z.B. bei **Phasenübergängen in elastischen Festkörpern** sogenannte **Mikrostrukturen** zu beobachten. Es handelt sich hier ganz offensichtlich nicht um “glatte” Strukturen. Ähnlich zu Abbildung 10.6 erkennt man in gewisse Richtungen reguläre Strukturen, in andere Richtungen beobachtet man Sprünge.

Die Betragsfunktion hat zwar eine “Knickstelle”, sie hat aber noch eine sehr gute Eigenschaft, die durch den Begriff der **Lipschitz-Stetigkeit** (vgl. Übungen) charakterisiert ist: Die Differenz des Abstandes der Bilder zweier Punkte ist durch den Abstand der Urbilder abgeschätzt.

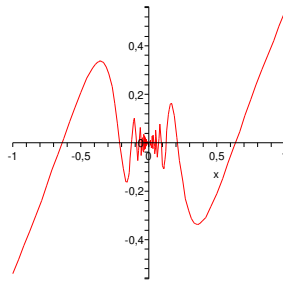
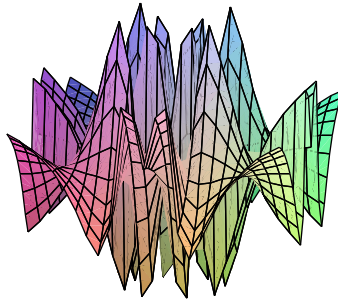

 Abbildung 10.7: Ist die Funktion $x \cos(1/x)$ stetig?


Abbildung 10.8: Ist diese Funktion stetig?

Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = |x|$, so folgt für alle $x, y \in \mathbb{R}$ nach der Dreiecksungleichung

$$|f(x) - f(y)| = ||x| - |y|| \leq |x - y| .$$

Bei einer Sprungfunktion gibt es keine Chance, eine analoge Abschätzung für alle $x, y \in \mathbb{R}$ zu erhalten (warum?).

Bemerkung. Eine Lipschitz-Bedingung wird eine wesentliche Rolle bei der Diskussion gewöhnlicher Differentialgleichungen spielen.

Genauer als bisher erläutert ist die Idee des Stetigkeitsbegriffes: “Kleine” Änderungen im Argument sollten bei einer stetigen (einer “kontinuierlichen”) Funktion nur “kleine” Änderungen im Funktionswert bewirken, so wie es bei Lipschitz-stetigen Funktionen in besonders



Abbildung 10.9: Mikrostrukturen (aus [Mü]).

kontrollierter Weise der Fall ist.

Zur Präzisierung dieser Idee muss zunächst der **Grenzwert einer Funktion** eingeführt werden (bisher wurden nur **Grenzwerte von Folgen** diskutiert). Dazu sei an den Begriff **Häufungspunkt** aus Definition 6.4.3 erinnert:

Per definitionem liegt in jeder Umgebung eines Häufungspunktes $\underline{\mathbf{x}}^{(0)}$ einer Menge $U \subset \mathbb{R}^m$ ein Punkt von U , der von $\underline{\mathbf{x}}^{(0)}$ verschieden ist. Es existiert also eine Folge $\{\underline{\mathbf{x}}^{(k)}\}$, $\underline{\mathbf{x}}^{(k)} \neq \underline{\mathbf{x}}^{(0)}$, $\underline{\mathbf{x}}^{(k)} \in U$ für alle $k \in \mathbb{N}$, mit $\lim_{k \rightarrow \infty} \underline{\mathbf{x}}^{(k)} = \underline{\mathbf{x}}^{(0)}$ (Übung).

Damit wird definiert (vgl. auch Satz 10.1.1):

Definition 10.1.1

Es sei $\underline{\mathbf{x}}^{(0)}$ ein Häufungspunkt der Definitionsmenge $U \subset \mathbb{R}^m$ einer Abbildung $f: U \rightarrow \mathbb{R}^n$. Man schreibt

$$\lim_{\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{(0)}} f(\underline{\mathbf{x}}) = \underline{\mathbf{a}},$$

wenn es zu jedem $\varepsilon > 0$ ein $\delta(\varepsilon) > 0$ gibt, sodass gilt

$$\underline{\mathbf{x}} \in U, \quad \underline{\mathbf{x}} \neq \underline{\mathbf{x}}^{(0)}, \quad \|\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}\| < \delta \quad \Rightarrow \quad \|f(\underline{\mathbf{x}}) - \underline{\mathbf{a}}\| < \varepsilon.$$

*Bezeichnung: f hat an der Stelle $\underline{\mathbf{x}}^{(0)}$ den **Grenzwert** $\underline{\mathbf{a}}$.*

Bemerkungen und Beispiele.

- i)* Hier bezeichnet $\|\cdot\|$ die Euklidische Norm im \mathbb{R}^m bzw. \mathbb{R}^n (den Betrag im Fall $m = 1$ bzw. $n = 1$).
- ii)* Ebenso wie in Kapitel 4 folgt: **Falls ein Grenzwert existiert, so ist er eindeutig.**
- iii)* Für Funktionen $f: \mathbb{C} \rightarrow \mathbb{R}$ oder $f: \mathbb{C} \rightarrow \mathbb{C}$ ist die Grenzwertdefinition identisch, ebenso wie die nachfolgenden Betrachtungen zur Stetigkeit. Deshalb wird dieser Fall nicht gesondert diskutiert.
- iv)* Betrachtet sei wieder die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$ für alle $x \in \mathbb{R}$. Um die Existenz eines Grenzwertes im Punkt $x^{(0)} = 0$ zu untersuchen, wird zu gegebenem $\varepsilon > 0$ beispielsweise $\delta(\varepsilon) = \varepsilon$ gewählt. Aus der Lipschitz-Stetigkeit folgt unmittelbar, dass f an der Stelle $x^{(0)} = 0$ den Grenzwert 0 hat (vgl. Abbildung 10.10).

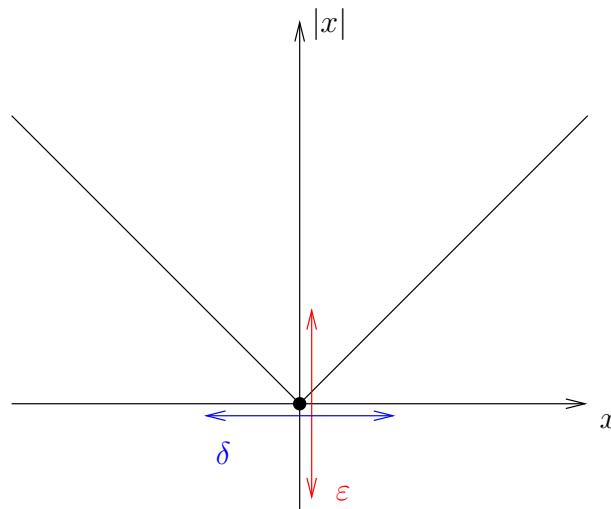


Abbildung 10.10: Der Grenzwertbegriff für die Betragsfunktion im Nullpunkt.

- v)* Natürlich kann δ auch kleiner gewählt werden.
- vi)* Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$ für alle $x \in \mathbb{R}$ und es sei $x^{(0)} = 0$. Für beliebiges, fixiertes $\varepsilon > 0$ wird $\delta = \sqrt{\varepsilon}$ gewählt und es folgt für alle $x \in \mathbb{R}$, $x \neq 0$, mit $|x - x^{(0)}| < \delta$:

$$|f(x) - 0| = |f(x)| = |x^2| < \delta^2 = \varepsilon ,$$

f hat an der Stelle $x^{(0)} = 0$ den Grenzwert 0.

vii) f muss an der Stelle $\underline{\mathbf{x}}^{(0)}$ nicht definiert zu sein.

viii) Es sei $U = \mathbb{R}^2 - \{\underline{\mathbf{0}}\}$ und $f: U \rightarrow \mathbb{R}$ sei gegeben durch

$$f(\underline{\mathbf{x}}) = \frac{x_1^2 x_2^2}{x_1^2 + x_2^2}.$$

In diesem Beispiel ist f im Punkt $\{\underline{\mathbf{0}}\}$ nicht definiert.

Der Punkt $\underline{\mathbf{0}}$ ist aber Häufungspunkt von U (warum?) und zu $\varepsilon > 0$ wähle man z.B. $\delta = \sqrt{\varepsilon}$. Es folgt für $\underline{\mathbf{x}} \in \mathbb{R}^2$, $\underline{\mathbf{x}} \neq \underline{\mathbf{0}}$, $\|\underline{\mathbf{x}}\| < \delta$:

$$|f(\underline{\mathbf{x}}) - \underline{\mathbf{0}}| = |f(\underline{\mathbf{x}})| = \left| \frac{x_1^2 x_2^2}{x_1^2 + x_2^2} \right| \leq \frac{\|x\|^4}{\|x\|^2} = \|x\|^2 < \delta^2 = \varepsilon,$$

f hat an der Stelle $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{0}}$ den Grenzwert 0 (der Graph der Funktion ist in Abbildung 10.11 dargestellt).

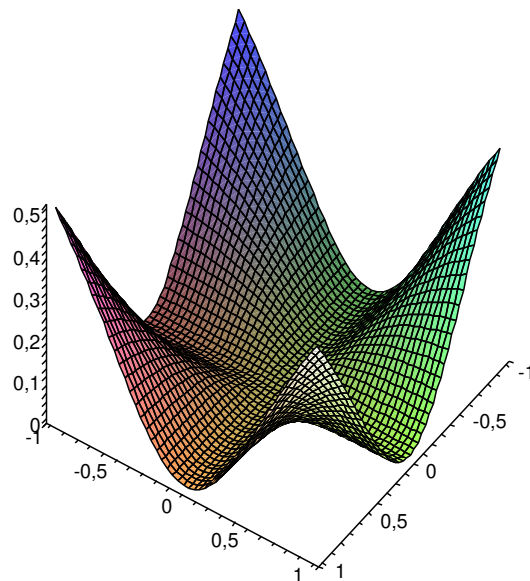


Abbildung 10.11: Im Nullpunkt existiert der Grenzwert in diesem Beispiel.

Der Grenzwert von Funktionen kann äquivalent über Folgen eingeführt werden:

Satz 10.1.1

Mit den Bezeichnungen aus Definition 10.1.1 gilt:

$$\lim_{\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{(0)}} f(\underline{\mathbf{x}}) = \underline{\mathbf{a}}$$

ist **genau dann** richtig, wenn **für jede** Folge $\{\underline{\mathbf{x}}^{(k)}\}$ von Punkten $\underline{\mathbf{x}}^{(k)} \in U - \{\underline{\mathbf{x}}^{(0)}\}$ mit $\underline{\mathbf{x}}^{(k)} \rightarrow \underline{\mathbf{x}}^{(0)}$ für $k \rightarrow \infty$ gilt

$$\lim_{k \rightarrow \infty} f(\underline{\mathbf{x}}^{(k)}) = \underline{\mathbf{a}} .$$

Bemerkungen und Beispiele.

- i) Mit Satz 10.1.1 können die bekannten Rechenregeln für Folgen ausgenutzt werden (vgl. Satz 4.1.4).
- ii) **Es reicht nicht aus, nur spezielle Folgen zu betrachten.**
- iii) Man betrachte die Sprungfunktion (vgl. Abbildung 10.12) $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 0 & \text{für } x < 0, \\ 1 & \text{für } x \geq 0. \end{cases}$$

Dann haben gegen 0 konvergente Folgen mit negativen Gliedern konstant den Funktionswert 0, nicht-negative hingegen konstant den Funktionswert 1.

Der Grenzwert in 0 existiert nicht. Es ist aber möglich, einen **einseitigen Grenzwert** zu definieren (siehe Übungen).

- iv) Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} \sin(1/x) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

Dann betrachte man einerseits die Folge $\{x_k\}$ mit $x_k = 1/k\pi$ für alle $k \in \mathbb{N}$, andererseits betrachte man die Folge $\{\tilde{x}_k\}$ mit $\tilde{x}_k = 2/((4k+1)\pi)$ für alle $k \in \mathbb{N}$. Wie sehen die Grenzwerte dieser Folgen aus?

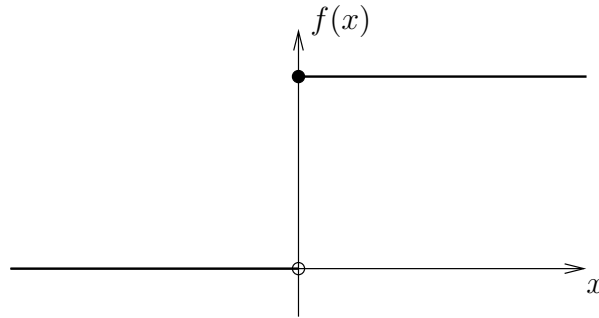


Abbildung 10.12: Eine Sprungfunktion.

v) Im Fall $m > 1$ kann sich f z.B. auf den Koordinatenachsen (aus allen möglichen Richtungen) unterschiedlich verhalten.

Es sei etwa $U = \mathbb{R}^2 - \{\mathbf{0}\}$ und $f: U \rightarrow \mathbb{R}$ sei gegeben durch

$$f(\underline{\mathbf{x}}) = \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2}.$$

Es ist $\underline{\mathbf{x}}^{(0)} = \mathbf{0}$ Häufungspunkt von U . Für eine Folge $\{\underline{\mathbf{x}}^{(k)}\}$ der Form $\underline{\mathbf{x}}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ 0 \end{pmatrix}$, $x_1^{(k)} \rightarrow 0$, $x_1^{(k)} \neq 0$, gilt

$$\lim_{k \rightarrow \infty} f(\underline{\mathbf{x}}^{(k)}) = 1.$$

Dahingegen gilt für eine Folge $\tilde{\underline{\mathbf{x}}}^{(k)}$ der Form $\tilde{\underline{\mathbf{x}}}^{(k)} = \begin{pmatrix} 0 \\ \tilde{x}_2^{(k)} \end{pmatrix}$,

$$\tilde{x}_2^{(k)} \rightarrow 0, \tilde{x}_2^{(k)} \neq 0$$

$$\lim_{k \rightarrow \infty} f(\tilde{\underline{\mathbf{x}}}^{(k)}) = -1.$$

f an der Stelle $\underline{\mathbf{x}}^{(0)} = \mathbf{0}$ **keinen Grenzwert**. Das Verhalten der Funktion ist in Abbildung 10.13 angedeutet.

Der Grenzwertbegriff führt zur folgenden Definition der Stetigkeit.

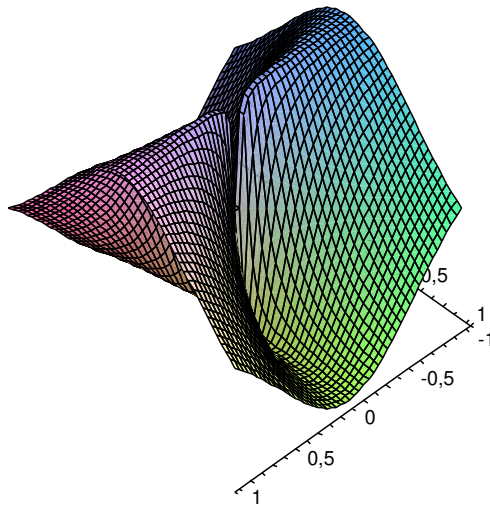


Abbildung 10.13: Im Nullpunkt existiert der Grenzwert in diesem Beispiel nicht.

Definition 10.1.2

Es sei $\underline{\mathbf{x}}^{(0)} \in U$ ein Häufungspunkt der Definitionsmenge $U \subset \mathbb{R}^m$ einer Abbildung $f: U \rightarrow \mathbb{R}^n$. Dann heißt f **stetig im Punkt $\underline{\mathbf{x}}^{(0)}$** , wenn gilt

$$\lim_{\underline{\mathbf{x}} \rightarrow \underline{\mathbf{x}}^{(0)}} f(\underline{\mathbf{x}}) = f(\underline{\mathbf{x}}^{(0)}) .$$

Die **Abbildung f** heißt **stetig**, wenn sie für alle $\underline{\mathbf{x}} \in U$ stetig ist.

Bemerkungen und Beispiele.

- i) Stetigkeit ist eine **lokale Eigenschaft**, die in jedem Punkt $\underline{\mathbf{x}}^{(0)} \in U$ zu untersuchen ist.
- ii) Konstante Abbildungen und die Identität sind stetig.
- iii) Man betrachte die **Koordinatenprojektion** $f: \mathbb{R}^m \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = x_j$ für ein festes $1 \leq j \leq m$. Ist $\varepsilon > 0$ gegeben, so wähle man $\delta = \varepsilon$

und erhält für alle $\underline{\mathbf{x}} \in \mathbb{R}^m$, $\underline{\mathbf{x}} \neq \underline{\mathbf{x}}^{(0)}$, mit $\|\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}\| < \delta = \varepsilon$:

$$|f(\underline{\mathbf{x}}) - f(\underline{\mathbf{x}}^{(0)})| = |x_j - x_j^{(0)}| \leq \|\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}\| < \varepsilon,$$

die Funktion ist somit stetig.

iv) Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 0 & \text{für } x \neq 1, \\ 1 & \text{für } x = 1, \end{cases}$$

ist im Punkt $x^{(0)} = 1$ nicht stetig (vgl. Abbildung 10.14). Es gilt nämlich

$$\lim_{x \rightarrow 1} f(x) = 0 \neq 1 = f(1).$$

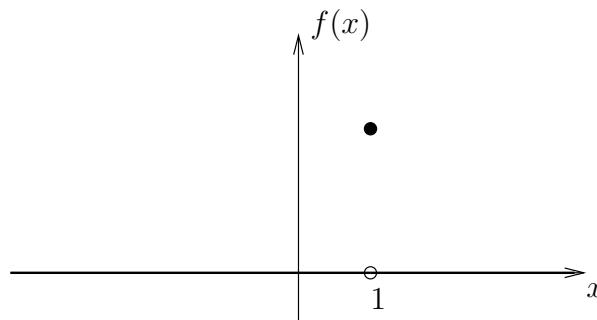


Abbildung 10.14: Diese Funktion ist im Punkt $x = 1$ nicht stetig.

v) Es sei $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$g(\underline{\mathbf{x}}) = \begin{cases} \frac{x_1^2 x_2^2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Wie bereits oben gezeigt, ist g stetig in $\underline{\mathbf{0}}$.

vi) Es sei $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$h(\underline{\mathbf{x}}) = \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Betrachtet man gegen $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{0}}$ konvergente Folgen entlang der x_1 -Achse und längs der Winkelhalbierenden $x_2 = x_1$, so erkennt man, dass diese Funktion **nicht stetig** in $\underline{\mathbf{0}}$ ist (vgl. die Abbildungen 10.15 und 10.16).

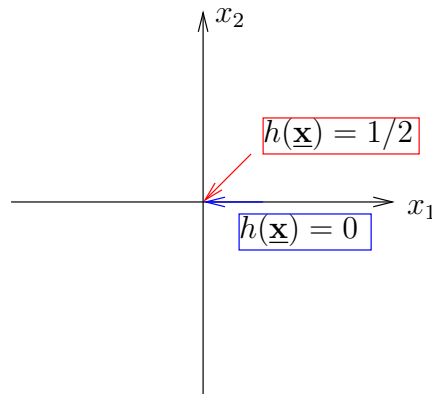


Abbildung 10.15: Unterschiedliches Verhalten der Funktion aus verschiedenen Richtungen.

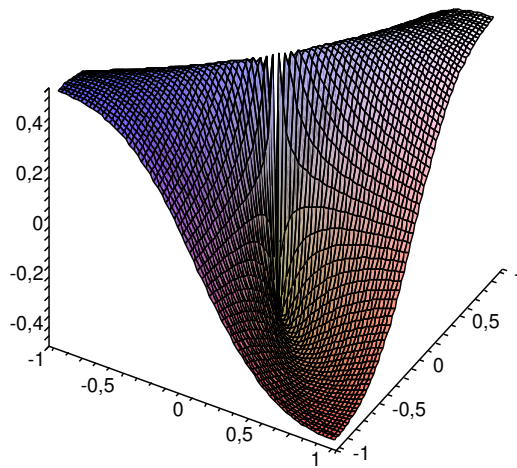


Abbildung 10.16: Im Nullpunkt ist diese Funktion nicht stetig.

vii) Die Abbildung $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ ist **genau dann** in $\underline{\mathbf{x}}^{(0)}$ stetig, wenn die

reellwertigen Funktionen f_1, f_2, \dots, f_m stetig in $\underline{\mathbf{x}}^{(0)}$ sind (vgl. Kapitel 6.3). Zu untersuchen ist in diesem Fall, ob **alle** Komponentenfunktionen stetig sind.

viii) Demnach ist die Abbildung $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = \begin{pmatrix} g(\underline{\mathbf{x}}) \\ h(\underline{\mathbf{x}}) \end{pmatrix}$, g, h wie in den beiden letzten Beispielen, nicht stetig in $\underline{\mathbf{0}}$.

Rechenregeln für stetige Funktionen?

Die folgenden Rechenregeln ergeben sich aus den bekannten Regeln für konvergente Folgen und sind wichtig, um die Stetigkeit einer Funktion evtl. auf die Stetigkeit bekannter Funktionen zurückführen zu können.

Beispielsweise ist die Stetigkeit von Polynomen eine unmittelbare Konsequenz (siehe nachfolgende Beispiele).

Satz 10.1.2

Mit stetigen $f: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ und $g: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ sind auch die folgenden Abbildungen stetig:

$$f + g \quad \text{und} \quad cf \quad (c \in \mathbb{R}).$$

Satz 10.1.3

Sind $f, g: U \subset \mathbb{R}^m \rightarrow \mathbb{R}$ reellwertige, stetige Funktionen, so ist die Funktion $f \cdot g$ überall und die Funktion $\frac{f}{g}$ an allen Stellen, in denen g nicht den Wert 0 annimmt, stetig.

Beispiele.

i) Da die konstante Funktion und die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x$, stetig sind, sind nach Satz 10.1.2 und nach Satz 10.1.3 alle Polynome einer Veränderlichen stetig.

ii) Polynome $p: \mathbb{R}^m \rightarrow \mathbb{R}$ vom Grad $\leq n$ in mehreren Veränderlichen sind von der Form

$$p(\underline{x}) = \sum_{k_1, \dots, k_m} a_{k_1, \dots, k_m} x_1^{k_1} \dots x_m^{k_m} \quad \text{mit} \quad \sum_{i=1}^m k_i \leq n .$$

Da die Koordinatenprojektionen stetig sind, sind auch Polynome in mehreren Veränderlichen stetig.

Auch die Verkettung stetiger Funktionen bereitet keine Probleme:

Satz 10.1.4

Es sei f eine stetige Abbildung aus dem \mathbb{R}^m in den \mathbb{R}^p und g eine stetige Abbildung aus dem \mathbb{R}^p in den \mathbb{R}^n , die verkettet werden können. Dann ist auch die Abbildung $h = g \circ f$ aus dem \mathbb{R}^m in den \mathbb{R}^n stetig.

Zu den wichtigsten Aufgaben bei der Analyse reellwertiger Funktionen gehört die Suche nach Maxima bzw. Minima (bei vektorwertigen Funktionen kann der Begriff Maximum bzw. Minimum nicht definiert werden).

Hier hilft die Stetigkeit einer Funktion allein nicht weiter. Man betrachte etwa die Funktion $f: (0, \infty) \rightarrow \mathbb{R}$, $f(x) = 1/x$. Diese Funktion ist auf $(0, \infty)$ stetig, nimmt jedoch weder ein Maximum noch ein Minimum an. Auf diese Problematik wurde bereits mehrfach hingewiesen (siehe etwa Abbildung 3.1).

Die Existenz von Maxima bzw. Minima ergibt sich erst aus dem Zusammenspiel von Stetigkeit und Kompaktheit.

Satz 10.1.5

i) Ist die Definitionsmenge einer stetigen Abbildung $f: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ kompakt, so ist auch die Bildmenge kompakt.

ii) Ist insbesondere $f: U \subset \mathbb{R}^m \rightarrow \mathbb{R}$ (*skalarer Fall*) stetig und U kompakt, so nimmt die Funktion f in U ihr *absolutes Maximum* und ihr *absolutes Minimum* an, d.h. es gibt Punkte $\underline{\mathbf{x}}^{(0)}, \underline{\mathbf{x}}^{(1)} \in U$, sodass für alle $\underline{\mathbf{x}} \in U$ gilt

$$f(\underline{\mathbf{x}}^{(0)}) \leq f(\underline{\mathbf{x}}) \leq f(\underline{\mathbf{x}}^{(1)}) .$$

Beweis von ii). O.E. sei hier der Fall $U = [a, b] \subset \mathbb{R}$ betrachtet, den allgemeinen Fall zeigt man mit identischen Argumenten. Es sei

$$s := \sup\{f(x) : a \leq x \leq b\} .$$

Dieses s existiert, wenn der “Wert $+\infty$ ” zugelassen wird.

Nach der Definition von s existiert eine Folge $\{x^{(k)}\}$, $a \leq x^{(k)} \leq b$ für alle $k \in \mathbb{N}$, mit

$$f(x^{(k)}) \xrightarrow{k \rightarrow \infty} s . \quad (*)$$

Der Satz von Bolzano-Weierstraß (Satz 4.1.8) liefert weiter eine konvergente Teilfolge $\{x^{(k_j)}\}$ und ein $x^{(0)} \in [a, b]$ mit

$$x^{(k_j)} \xrightarrow{j \rightarrow \infty} x^{(0)} . \quad (**)$$

Natürlich bleibt die Konvergenz $(*)$ beim Übergang zu der Teilfolge erhalten.

Andererseits ist f nach Voraussetzung stetig und aus $(**)$ ergibt sich

$$f(x^{(k_j)}) \xrightarrow{j \rightarrow \infty} f(x^{(0)}) .$$

Es ist gezeigt

$$s = f(x^{(0)}) = \max_{a \leq x \leq b} f(x) ,$$

in der Tat ist s also endlich und f nimmt sein Maximum in $x^{(0)}$ an. Die Annahme des Minimums folgt ebenso. \square

Bemerkung. Die erste Aussage des Satzes 10.1.5 impliziert insbesondere die *Beschränktheit stetiger Funktionen auf einem Kompaktum*.

Anhand der Funktion $f(x) = 1/x$ lässt sich auch der Begriff der **gleichmäßigen Stetigkeit** erklären.

Man betrachte die Funktion auf dem Intervall $[a, b]$ mit $a > 0$. Desweiteren betrachte man einen Punkt $x^{(0)} \in [a, b]$, in dem die Stetigkeit von f untersucht werden soll. Es ist für alle $x \in [a, b]$

$$|f(x) - f(x^{(0)})| = \left| \frac{1}{x} - \frac{1}{x^{(0)}} \right| = \frac{|x^{(0)} - x|}{xx^{(0)}} \leq \frac{|x^{(0)} - x|}{a^2}.$$

Zu gegebenem $\varepsilon > 0$ ist δ also kleiner oder gleich εa^2 zu wählen, um die Stetigkeit nachzuweisen.

Auf dem Intervall $(0, b)$ hingegen muss δ zusätzlich in Abhängigkeit von $x^{(0)}$ gewählt werden, $\delta = \delta(\varepsilon, x^{(0)})$, wobei bei festem $\varepsilon > 0$ gilt $\delta(\varepsilon, x^{(0)}) \rightarrow 0$, falls $x^{(0)} \rightarrow 0$.

Zu gegebenem $\varepsilon > 0$ kann nicht für alle $x^{(0)} \in (0, b]$ das gleiche δ gewählt werden, d.h. die Funktion ist auf $[a, b]$ gleichmäßig stetig, nicht aber auf $(0, b]$

Definition 10.1.3

Eine Abbildung $f: U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ heißt **gleichmäßig stetig**, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in U$ mit $\|\underline{\mathbf{x}} - \underline{\mathbf{y}}\| < \delta$ gilt:

$$\|f(\underline{\mathbf{x}}) - f(\underline{\mathbf{y}})\| < \varepsilon.$$

Wie im obigen Beispiel ist auch i.A. richtig:

Satz 10.1.6

Jede auf einer **kompakten Menge** $U \subset \mathbb{R}^m$ definierte stetige Funktion ist **gleichmäßig stetig**.

10.2 Stetige Funktionen einer Variablen (Stetigkeit von Potenzreihen; Umkehrfunktion einer stetigen, streng monotonen Funktion; Zwischenwertsatz)

Es bezeichne nun $I \subset \mathbb{R}$ stets ein verallgemeinertes Intervall¹

In Kapitel 5 und Kapitel 7 wurden die Exponentialfunktion und die trigonometrischen Funktionen als Potenzreihen eingeführt. In diesem Paragraphen soll zunächst gezeigt werden, dass auf diese Weise stetige Funktionen definiert wurden.

Zur Diskussion von Potenzreihen muss vorausgeschickt werden (Stetigkeit in Randpunkten ist immer einseitig zu interpretieren):

Satz 10.2.1

Es sei $\{f_n\}$ eine gleichmäßig konvergente Folge stetiger Funktionen $f_n: I \rightarrow \mathbb{R}$. Dann hat diese Folge eine stetige Grenzfunktion $f: I \rightarrow \mathbb{R}$.

Beweis. Es seien $x_0 \in I$ und $\varepsilon > 0$ fixiert. Nach der Dreiecksungleichung gilt für alle $x \in I$ und für alle $n \in \mathbb{N}$

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|.$$

Gleichmäßige Konvergenz der Funktionenfolge bedeutet aber

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{\infty} = \lim_{n \rightarrow \infty} \sup_{x \in U} |f_n(x) - f(x)| = 0.$$

Ist also $\tilde{\varepsilon} = \varepsilon/3$ gegeben, so existiert ein $N = N(\tilde{\varepsilon}) \in \mathbb{N}$ mit

$$\sup_{x \in I} |f_n(x) - f(x)| < \tilde{\varepsilon} = \frac{\varepsilon}{3} \quad \text{für alle } n \geq N.$$

Ein solches n sei nun **fixiert**. Dann gilt für alle $x \in I$

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3}, \quad \text{und} \quad |f_n(x_0) - f(x_0)| < \frac{\varepsilon}{3}.$$

¹Ein **uneigentliches Intervall** ist von der Form (a, ∞) , $[a, \infty)$, $(-\infty, b)$, $(-\infty, b]$, $(-\infty, \infty)$. Ein Intervall heißt verallgemeinertes Intervall, wenn es entweder ein Intervall oder ein uneigentliches Intervall ist.

Zudem ist die Funktion f_n nach Voraussetzung stetig in x_0 , d.h. zu $\tilde{\varepsilon}$ existiert ein $\delta > 0$, sodass für alle $x \in I$ mit $|x - x_0| < \delta$

$$|f_n(x) - f_n(x_0)| < \tilde{\varepsilon} = \frac{\varepsilon}{3}.$$

Insgesamt folgt für alle $x \in I$ mit $|x - x_0| < \delta$:

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

und die Stetigkeit der Grenzfunktion ist bewiesen. \square

Bemerkung. Anhand der Funktionenfolge $\{f_n\}$, $f_n(x) = (1 - x^2)^{n/2}$ für $-1 \leq x \leq 1$ und für alle $n \in \mathbb{N}$ erkennt man, dass die Voraussetzung der gleichmäßigen Konvergenz in Satz 10.2.1 nicht weggelassen werden kann. Die Folge konvergiert nämlich punktweise (aber nicht gleichmäßig) gegen die **unstetige** Funktion

$$f(x) = \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{für } 0 < |x| \leq 1. \end{cases}$$

Analog zu Satz 10.2.1 gilt für Funktionenreihen

Satz 10.2.2

*Betrachtet sei eine auf I **gleichmäßig konvergente** Funktionenreihe $\sum_{k=1}^{\infty} f_k$, wobei alle Glieder $f_k: I \rightarrow \mathbb{R}$ **stetig** seien. Dann ist die Grenzfunktion $s = \sum_{k=1}^{\infty} f_k$ **ebenfalls stetig** auf I .*

Korollar 10.2.1

Die Exponentialfunktion und die trigonometrischen Funktionen sind stetig auf ganz \mathbb{R} .

Bemerkung. Ebenso sieht man die Stetigkeit der komplexen Exponentialfunktion und der komplexen trigonometrischen Funktionen ein.

An dieser Stelle schließt sich die Aussage an, dass die Exponentialfunktion eine bijektive Abbildung auf $(0, \infty)$ ist, so wie es in Abschnitt 5.3 schon behauptet wurde. Es gilt nämlich

Satz 10.2.3

Eine stetige, streng monotone Funktion $f: I \rightarrow \mathbb{R}$ besitzt eine stetige, streng monotone Umkehrfunktion f^{-1} , definiert auf dem verallgemeinerten Intervall $f(I)$.

Bemerkung. Ist f nicht auf einem (verallgemeinerten) Intervall definiert, so ist die Aussage des Satzes im Allgemeinen nicht mehr richtig (Übungsaufgabe als Beispiel).

Der Beweis von Satz 10.2.3 basiert auf dem [Zwischenwertsatz](#):

Satz 10.2.4

Ist $f: \mathbb{R} \supset [a, b] \rightarrow \mathbb{R}$ stetig, so nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an, d.h.: Aus $f(a) < \xi < f(b)$ folgt, dass ein $x_0 \in (a, b)$ existiert mit $f(x_0) = \xi$.

Bemerkung.

- i)* Der Beweis folgt recht leicht z.B. aus einer geeigneten Intervallschachtelung und der Stetigkeit von f .
- ii)* Es kann natürlich verschiedene Punkte x_0 mit obiger Eigenschaft geben. Im zweiten Beispiel aus Abbildung 10.17 stehen 3 Möglichkeiten zur Wahl von x_0 zur Verfügung.
- iii)* Ebenso können Werte außerhalb des Intervalls $[f(a), f(b)]$ (o.E. $f(a) < f(b)$) angenommen werden (vgl. wieder Abbildung 10.17).
- iv)* Ohne die Voraussetzung der Stetigkeit kann kein Zwischenwertsatz gelten (siehe Abbildung 10.18).

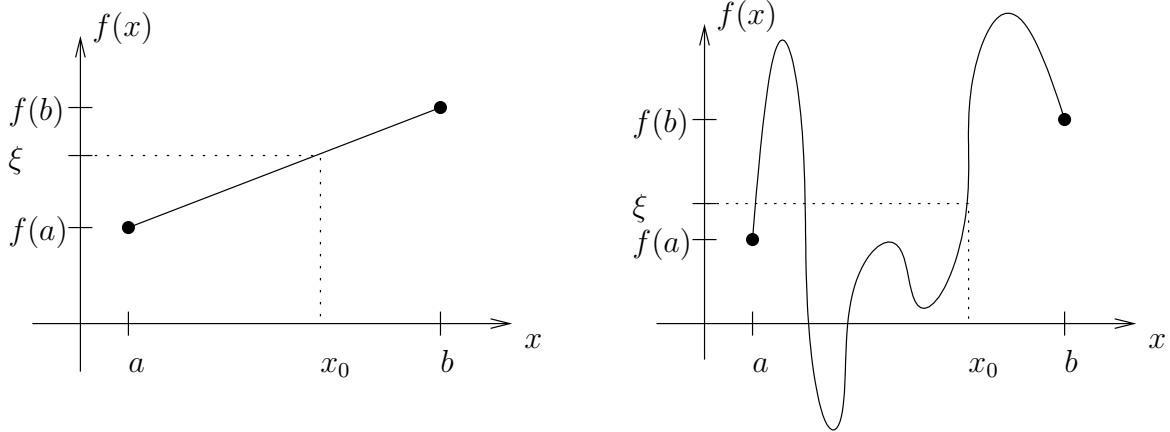


Abbildung 10.17: Zum Zwischenwertsatz.

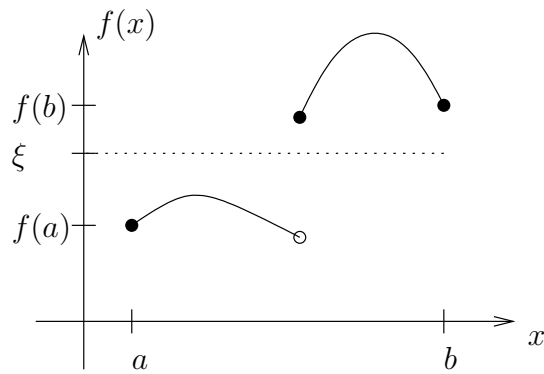


Abbildung 10.18: Die Stetigkeit muss im Zwischenwertsatz vorausgesetzt werden.

Kapitel 11

Differentialrechnung in einer Veränderlichen

11.1 Differenzierbarkeit von Funktionen einer Veränderlichen (Differenzenquotient; Ableitung; Produktquotientenregel; Kettenregel; Ableitung der Umkehrfunktion; gliedweises Differenzieren von Potenzreihen; höhere Ableitungen; die Klasse $C^k(I)$)

Der nächste Schritt zu einer infinitesimalen Analyse von Funktionen (hier einer Veränderlichen) ist die Untersuchung von Differenzierbarkeitseigenschaften einer Funktion.

Anschaulich gesprochen fallen Funktionen mit einer “Knickstelle” wie die Betragsfunktion aus der [Klasse der differenzierbaren Funktionen](#) heraus (zur Erinnerung: Die Betragsfunktion hat vom Standpunkt der Stetigkeit aus gesehen als Lipschitz-stetige Funktion noch besonders schöne Eigenschaften).

Die [Differential- und Integralrechnung](#) der heutigen Form geht auf Newton (ab 1665/66) und Leibniz (ab 1672) zurück. Der enge Zusammenhang zwischen diesen beiden Disziplinen wird im nächsten Kapitel deutlich werden.

Hier wird zunächst der Begriff der [Ableitung einer Funktion](#) eingeführt. Anschaulich gibt es zwei unterschiedliche Interpretationen des Ableitungsbegriffes.

Die **geometrische Interpretation** betrachtet in einem festen Punkt x_0 die **Tangente** an den Graphen von $f(x)$ ((affin) lineare Approximation von $f(x)$). Die Ableitung entspricht dann der Steigung dieser Tangente (vgl. Abbildung 11.1).

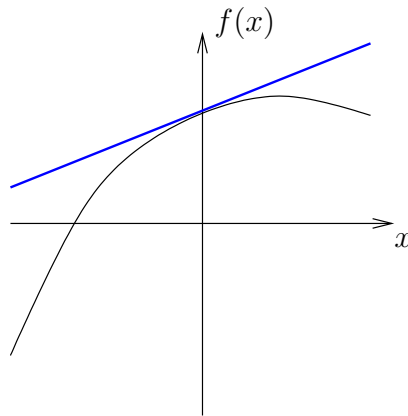


Abbildung 11.1: Zur geometrischen Interpretation der Ableitung.

Bei der **kinematischen Interpretation** betrachtet man etwa die Bewegung eines Massenpunktes im dreidimensionalen Raum in Abhängigkeit von der Zeit. Beschrieben wird diese Bewegung durch eine **Kurve** im \mathbb{R}^3 , das ist eine Abbildung $\gamma: \mathbb{R} \supset I \rightarrow \mathbb{R}^3$, wobei I ein (Zeit-) Intervall bezeichne. Die Ableitung entspricht in diesem Fall genau dem **Geschwindigkeitssvektor** des Massenpunktes zu einem gegebenen Zeitpunkt t_0 (vgl. Abbildung 11.2).

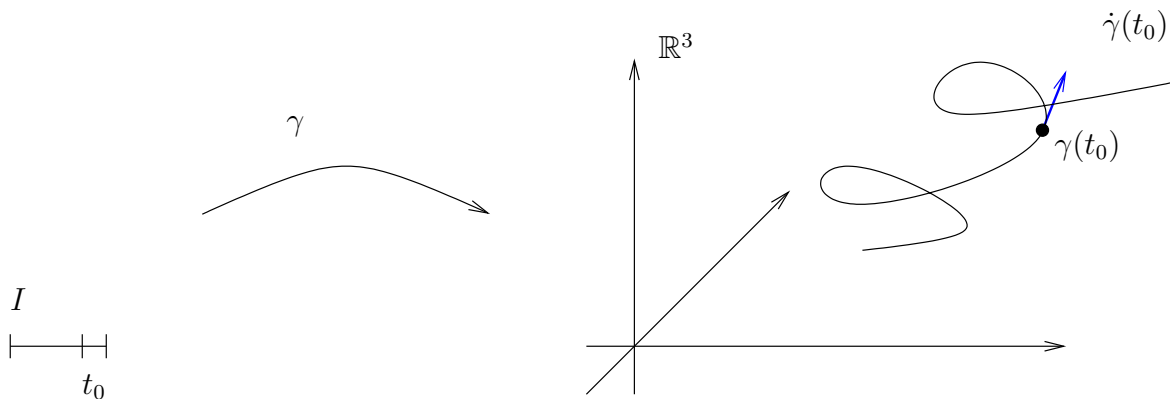


Abbildung 11.2: Zur kinematischen Interpretation der Ableitung.

Zur präzisen Einführung des Ableitungsbegriffes werden im Folgenden **stets Funktionen** $f: I \rightarrow \mathbb{R}$ betrachtet, wobei $I \subset \mathbb{R}$ wieder ein **(verallgemeinertes) Intervall bezeichnet**. Kurven im \mathbb{R}^n , d.h. Funktionen $f: I \rightarrow \mathbb{R}^n$ können völlig analog **komponentenweise** untersucht werden.

Definition 11.1.1

Es sei f eine Funktion von I nach \mathbb{R} und es sei $x_0 \in I$ ein innerer Punkt.

i) Ist $x \in I$, $x \neq x_0$, so heißt der Quotient

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

***Differenzenquotient** von f bzgl. der Punkte x und x_0 .*

*ii) Die Funktion f heißt **differenzierbar im Punkt x_0** , wenn der Grenzwert*

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

*existiert. Der Grenzwert heißt die **Ableitung** oder der **Differentialquotient von f bei x_0** und wird mit*

$$f'(x_0) \quad \text{oder} \quad \frac{df}{dx}(x_0) \quad \text{oder} \quad \frac{df}{dx}|_{x=x_0}$$

bezeichnet.

*iii) Ist f differenzierbar in jedem Punkt $x_0 \in I$, so heißt f **differenzierbar auf I** . In diesem Fall kann man die Funktion betrachten, die jedem Punkt $x \in I$ die Ableitung von f an dieser Stelle zuordnet: $f': I \rightarrow \mathbb{R}$, $I \ni x \mapsto f'(x)$. Diese Funktion heißt die **Ableitung von f** . Notation auch: $\frac{d}{dx}f$.*

Bemerkungen.

- i) Geometrisch entspricht der Differenzenquotient der Steigung der **Sekante** durch die Punkte $(x, f(x))$ und $(x_0, f(x_0))$, gegeben durch

$$s(y) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(y - x_0) .$$

Die Ableitung ist als Grenzwert der Sekantensteigungen die Steigung der Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$ (vgl. Abbildung 11.3). Die Tangente ist gegeben durch die affine Funktion

$$t(x) = f(x_0) + f'(x_0)(x - x_0) .$$

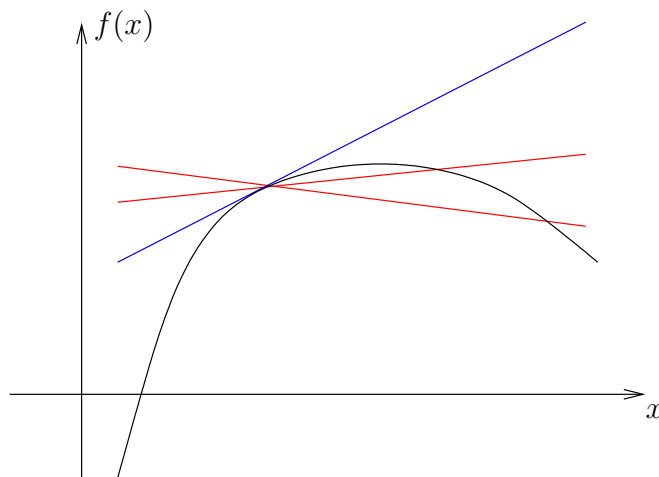


Abbildung 11.3: Die Sekantensteigungen gehen in die Steigung der Tangenten über.

- ii) Kinematisch entspricht der Differenzenquotient der **Durchschnittsgeschwindigkeit** im Zeitraum zwischen x_0 und x . Der Grenzwert gibt die **Momentangeschwindigkeit** zum Zeitpunkt x_0 .
- iii) Wie die Stetigkeit ist die Differenzierbarkeit eine lokale Eigenschaft der Funktion f .
- iv) Diese Definition **kann nicht auf Funktionen mehrerer Veränderlicher übertragen werden**, man kann nicht “durch einen Vektor teilen”.

v) Notation: Eine häufige Schreibweise ist

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} .$$

vi) Wieder können die einseitigen Grenzwerte

$$f'(x_0^+) := \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} ,$$

$$f'(x_0^-) := \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}$$

betrachtet werden. Diese heißen (falls existent) die **rechtsseitige** bzw. die **linksseitige Ableitung**.

vii) Alle folgenden Differenzierbarkeitsaussagen beziehen sich, wenn es nicht explizit anders erwähnt wird, auf innere Punkte von I . In Randpunkten müssen einseitige Grenzwerte untersucht werden.

Beispiele.

i) Eine konstante Funktion ist überall differenzierbar mit Ableitung 0.

ii) Es sei $f(x) = x$ für alle $x \in \mathbb{R}$. Der Differenzenquotient für fixiertes x_0 ist

$$\frac{x - x_0}{x - x_0} \equiv 1 ,$$

die Funktion ist in jedem Punkt differenzierbar mit Ableitung 1.

iii) Es sei $f(x) = |x|$ und $x_0 = 0$. Für die Folge $\{x_n\} = \{1/n\}$ gilt

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} = \frac{(1/n) - 0}{(1/n) - 0} = 1 .$$

Für die Folge $\{\tilde{x}_n\} = \{-1/n\}$ gilt hingegen

$$\frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} = \frac{(1/n) - 0}{(-1/n) - 0} = -1 ,$$

die Funktion ist in $x_0 = 0$ **nicht differenzierbar**.

Das letzte Beispiel belegt, dass aus der **Stetigkeit nicht die Differenzierbarkeit folgt**. Wie sieht das umgekehrt aus?

Satz 11.1.1

*Es sei $f: I \rightarrow \mathbb{R}$ im Punkt $x_0 \in I$ differenzierbar. Dann ist die Funktion f in x_0 auch stetig. Mit anderen Worten: **Differenzierbarkeit impliziert Stetigkeit.***

Beweis. Es sei $\{x_n\}$ eine Folge aus I , $x_n \neq x_0$, mit $x_n \rightarrow x_0$ für $n \rightarrow \infty$. Zu zeigen ist:

$$f(x_n) \xrightarrow{n \rightarrow \infty} f(x_0). \quad (*)$$

Nach Voraussetzung ist f in x_0 differenzierbar, dies bedeutet

$$\frac{f(x_n) - f(x_0)}{x_n - x_0} \xrightarrow{n \rightarrow \infty} f'(x_0).$$

Aus den Rechenregeln für Folgen (Satz 4.1.4) ergibt sich

$$f(x_n) - f(x_0) = \frac{f(x_n) - f(x_0)}{x_n - x_0} (x_n - x_0) \xrightarrow{n \rightarrow \infty} f'(x_0) \cdot 0 = 0,$$

also genau die Behauptung (*) und der Satz ist bewiesen. □

Rechenregeln für differenzierbare Funktionen?

Satz 11.1.2

Sind $f, g: I \rightarrow \mathbb{R}$ differenzierbare Funktionen, so sind auch cf , $c \in \mathbb{R}$, $f + g$ und fg auf I differenzierbar und es gilt:

i) $(cf)' = cf'$;

ii) $(f + g)' = f' + g'$;

iii) **Produktregel:** $(fg)' = f'g + fg'$.

iv) **Quotientenregel:**

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)},$$

falls $g(x_0) \neq 0$ in $x_0 \in I$.

Beweis. Zum Beweis wird hier exemplarisch die Produktregel betrachtet: Um die Information der Differenzierbarkeit von f und g auszunutzen, werden wie üblich Terme eingefügt und wieder abgezogen. Es ist für $x, x_0 \in I$:

$$\begin{aligned} & \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0}g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0} \\ &\xrightarrow{x \rightarrow x_0} f'(x_0)g(x_0) + f(x_0)g'(x_0) . \end{aligned}$$

Das ist gerade die Behauptung. Zu beachten ist in der letzten Zeile, dass aus Satz 11.1.1 die Stetigkeit von f bzw. g in x_0 folgt. \square

Aus Satz 11.1.2 folgt die Differenzierbarkeit von Polynomen. Für deren Ableitung gilt

Satz 11.1.3

Für alle $k \in \mathbb{Z}$ folgt (falls $x \neq 0$ für $k < 0$)

$$\frac{d}{dx}x^k = kx^{k-1} .$$

Beweis. Für $k \in \mathbb{N}_0$ wird der Beweis mittels vollständiger Induktion geführt.

Induktionsanfang: Für $k = 0$ und $k = 1$ ist die Aussage bereits verifiziert.

Induktionsschluss: Es sei nun angenommen, dass die Aussage für ein $k \in \mathbb{N}$ richtig sei. Dann folgt aus der Produktregel nach der Induktionsannahme

$$\frac{d}{dx}x^{k+1} = \frac{d}{dx}(x^k x) = kx^{k-1}x + x^k 1 = (k+1)x^k ,$$

die Aussage gilt somit auch für $k + 1$ und die Behauptung ist für alle $k \in \mathbb{N}_0$ bewiesen.

Für negative $k \in \mathbb{Z}$ folgt die Behauptung aus der Quotientenregel. \square

Wie sieht die Ableitung einer Wurzelfunktion aus?

Es sei $f(x) = \sqrt{x}$ für $x > 0$ und $x_0 > 0$ fixiert. Für $|h|$ hinreichend klein gilt

$$\begin{aligned} \frac{f(x_0 + h) - f(x_0)}{h} &= \frac{1}{h}(\sqrt{x_0 + h} - \sqrt{x_0}) \\ &= \frac{1}{h} \frac{(\sqrt{x_0 + h} - \sqrt{x_0})(\sqrt{x_0 + h} + \sqrt{x_0})}{\sqrt{x_0 + h} + \sqrt{x_0}} \\ &= \frac{1}{\sqrt{x_0 + h} + \sqrt{x_0}} \\ &\xrightarrow{h \rightarrow 0} \frac{1}{2\sqrt{x_0}}. \end{aligned}$$

Die Funktion ist demnach in alle Punkten $x_0 > 0$ differenzierbar mit Ableitung $1/(2\sqrt{x_0})$.

Im Punkt $x_0 = 0$ ist die Wurzelfunktion aber nicht differenzierbar: Die Tangentensteigungen werden für beliebig kleine x beliebig groß.

In Verallgemeinerung des Beispiels gilt

Satz 11.1.4

Ist $\alpha \in \mathbb{R}$ und $x > 0$, so ist die Funktion $f(x) = x^\alpha$ differenzierbar und es gilt

$$\frac{d}{dx} x^\alpha = \alpha x^{\alpha-1}.$$

Um die Klasse der bekannten differenzierbaren Funktionen nochmals zu erweitern, fehlt insbesondere

Satz 11.1.5 (*Kettenregel*)

Sind I_1 und I_2 zwei (verallgemeinerte) Intervalle, ist die Komposition $g \circ f$ bei $x_0 \in I_1$ definiert und sind $f: I_1 \rightarrow \mathbb{R}$, $g: I_2 \rightarrow \mathbb{R}$, differenzierbar (f bei x_0 und g bei $f(x_0)$), so ist $g \circ f$ in x_0 differenzierbar mit

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0) .$$

Beispiel. Für alle $x \in \mathbb{R}$ sei $h(x) = \sqrt{1+x^2}$. Es ist $h = g \circ f$ mit $g(y) = \sqrt{y}$ und $f(x) = 1+x^2$. Die Kettenregel gibt

$$\begin{aligned} h'(x) &= g'(f(x))f'(x) \\ &= \frac{1}{2\sqrt{1+x^2}} \cdot 2x \\ &= \frac{x}{\sqrt{1+x^2}} . \end{aligned}$$

Bemerkung. Im “Leibnizkalkül” schreibt man **formal** für die Ableitung der Funktion $z(y(x))$:

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx} ,$$

d.h. formal kürzt sich dy heraus. Diese Notation wird jedoch im Folgenden nicht verwendet.

Kann über die Ableitung einer Umkehrfunktion etwas gesagt werden?

Dazu sei $f: I \rightarrow f(I)$ bijektiv und differenzierbar, die Umkehrfunktion sei mit f^{-1} bezeichnet. Weiterhin sei für $x \in I$ $f'(x) \neq 0$ und f^{-1} differenzierbar. Es gilt nach Definition der Umkehrfunktion

$$f^{-1}(f(x)) = x .$$

Leitet man diese Gleichung mit Hilfe der Kettenregel ab, so ergibt sich

$$(f^{-1})'(f(x))f'(x) = 1 ,$$

und damit

$$(f^{-1})'(y = f(x)) = \frac{1}{f'(x)}.$$

Satz 11.1.6

Die Funktion $y = f(x)$ sei bijektiv und differenzierbar mit $f'(x) \neq 0$. Dann ist

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Beispiel. Für $x > 0$ sei die Funktion $f(x) = x^2$ betrachtet. Die Umkehrfunktion f^{-1} ist die Wurzelfunktion und Satz 11.1.6 liefert

$$(f^{-1})'(y) = \frac{d}{dy} \sqrt{y} = \frac{1}{2f^{-1}(y)} = \frac{1}{2\sqrt{y}}.$$

Das die bekannte Ableitungsregel für die Wurzelfunktion, wobei üblicherweise die unabhängige Variable wieder mit x bezeichnet wird.

Was ist mit der Exponentialfunktion und den trigonometrischen Funktionen?

Es wird analog zur Diskussion im letzten Paragraphen vorgegangen und Satz 11.1.7 vorweggeschickt.

Satz 11.1.7

Es sei $I = [a, b] \subset \mathbb{R}$ und alle unten auftretenden Ableitungen der Funktionenfolgen mögen existieren.

*i) Sind die Folgen $\{g_n\}$ und $\{g'_n\}$ auf I **gleichmäßig konvergent** und bezeichnet g den Grenzwert der Folge $\{g_n\}$, so ist auch g differenzierbar und es gilt*

$$\lim_{n \rightarrow \infty} g'_n = g'.$$

ii) Man betrachte die Reihe $\sum_{k=0}^{\infty} f_k$, wobei $s(x) = \sum_{k=0}^{\infty} f_k(x)$ (zumindest für ein $x \in I$) existiere und die Reihe der Ableitungen $\sum_{k=0}^{\infty} f'_k$ **gleichmäßig** auf I konvergiere. Dann konvergiert auch die Reihe $\sum_{k=0}^{\infty} f_k$ gleichmäßig auf I mit Grenzwert $s(x)$ und es gilt (**gliedweises Differenzieren**)

$$s'(x) = \left[\sum_{k=0}^{\infty} f_k(x) \right]' = \sum_{k=0}^{\infty} f'_k(x).$$

Beispiele.

i) Für die Exponentialfunktion sind die Voraussetzungen aus Satz 11.1.7 leicht mit den bereits bekannten Ergebnissen zu verifizieren (Übungsaufgabe) und für alle $x \in \mathbb{R}$ folgt

$$\begin{aligned} \frac{d}{dx} \exp(x) &= \left[\sum_{k=0}^{\infty} \frac{1}{k!} x^k \right]' = \sum_{k=0}^{\infty} \frac{1}{k!} [x^k]' = \sum_{k=1}^{\infty} \frac{1}{k!} k x^{k-1} \\ &= \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} = \sum_{j=0}^{\infty} \frac{1}{j!} x^j = \exp(x). \end{aligned}$$

ii) Der Logarithmus ist als die Umkehrfunktion der Exponentialfunktion definiert. Satz 11.1.6 zeigt für alle $y > 0$:

$$\begin{aligned} \frac{d}{dy} \ln(y) &= \frac{1}{\exp'(\ln(y))} \\ &= \frac{1}{\exp(\ln(y))} = \frac{1}{y}, \end{aligned}$$

wieder lautet die übliche Schreibweise

$$\ln'(x) = \frac{d}{dx} \ln(x) = \frac{1}{x} \quad \text{für alle } x > 0.$$

iii) Für fixiertes $a > 0$ und für alle $x \in \mathbb{R}$ ist die Funktion $f(x) = a^x$ definiert als

$$f(x) = \exp(x \ln(a)) .$$

Mit Hilfe der Kettenregel folgt

$$f'(x) = \exp(x \ln(a)) \ln(a) = a^x \ln(a) .$$

Analog wird die Funktion x^x abgeleitet (Übungsaufgabe).

Bemerkung. Wie im Fall der Exponentialfunktion werden die Ableitungen von \sin , \cos , \sinh und \cosh berechnet (Übungsaufgabe).

Ist eine Funktion $f: I \rightarrow \mathbb{R}$ auf ganz I differenzierbar, so ist die Funktion f' wieder eine Funktion von I nach \mathbb{R} und kann selbst auf Differenzierbarkeit untersucht werden.

Definition 11.1.2 (*höhere Ableitungen*)

i) Es sei $f: I \rightarrow \mathbb{R}$ differenzierbar mit Ableitung f' auf I . Ist f' selbst wieder differenzierbar, so heißt

$$f^{(2)}(x) := f''(x) := [f'(x)]' =: \frac{d^2}{dx^2} f(x)$$

die *zweite Ableitung* der Funktion $f(x)$.

ii) Induktiv wird definiert: Ist f auf I $(n-1)$ -mal differenzierbar, und existiert die Ableitung von $f^{(n-1)}$, so heißt

$$f^{(n)} := [f^{(n-1)}(x)]' =: \frac{d^n}{dx^n} f(x)$$

die *n-te Ableitung* der Funktion f .

iii) Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt *von der Klasse $C^k(I)$* , falls f k -mal differenzierbar ist und die Ableitung $f^{(k)}$ eine *stetige Funktion* ist.

Bemerkungen und Beispiele.

- i) $C^0(I)$ bezeichnet die Klasse der auf I stetigen Funktionen.
- ii) Nicht alle auf I differenzierbaren Funktionen sind von der Klasse $C^1(I)$. Dies belegen Beispiele wie $x^2 \sin(1/x)$, die zwar überall differenzierbar sind, deren Ableitungen aber nicht stetig sind.
- iii) Es sei $f(x) = \sqrt{1+x^2}$. Dann ist (siehe oben)

$$\begin{aligned} f''(x) &= \frac{d}{dx} \frac{x}{\sqrt{1+x^2}} = \frac{d}{dx} (1+x^2)^{-\frac{1}{2}} x \\ &= \left(-\frac{1}{2}\right) (1+x^2)^{-\frac{3}{2}} 2xx + (1+x^2)^{-\frac{1}{2}} \\ &= \frac{1}{(1+x^2)^{\frac{3}{2}}}. \end{aligned}$$

11.2 Lokale Extrema, Satz von Rolle, Mittelwertsätze (notwendige und hinreichende Bedingungen für lokale Extrema; Satz von Rolle; Ableitung und Monotonie; konvexe und konkave Funktionen; Regeln von l'Hospital)

In diesem Paragraphen wird zunächst eine **notwendige** Bedingung dafür angegeben, dass eine Funktion $f: I \rightarrow \mathbb{R}$ ein **lokales Extremum** annimmt. Wie im letzten Paragraphen ist $I \subset \mathbb{R}$ dabei stets ein (verallgemeinertes) Intervall.

Definition 11.2.1

*Eine Funktion $f: I \rightarrow \mathbb{R}$ hat an der Stelle $x_0 \in I$ ein **lokales Minimum (lokales Maximum)**, falls es ein $r > 0$ gibt mit*

$$f(x_0) \leq f(x) \quad (\text{lok. Min.}) \quad \text{bzw.} \quad f(x_0) \geq f(x) \quad (\text{lok. Max.})$$

*für alle $x \in I \cap (x_0 - r, x_0 + r)$. Gilt jeweils die strikte Ungleichung, so spricht man von einem **strengen lokalen Minimum (Maximum)**.*

Bemerkungen.

- i) Wie Abbildung 11.4 zeigt, ist ein lokales Minimum (Maximum) nicht zu verwechseln mit einem globalen Minimum (Maximum).

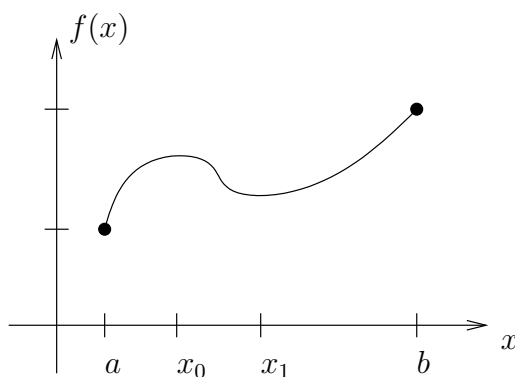


Abbildung 11.4: Ein lokales Minimum (Maximum) muss kein globales Minimum (Maximum) sein: Im Punkt x_0 hat f ein lokales Maximum, im Punkt x_1 ein lokales Minimum, in beiden Punkten liegt kein globales Extremum vor.

- ii) Lokale Minima oder lokale Maxima werden auch (lokale) **Extrema** genannt.

Satz 11.2.1

Die Funktion $f: I \rightarrow \mathbb{R}$ habe in einem *inneren Punkt* $x_0 \in I$ ein lokales Extremum und sei dort differenzierbar. Dann gilt

$$f'(x_0) = 0 .$$

Beweis. Es sei x_0 ein lokaler Minimierer von f (Maximierer analog) und es sei x_0 ein innerer Punkt von I . Dann existiert ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \subset I$.

Aus der Eigenschaft “lokaler Minimierer” folgt

$$f(x_0 + h) - f(x_0) \geq 0$$

für alle h mit $0 < |h| < \delta$, wobei δ evtl. zu verkleinern ist. D.h.

$$\frac{f(x_0 + h) - f(x_0)}{h} \geq 0 \quad \text{für alle } 0 < h < \delta,$$

$$\frac{f(x_0 + h) - f(x_0)}{h} \leq 0 \quad \text{für alle } -\delta < h < 0.$$

Dabei hat sich in der zweiten Zeile das Vorzeichen vertauscht, weil die Abschätzung $f(x_0 + h) - f(x_0) \geq 0$ durch die negative Zahl h geteilt wurde.

Aus der ersten Zeile und der Differenzierbarkeit von f folgt aber $f'(x_0) \geq 0$, aus der zweiten $f'(x_0) \leq 0$, und wie behauptet ist $f'(x_0) = 0$. \square

Bemerkungen.

- i) Anschaulich besagt der Satz, dass in einem lokalen Extremum die Tangente an den Graphen von f die Steigung 0 haben muss, d.h. die Tangente ist horizontal (vgl. Abbildung 11.5).

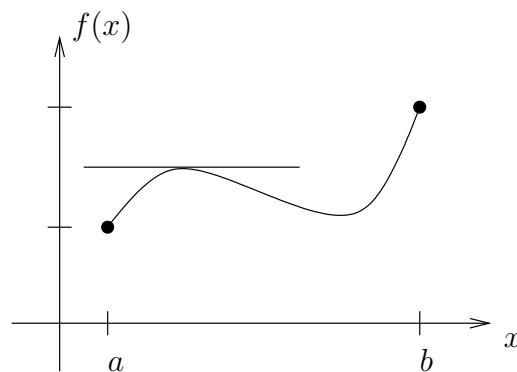


Abbildung 11.5: Die Tangente in einem lokalen Extremum.

- ii) Aus $f'(x_0) = 0$ **folgt jedoch nicht**, dass in x_0 ein lokales Extremum vorliegt (vgl. Abbildung 11.6). Ein konkretes Beispiel ist die Funktion $f(x) = x^3$ im Punkt $x_0 = 0$. Punkte mit $f'(x_0) = 0$ heißen **stationäre** oder **kritische Punkte**.
- iii) Bei der Suche nach einem **globalen** Extremum einer differenzierbaren Funktion werden zunächst **alle kritischen Punkte** ermittelt. Deren Funktionswerte sowie die Funktionswerte in den **Randpunkten** werden anschließend verglichen.

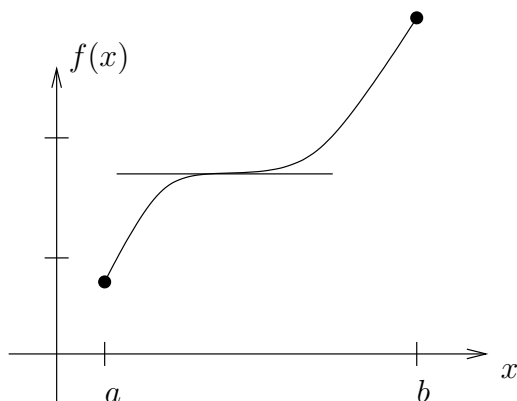


Abbildung 11.6: Kein lokales Extremum trotz horizontaler Tangente.

Ist f nicht überall differenzierbar, so müssen die **Punkte, in denen f nicht differenzierbar ist**, gesondert betrachtet werden, wie das Beispiel der Betragsfunktion zeigt. Diese nimmt im Nullpunkt ihr absolutes Minimum an, wo sie aber nicht differenzierbar ist.

Beispiel. Betrachtet sei die Funktion $f: [-2, 2] \rightarrow \mathbb{R}$, $f(x) = (1 - x^2)^2$. Es ist

$$f'(x) = -4x(1 - x^2),$$

stationäre Punkte sind $x_1 = -1$, $x_2 = 0$ und $x_3 = 1$. Es gilt

$$f(x_1) = 0, \quad f(x_2) = 1, \quad f(x_3) = 0.$$

Hier liegen in der Tat lokale Extrema vor: Ist nämlich $x \in [-1, 1]$, so ist $f(x) \leq 1 = f(x_2)$, in x_2 nimmt die Funktion ein lokales Maximum an.

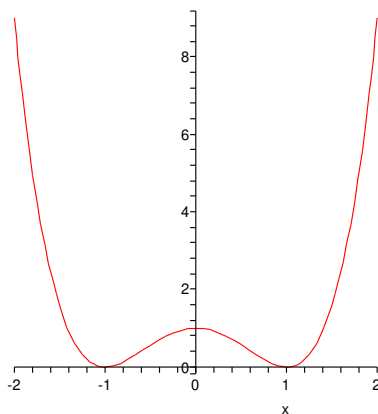
Ähnlich sieht man, dass die Funktion in x_1 und x_3 jeweils ein lokales Minimum annimmt.

Bei der Suche nach absoluten (d.h. globalen) Extrema müssen die obigen Funktionswerte mit den Werten von f in den Randpunkten,

$$f(-2) = 9, \quad f(2) = 9$$

verglichen werden: In x_1 und x_3 hat f ein globales Minimum, in x_2 hat f jedoch kein globales Maximum, die globalen Maxima werden in den Randpunkten realisiert (vgl. Abbildung 11.7).

Aus der Diskussion von Extremwerten ergibt sich direkt

Abbildung 11.7: Die Funktion $f(x) = (1 - x^2)^2$ in $[-2, 2]$.**Satz 11.2.2** (*Satz von Rolle*)¹

Es sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar auf (a, b) . Ist $f(a) = f(b)$, so existiert ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.

Beweis. Nach Satz 10.1.5 , ii), nimmt die stetige Funktion f auf dem Kompaktum $[a, b]$ sowohl ein (globales) Maximum M als auch ein Minimum m an.

Im Fall $m = M$ ist die Funktion konstant und es gilt $f' \equiv 0$.

Nun zum Fall $m < M$: In diesem Fall kann wegen $f(a) = f(b)$ höchstens einer dieser Werte in einem Randpunkt angenommen werden. In (a, b) wird also entweder das absolute Maximum oder das absolute Minimum (oder beides) angenommen. Nach Satz 11.2.1 verschwindet dort die Ableitung von f . \square

Der Satz von Rolle ist wesentlich für den Beweis der folgenden [Mittelwertsätze](#).

Der erste sagt beispielsweise aus, dass im Intervall (a, b) ein Punkt x_0 existiert, in dem die Tangente an den Graphen von f die gleiche Steigung hat wie die Sekante durch $(a, f(a))$ und $(b, f(b))$ (vgl. Abbildung 11.8).

¹M. Rolle, 1652-1719; Paris.

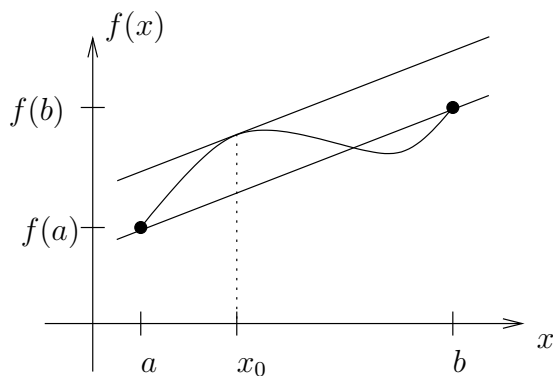


Abbildung 11.8: Zum Mittelwertsatz.

Satz 11.2.3

Die Funktionen $f, g: [a, b] \rightarrow \mathbb{R}$ seien stetig und auf (a, b) differenzierbar.

i) **Mittelwertsatz:** Dann existiert ein x_0 in (a, b) mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

ii) **Verallgemeinerter Mittelwertsatz:** Ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so existiert ein $x_0 \in (a, b)$ mit

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Eine erste wichtige Konsequenz ist

Korollar 11.2.1

Es sei f wie in Satz 11.2.3.

i) Ist $f' = 0$ für alle $x \in (a, b)$, so ist f konstant.

ii) Es ist $f'(x) \geq 0$ für alle $x \in (a, b)$ **genau dann**, wenn f **monoton wachsend** auf $[a, b]$ ist (analog: “ \leq und monoton fallend”).

iii) Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f **streng monoton wachsend** auf $[a, b]$ (analog: “< und streng monoton fallend”)

Beweis. Der Beweis ergibt sich unmittelbar aus dem Mittelwertsatz (siehe Übungen). \square

Der Mittelwertsatz liefert aber auch eine Charakterisierung zweier wichtiger Funktionenklassen: **Konvexe** und **konkave** Funktionen.

Die geometrische Vorstellung ist dabei, dass der Graph konvexer (konkaver) Funktionen stets oberhalb (unterhalb) seiner Tangenten liegt (vgl. Abbildung 11.9)

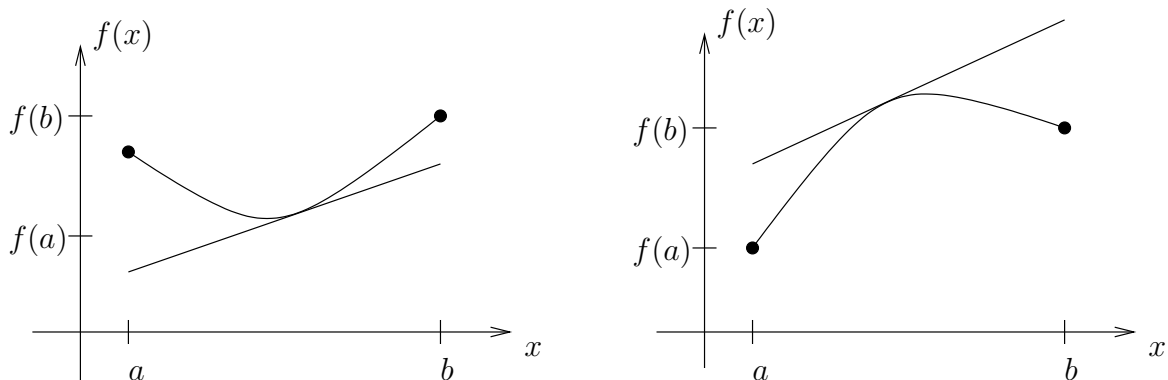


Abbildung 11.9: Der Graph einer konvexen (konkaven) Funktion liegt oberhalb (unterhalb) seiner Tangenten.

Entsprechend liegt die Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ bei einer konvexen Funktion im Intervall (a, b) oberhalb des Graphen, bei einer konkaven Funktion unterhalb (vgl. Abbildung 11.10).

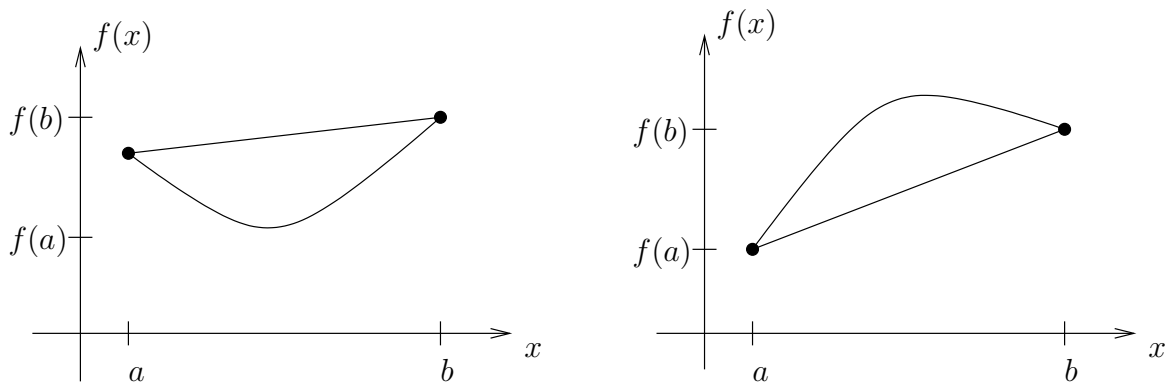


Abbildung 11.10: Die Sekante liegt in (a, b) bei einer konvexen (konkaven) Funktion oberhalb (unterhalb) des Graphen.

Definition 11.2.2

i) Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt **konvex**, falls für alle $x, y \in I$ und für alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) .$$

ii) Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt **konkav**, falls für alle $x, y \in I$ und für alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y) .$$

iii) Bei **streng konvexen** bzw. **streng konkaven** Funktionen gelten die strikten Ungleichungen.

Für Funktion der Klasse C^2 gilt

Satz 11.2.4

Es sei $f: I \rightarrow \mathbb{R}$ von der Klasse $C^2(I)$.

- i) Ist $f''(x) > 0$ für alle $x \in I$, so ist f **streng konvex**.*
- ii) Ist $f''(x) < 0$ für alle $x \in I$, so ist f **streng konkav**.*

Beispiele.

- i) Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$, ist streng konvex.*
- ii) Die Funktion f , $f(x) = \ln(x)$ für $x > 0$, ist streng konkav.*

Für eine differenzierbare konvexe Funktion $f: I \rightarrow \mathbb{R}$ zeigt man leicht (dies bedeutet gerade, dass der Graph einer differenzierbaren konvexen Funktion oberhalb seiner Tangenten liegt)

$$f(y) \geq f(x) + f'(x)(y - x) \quad \text{für alle } x, y \in I .$$

Ist $x \in I$ zusätzlich ein kritischer Punkt von f , so ist $f'(x) = 0$ und für alle $y \in I$ folgt

$$f(y) \geq f(x) .$$

In x ist ein (lokales) Minimum gefunden: Ist eine Funktion in der Nähe eines kritischen Punktes konvex (konkav), so liegt ein lokales Minimum (Maximum) vor. Dies wird als **hinreichende Bedingung** für lokale Extrema festgehalten in:

Satz 11.2.5

Es sei $f: I \rightarrow \mathbb{R}$ eine Funktion der Klasse C^2 und $x_0 \in I$ sei ein innerer Punkt von I mit $f'(x_0) = 0$.

- i) Ist $f''(x_0) > 0$, so hat f in x_0 ein **strenges lokales Minimum**.*
- ii) Ist $f''(x_0) < 0$, so hat f in x_0 ein **strenges lokales Maximum**.*

Bemerkung. Der Satz macht **keine Aussage** über den Fall $f''(x_0) = 0$.

Beispiel. Im oben diskutierten Beispiel $f(x) = (1 - x^2)^2$ ist

$$f''(x) = 12x^2 - 4,$$

in den Punkten $x_1 = -1$, $x_3 = 1$ liegen also tatsächlich (strenge) lokale Minima vor, in $x_2 = 0$ ein (strenges) lokales Maximum (vgl. Abbildung 11.7).

Bemerkung. Charakteristisch für eine Funktion sind auch **Wendepunkte**: Eine Funktion $f: I \rightarrow \mathbb{R}$ hat in einem inneren Punkt x_0 einen Wendepunkt, wenn die Funktion in einer Umgebung von x_0 auf einer Seite konvex ist, auf der anderen konkav. Es geht also eine Linkskurve in eine Rechtskurve über (oder umgekehrt, vgl. Abbildung 11.11).

Ist f von der Klasse C^3 , so gilt in einem Wendepunkt immer $f''(x_0) = 0$, ist umgekehrt $f''(x_0) = 0$ und $f'''(x_0) \neq 0$, so ist x_0 ein Wendepunkt.

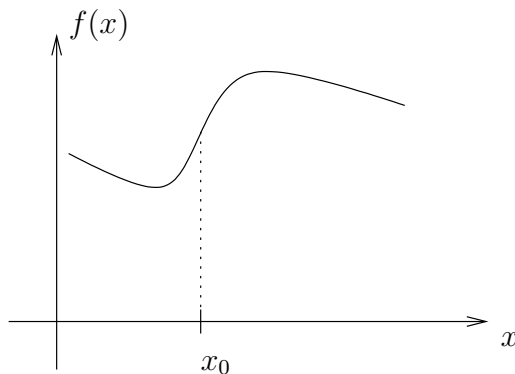


Abbildung 11.11: Ein Wendepunkt.

Eine Folgerung aus dem zweiten Mittelwertsatz ist schließlich, dass Grenzwerte von Quotienten berechnet werden können, die sich nicht aus den bisher bekannten Regeln erschließen: Beispielsweise kann

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$$

berechnet werden mit Hilfe von

Satz 11.2.6 (*Regeln von l'Hospital*)

Es seien $f, g: I \rightarrow \mathbb{R}$ differenzierbar und x_0 sei ein innerer Punkt aus I . Ist $f(x_0) = g(x_0) = 0$ und ist $g'(x_0) \neq 0$ für $x \neq x_0$, so gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

falls der zweite Grenzwert existiert.

Beweisidee. Wie bereits erwähnt folgt der Satz aus dem verallgemeinerten Mittelwertsatz. \square

Bemerkungen und Beispiele.

i) Es ist

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1.$$

ii) Die Regel kann auch mehrfach hintereinander angewandt werden, so gilt

$$\lim_{x \rightarrow x_0} \frac{1 - \cos(x)}{x^2} = \lim_{x \rightarrow x_0} \frac{\sin(x)}{2x} = \lim_{x \rightarrow x_0} \frac{\cos(x)}{2} = \frac{1}{2}.$$

iii) Eine analoge Aussage gilt auch für Grenzwerte " $x \rightarrow \infty$ ", falls f und g im Unendlichen gegen Null konvergieren, z.B.

$$\begin{aligned} \lim_{x \rightarrow \infty} [\sqrt{x(x+1)} - x] &= \lim_{x \rightarrow \infty} \frac{\sqrt{1 + \frac{1}{x}} - 1}{\frac{1}{x}} \\ &= \lim_{x \rightarrow \infty} \frac{\frac{1}{2\sqrt{1 + \frac{1}{x}}} \left[-\frac{1}{x^2}\right]}{\left[-\frac{1}{x^2}\right]} \\ &= \lim_{x \rightarrow \infty} \frac{1}{2\sqrt{1 + \frac{1}{x}}} = \frac{1}{2}. \end{aligned}$$

iv) Ebenso gelten die Regeln für Grenzwerte der Form $x \rightarrow x_0$, $f(x)$, $g(x) \rightarrow \pm\infty$. Wie überall sind dabei auch einseitige Grenzwerte zugelassen:

$$\lim_{x \rightarrow 0^+} x \ln(x) = \lim_{x \rightarrow 0^+} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \rightarrow 0^+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \rightarrow 0^+} (-x) = 0 .$$

11.3 Numerische Differentiation (Extrapolation zum Limes $h \rightarrow 0$; zentraler Differenzenquotient)

In diesem Paragraphen wird kurz aufgezeigt, wie mit Hilfe der Polynominterpolation (vgl. Kapitel 3.2) Ableitungen auch numerisch berechnet werden können.

Die Grundidee ist eine [Extrapolation zum Limes \$h \rightarrow 0\$](#) .

Die allgemeine Aufgabenstellung lautet: Gegeben sei eine Funktion $g: (0, b] \rightarrow \mathbb{R}$, $b > 0$, wobei für $0 < h \leq b$ die Funktion g explizit ausgewertet werden könne. Zudem sei g als stetige Funktion in den Nullpunkt fortsetzbar. Gesucht ist eine numerische Näherung für

$$\lim_{h \rightarrow 0^+} g(h) =: g(0) .$$

Beispiele.

i) Man approximiere numerisch

$$\lim_{h \rightarrow 0^+} \frac{\cosh(h) - 1}{\sinh(h)} .$$

ii) Für eine differenzierbare Funktion $f: I \rightarrow \mathbb{R}$ setzt man in einem fixierten inneren Punkt $x_0 \in I$

$$g(h) := \frac{f(x_0 + h) - f(x_0)}{h} .$$

Gesucht ist hier eine Näherung für den Grenzwert des Differenzenquotienten, d.h. für die Ableitung von f in x_0 .

Lösungsidee.

- i)* Zunächst werden die Werte $g(h_i)$, $i = 0, \dots, n$ für geeignete Stützstellen $h_0 > h_1 > \dots > h_n > 0$ und für geeignetes $n \in \mathbb{N}$ (nicht zu groß) berechnet.
- ii)* Man löse die Interpolationsaufgabe von Lagrange (vgl. Kapitel 3.2) zu diesen Daten, d.h. betrachtet sei das eindeutig bestimmte Interpolationpolynom vom Grad n durch die Punkte $(h_0, g(h_0)), \dots, (h_n, g(h_n))$.
- iii)* Ist die Güte der Approximation hinreichend gut, so sollte $p_n(0)$ eine gute Näherungslösung für $g(0)$ sein.

Beispiele.

- i)* Gesucht ist eine Näherung für

$$\lim_{h \rightarrow 0^+} \frac{h^{3/2}}{\sin(h)}.$$

Bekanntlich ist dieser Grenzwert 0. Die Rechnung liefert

| h_i | 1/8 | 1/16 | 1/32 | 1/64 |
|-------------------------|------------|------------|------------|------------|
| $h_i^{3/2} / \sin(h_i)$ | 0.35447578 | 0.25016283 | 0.17680547 | 0.12500509 |

Anstatt den Wert 0.12500509 als Näherung zu wählen, legt man durch die berechneten Werte das Interpolationspolynom vom Grad 3, wie es in der Abbildung 11.12 rot angedeutet ist. Der exakte Verlauf ist blau gekennzeichnet, eine recht gute Übereinstimmung ist zu erkennen.

Für $p_3(0)$ ergibt sich der Wert 0.05938205, der zwar nicht um Größenordnungen aber doch deutlich näher bei der 0 liegt. Dieses Beispiel (wie auch das nächste) wurde mit “Maple” ausgewertet.

- ii)* Mit den gleichen Stützstellen wie oben soll

$$\lim_{h \rightarrow 0^+} \frac{\cosh(h) - 1}{\sinh(h)} = 0$$

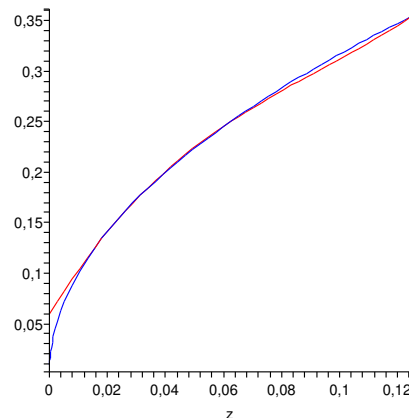


Abbildung 11.12: Extrapolation von $h^{3/2}/\sin(h)$ im Limes $h \rightarrow 0$.

näherungsweise berechnet werden. Die Daten lauten:

| h_i | 1/8 | 1/16 | 1/32 | 1/64 |
|-----------------------------------|------------|------------|------------|------------|
| $\frac{\cosh(h_i)-1}{\sinh(h_i)}$ | 0.06241875 | 0.03123983 | 0.01562373 | 0.00781235 |

Die Extrapolation zum Limes $h \rightarrow 0$ liefert in diesem Fall $p_3(0) = 3.638 \cdot 10^{-8}$, die Approximation ist also tatsächlich deutlich besser als der Wert 0.00781235. In Abbildung 11.13 kann der Unterschied zwischen dem Approximationspolynom und der exakten Lösung nicht mehr aufgelöst werden (es sind tatsächlich beide Funktionsverläufe eingetragen).

Natürlich approximiert das Interpolationspolynom die Funktion nur in der Nähe des Ursprungs so gut. Zur Illustration dieser Tatsache sind die Verläufe von Interpolationspolynom und exakter Lösung in Abbildung 11.14 in einem größeren Bereich dargestellt. Diese Abweichungen für große h sind für die gegebene Fragestellung jedoch irrelevant.

Zur numerischen Differentiation einer C^1 -Funktion $f: I \rightarrow \mathbb{R}$ kann wie oben angemerkt der Differenzenquotient $g(h) = (f(x_0 + h) - f(x_0))/h$ in einem fixierten (inneren) Punkt $x_0 \in I$ herangezogen werden.

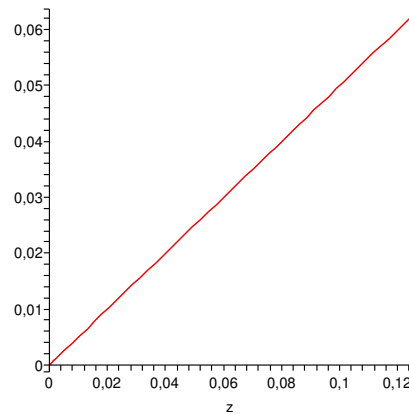
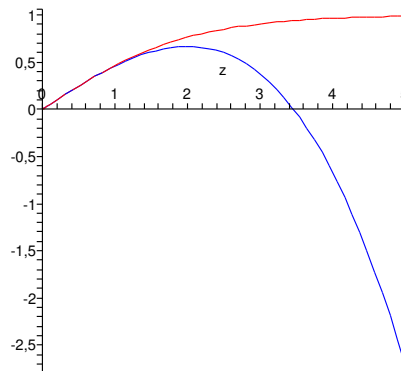

 Abbildung 11.13: Extrapolation von $(\cosh(h) - 1) / \sinh(h)$ im Limes $h \rightarrow 0$.


Abbildung 11.14: Interpolation und exakte Lösung in einem größeren Bereich.

Die Extrapolation zum Limes $h \rightarrow 0$ soll einen Näherungswert für die Ableitung von f im Punkt x_0 liefern. Aber:

Beobachtung. Betrachtet sei der **zentrale Differenzenquotient**

$$\delta(h) := \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Es konvergiert $\delta(h)$ im Limes $h \rightarrow 0$ ebenfalls gegen $f'(x_0)$ (Beweis?).

Der wesentliche Vorteil im Hinblick auf diesen Paragraphen ist nun, dass der zentrale Differenzenquotient ein **gerade Funktion in h** ist, es gilt

$$\frac{f(x_0 + (-h)) - f(x_0 - (-h))}{-2h} = \frac{f(x_0 + h) - f(x_0 - h)}{2h}.$$

Dies legt die Idee nahe, mit geraden Polynomen, d.h. mit **Polynomen in Abhängigkeit von h^2 zu approximieren.**

Beispiel. Es sei $f(x) = \sin(x)$. Mittels des zentralen Differenzenquotienten soll eine Näherung für $f'(0) = 1$ berechnet werden.

Der zentrale Differenzenquotient von f in 0 lautet

$$\delta(h) = \frac{\sin(h) - \sin(-h)}{2h} = \frac{\sin(h)}{h} = g(h^2) .$$

Obwohl in den Differenzenquotienten das Argument h einzusetzen ist, handelt es sich tatsächlich um eine Funktion von h^2 . (Beispiel: Ist $\psi(h) = h^3$, so ist $\psi(h)/h = h^2$, also $\psi(h)/h = g(h^2)$ mit $g(x) = x$.)

Es sei jetzt $h_0 = 1/8$, $h_1 = 1/16$, $h_2 = 1/32$. Zu berechnen ist als Näherungswert für $\sin'(0)$ das Interpolationspolynom der Ordnung 2 von g zu den Stützstellen h_i^2 , $i = 0, 1, 2$. Dies ist an der Stelle 0 auszuwerten. Der Algorithmus von Neville (vgl. Kapitel 3.2) liefert die in Tabelle 11.1 angegebenen Werte. Man erkennt deutlich die Güte der berechneten Approximation.

| h_i^2 | $g(h_i^2) = \sin(h_i)/h_i$ | $k = 1$ | $k = 2$ |
|------------|----------------------------|----------------------------|----------------------------|
| $(1/8)^2$ | $p_{0,0}(0) = 0.997397867$ | $p_{0,1}(0) = 0.999999491$ | $p_{0,2}(0) = 1.000000001$ |
| $(1/16)^2$ | $p_{1,1}(0) = 0.999349085$ | $p_{1,2}(0) = 0.999999969$ | |
| $(1/32)^2$ | $p_{2,2}(0) = 0.999837248$ | | |

Tabelle 11.1: Numerische Differentiation nach dem Algorithmus von Neville.

Weitere Beispiele werden in den Übungen besprochen.

Kapitel 12

Eindimensionale Integration

12.1 Das bestimmte (Riemannsches) Integral (Zerlegung; Feinheit; Riemannsches Zwischensumme; Untersumme; Ober-summe; Verfeinerung; Unterintegral; Oberintegral; Riemann integrierbare Funktionen; Integrabilitätskriterien; Rechenregeln für integrierbare Funktionen; stückweise stetige Funktionen; Integration auf Teilintervallen; Mittelwertsatz der Integralrechnung; orientiertes Riemannsches Integral; Flächeninhalt)

Hier sei **stets** $I = [a, b]$ und $f: I \rightarrow \mathbb{R}$ eine beschränkte Funktion.

Idee zur Einführung des bestimmten Integrals. Das bestimmte Integral $\int_a^b f(x) dx$ ist der Flächeninhalt der vom Graphen von f und der x -Achse eingeschlossenen Punktmenge F , so wie es in Abbildung 12.1 angedeutet ist.

Problem. Das Problem bei dieser suggestiven Vorstellung ist aber: Wie ist der Flächeninhalt überhaupt definiert, wenn es sich nicht um ein elementargeometrisches Objekt handelt?

Vorgehensweise. In der Tat sieht die Vorgehensweise wie folgt aus:

- i)* Man definiere zunächst das bestimmte Integral.
- ii)* Mit Hilfe des bestimmten Integrals können Flächeninhalte definiert werden.
- iii)* Dabei soll “einfachen” geometrischen Objekten (Rechtecken,

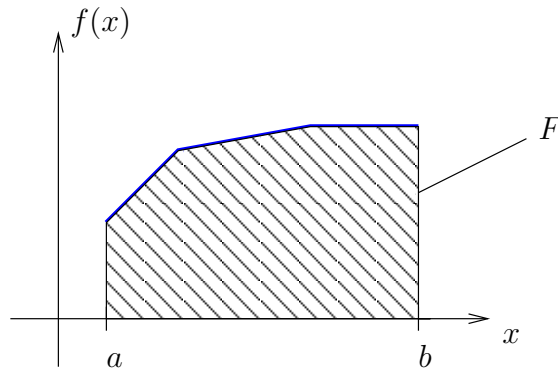


Abbildung 12.1: Zur suggestiven Vorstellung des bestimmten Integrals.

Dreiecken etc.) der Flächeninhalt zugeordnet werden, der mit der Elementargeometrie konsistent ist.

- iv)* Um dies zu erreichen, approximiert man bei der Definition des Integrals die Menge F mit Rechtecken, denen elementargeometrisch ein Flächeninhalt zugeordnet werden kann.
- v)* Schließlich ist zu untersuchen, ob zum Grenzwert übergegangen werden kann.

Konkret wird definiert:

Definition 12.1.1 ¹

- i)* Eine **Zerlegung** \mathcal{Z} des Intervalls $I = [a, b]$ in Teilintervalle I_j , $j = 1, \dots, n$, der Länge $|I_j|$ ist eine **Menge von Punkten**

$$\mathcal{Z} = \{a = x_0 < x_1 < x_2 < \dots < x_n = b\} .$$

Hierbei ist $I_j := [x_{j-1}, x_j]$, $\Delta x_j := x_j - x_{j-1} = |I_j|$, $j = 1, \dots, n$, und

$$\Delta(\mathcal{Z}) := \max\{\Delta x_1, \dots, \Delta x_n\}$$

heißt die **Feinheit der Zerlegung** \mathcal{Z} .

¹B. Riemann, 1826-1866; Göttingen.

ii) Für alle $j = 1, \dots, n$, sei $\xi_j \in I_j$ gewählt. Setzt man $\xi = (\xi_1, \dots, \xi_n)$, so heißt

$$S_{\mathcal{Z}}(f) := S_{\mathcal{Z}}(f, \xi) := \sum_{j=1}^n f(\xi_j) \Delta x_j$$

eine **Riemannsche Zwischensumme**. Dabei sei $f: I \rightarrow \mathbb{R}$ eine beschränkte Funktion.

iii) Ist für $j = 1, \dots, n$

$$\begin{aligned} \underline{m}_j &:= \inf\{f(x) : x \in I_j\}, \\ \bar{m}_j &:= \sup\{f(x) : x \in I_j\}, \end{aligned}$$

so heißt

$$\underline{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \underline{m}_j \Delta x_j$$

die **Untersumme**,

$$\bar{S}_{\mathcal{Z}}(f) := \sum_{j=1}^n \bar{m}_j \Delta x_j$$

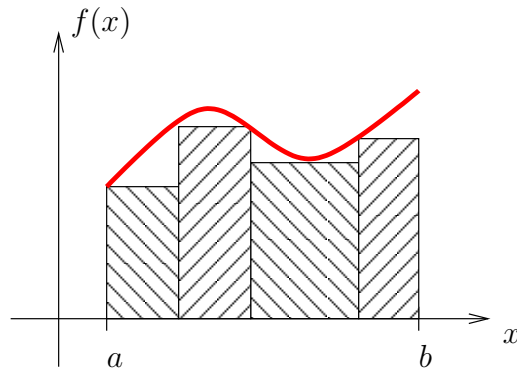
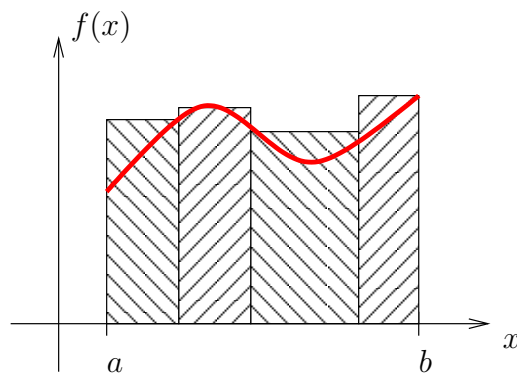
die **Obersumme** von f zur Zerlegung \mathcal{Z} .

Bemerkung. Nach Definition gilt für jede Zerlegung \mathcal{Z} und für jedes ξ wie in Definition 12.1.1

$$\underline{S}_{\mathcal{Z}}(f) \leq S_{\mathcal{Z}}(f, \xi) \leq \bar{S}_{\mathcal{Z}}(f).$$

Definition 12.1.2

Eine Zerlegung \mathcal{Z}^* von I heißt **Verfeinerung** der Zerlegung \mathcal{Z} von I ($\mathcal{Z} \subseteq \mathcal{Z}^*$), falls alle Teilpunkte von \mathcal{Z} auch Teilpunkte von \mathcal{Z}^* sind.

Abbildung 12.2: Eine Untersumme von f .Abbildung 12.3: Eine Obersumme von f .**Übung.**

i) Ist $\mathcal{Z} \subseteq \mathcal{Z}^*$, so folgt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{S}_{\mathcal{Z}^*}(f) \leq \overline{S}_{\mathcal{Z}^*}(f) \leq \overline{S}_{\mathcal{Z}}(f) .$$

ii) Für zwei beliebige Zerlegungen $\mathcal{Z}_1, \mathcal{Z}_2$ von I gilt

$$\underline{S}_{\mathcal{Z}_1}(f) \leq \overline{S}_{\mathcal{Z}_2}(f) .$$

iii) Für jede Zerlegung \mathcal{Z} von I ist

$$|I| \inf_I f \leq \underline{S}_{\mathcal{Z}}(f) \leq \overline{S}_{\mathcal{Z}}(f) \leq |I| \sup_I f .$$

Mit diesen Vorüberlegungen kann sinnvoll definiert werden (warum?):

Definition 12.1.3

Ist $f: I = [a, b] \rightarrow \mathbb{R}$ beschränkt, so sind das *Unteriorintegral* $\underline{\mathcal{I}}$ und das *Oberintegral* $\overline{\mathcal{I}}$ definiert als

$$\begin{aligned}\underline{\mathcal{I}}(f) &:= \sup\{\underline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\}, \\ \overline{\mathcal{I}}(f) &:= \inf\{\overline{S}_{\mathcal{Z}}(f) : \mathcal{Z} \text{ ist Zerlegung von } I\}.\end{aligned}$$

Die Funktion heißt *(Riemann) integrierbar* auf I , falls gilt

$$\underline{\mathcal{I}}(f) = \overline{\mathcal{I}}(f) =: \mathcal{I}(f) =: \int_I f(x) \, dx =: \int_a^b f(x) \, dx.$$

Bemerkungen.

i) Für eine beliebige Zerlegung \mathcal{Z} von I gilt

$$\underline{S}_{\mathcal{Z}}(f) \leq \underline{\mathcal{I}}(f) \leq \overline{\mathcal{I}}(f) \leq \overline{S}_{\mathcal{Z}}(f).$$

ii) Die Klasse der auf I integrierbaren Funktionen wird mit $\mathcal{R}(I)$ bezeichnet.

iii) Es heißt $\int_a^b f(x) \, dx$ das bestimmte Integral von f zwischen den Grenzen a und b .

Beispiele.

i) Es sei $f(x) = c = \textit{konstant}$ auf $[a, b]$. Es folgt für jede Zerlegung \mathcal{Z} von I

$$\underline{S}_{\mathcal{Z}}(f) = \sum_{j=1}^n c \Delta x_j = c(b-a) = \overline{S}_{\mathcal{Z}}(f).$$

Somit ist f auf $[a, b]$ integrierbar mit

$$\int_a^b c \, dx = c(b-a).$$

ii) Es sei $I = [0, 1]$ und $f(x) = x$ auf I . Zu $n \in \mathbb{N}$ betrachte man die Zerlegung

$$\mathcal{Z}_n = \left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}.$$

Für die Ober- bzw. Untersummen gilt

$$\begin{aligned} \underline{S}_{\mathcal{Z}_n}(f) &= \sum_{j=1}^n \frac{j-1}{n} \left[\frac{j}{n} - \frac{j-1}{n} \right] = \frac{1}{n^2} \sum_{j=1}^n (j-1) \\ &= \frac{1}{n^2} \frac{n(n-1)}{2} = \frac{1}{2} - \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}, \end{aligned}$$

$$\overline{S}_{\mathcal{Z}_n}(f) = \frac{1}{2} + \frac{1}{2n} \xrightarrow{n \rightarrow \infty} \frac{1}{2}.$$

Um daraus tatsächlich die Integrierbarkeit von f auf I mit

$$\int_0^1 x \, dx = \frac{1}{2}$$

zu folgern, wird das erste Kriterium zitiert aus

Satz 12.1.1 (*Integrabilitätskriterien*)

Für eine beschränkte Funktion $f: I \rightarrow \mathbb{R}$ gilt

$$f \in \mathcal{R}(I) \Leftrightarrow \left\{ \begin{array}{l} \text{Zu jedem } \varepsilon > 0 \text{ gibt es eine} \\ \text{Zerlegung } \mathcal{Z} \text{ von } I \text{ mit} \\ \overline{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon. \end{array} \right.$$

$$f \in \mathcal{R}(I) \Leftrightarrow \left\{ \begin{array}{l} \text{Zu jedem } \varepsilon > 0 \text{ existiert ein } \delta > 0 \\ \text{derart, dass } \overline{S}_{\mathcal{Z}}(f) - \underline{S}_{\mathcal{Z}}(f) < \varepsilon \text{ für} \\ \text{jede Zerlegung } \mathcal{Z} \text{ von } I \text{ mit } \Delta(\mathcal{Z}) < \delta. \end{array} \right.$$

Bemerkung. Aus dem zweiten Integrabilitätskriterium folgt: Ist f integrierbar auf I , so konvergiert jede Folge Riemannscher Zwischensummen gegen $\mathcal{I}(f)$, sofern die Feinheit der Zerlegungsfolge gegen Null konvergiert.

Weitere integrierbare Funktionen?

Satz 12.1.2

Es seien $f, g \in \mathcal{R}(I)$ und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

i) $\alpha f + \beta g \in \mathcal{R}(I)$ mit

$$\mathcal{I}(\alpha f + \beta g) = \alpha \mathcal{I}(f) + \beta \mathcal{I}(g) .$$

ii) $fg \in \mathcal{R}(I)$.

iii) $|f| \in \mathcal{R}(I)$, wobei $|f|(x) := |f(x)|$.

iv) Ist $|g| \geq c$ für eine Konstante $c > 0$, so ist auch $f/g \in \mathcal{R}(I)$.

v) Ist für alle $x \in I$ die Ungleichung $f(x) \leq g(x)$ richtig, so folgt

$$\mathcal{I}(f) \leq \mathcal{I}(g) .$$

vi) $|\mathcal{I}(f)| \leq \mathcal{I}(|f|)$.

Beweis. Der Beweis folgt leicht aus der Definition des Riemannsches Intergals. □

Aus Satz 12.1.2 folgt insbesondere die Integrierbarkeit von Polynomen auf $[a, b]$. Die Frage nach weiteren bekannten Funktionenklassen in $\mathcal{R}(I)$ beantwortet

Satz 12.1.3

Für eine beschränkte Funktion $f: [a, b] \rightarrow \mathbb{R}$ gilt:

- i) Ist f **monoton**, so folgt $f \in \mathcal{R}(I)$.
- ii) Ist f **stetig**, so folgt $f \in \mathcal{R}(I)$.
- iii) Ist f **stückweise stetig** (vgl. Definition 12.1.4), so folgt $f \in \mathcal{R}(I)$.

Definition 12.1.4

Eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ heißt **stückweise stetig**, wenn es eine Zerlegung $\mathcal{Z} = \{a = x_0 < x_1 < \dots < x_n = b\}$ gibt, sodass f in jedem Teilintervall (x_{j-1}, x_j) , $j = 1, \dots, n$, stetig ist und die **einseitigen Grenzwerte** $f(a^+)$, $f(b^-)$, $f(x_j^+)$, $f(x_j^-)$, $j = 1, \dots, n-1$, **existieren**.

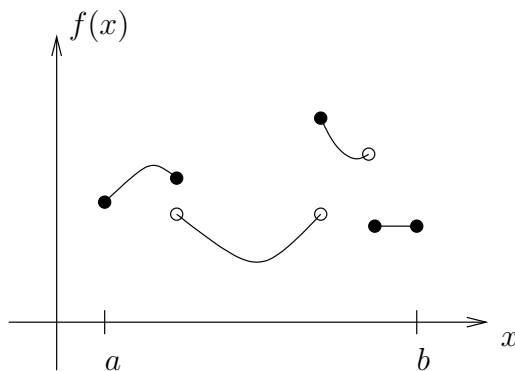


Abbildung 12.4: Eine stückweise stetige Funktion.

Integration auf Teilintervallen?

Satz 12.1.4

Es sei $f \in \mathcal{R}(I)$.

i) Dann ist f auch auf jedem Teilintervall $I' \subset I$ integrierbar.

ii) Ist I in endlich viele Teilintervalle I_1, I_2, \dots, I_n zerlegt, die höchstens Randpunkte gemeinsam haben, so gilt

$$\int_I f(x) \, dx = \sum_{j=1}^n \int_{I_j} f(x) \, dx .$$

Wie in der Differentialrechnung gibt es auch einen [Mittelwertsatz der Integralrechnung](#). Er folgt aus dem Zwischenwertsatz für stetige Funktionen.

Satz 12.1.5

Es sei f stetig auf $I = [a, b]$ und es sei $p \in \mathcal{R}(I)$ mit $p(x) \geq 0$ für alle $x \in I$. Dann gibt es ein $\xi \in (a, b)$ mit

$$\int_a^b f(x)p(x) \, dx = f(\xi) \int_a^b p(x) \, dx .$$

Bemerkung. Wie lautet der Spezialfall $p \equiv 1$ und wie ist dieser geometrisch zu interpretieren?

Den Abschluss dieses Abschnittes bilden zwei Bemerkungen zum [orientierten Riemannschen Integral](#) und zum eingangs diskutierten [Flächeninhalt](#).

Bemerkungen.

- i) Das orientierte Riemannsche Integral ist wie folgt definiert: Ist $f \in \mathcal{R}(I)$ und sind α, β zwei beliebige Punkte aus $[a, b]$, so wird gesetzt (mit $I' = [\alpha, \beta]$, falls $\alpha < \beta$, $I'' = [\beta, \alpha]$, falls $\beta < \alpha$)

$$\int_{\alpha}^{\beta} f(x) \, dx := \int_{I'} f(x) \, dx, \quad \text{falls } \alpha < \beta;$$

$$\int_{\alpha}^{\beta} f(x) \, dx := - \int_{I''} f(x) \, dx, \quad \text{falls } \alpha > \beta;$$

$$\int_{\alpha}^{\beta} f(x) \, dx := 0, \quad \text{falls } \alpha = \beta.$$

- ii) Ist f **stetig** auf $I = [a, b]$, sind $\alpha, \beta \in I$, $\alpha < \beta$ und ist $f(x) > 0$ für alle $\alpha < x < \beta$, so ist

$$A(\alpha, \beta) := \int_{\alpha}^{\beta} f(x) \, dx$$

positiv und als der Flächeninhalt der **ebenen Menge**

$$F(\alpha, \beta) := \{\underline{\mathbf{x}} \in \mathbb{R}^2 : \alpha \leq x_1 \leq \beta, 0 \leq x_2 \leq f(x_1)\}$$

definiert (vgl. Abbildung 12.1).

12.2 Der Hauptsatz der Differential- und Integralrechnung (Stammfunktion; unbestimmtes Integral)

Die Idee des **Hauptsatzes der Differential- und Integralrechnung** sei hier anhand eines einfachen Beispiels erläutert: Man betrachte für $x \geq 0$ die Funktion $f(x) = \alpha x$, $\alpha > 0$, den Flächeninhalt ($x_0 > 0$)

$$A(x_0) = \int_0^{x_0} f(x) \, dx$$

und ebenso für $h > 0$

$$A(x_0 + h) = \int_0^{x_0+h} f(x) \, dx.$$

Untersucht werden soll die relative Änderung “ $\Delta A/\Delta x$ ”, d.h. gesucht ist der Differenzenquotient von A an der Stelle x_0 ,

$$\frac{1}{h} \left[\int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right].$$

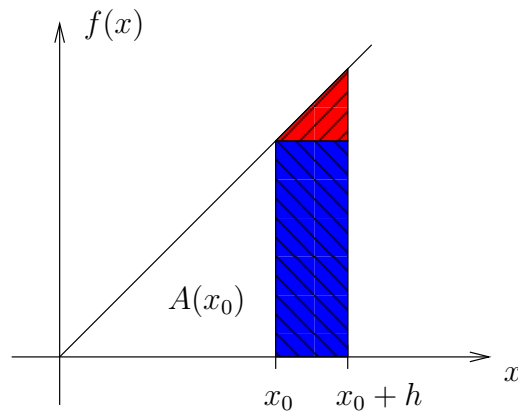


Abbildung 12.5: Zum Hauptsatz der Differential- und Integralrechnung.

Die Differenz $A(x_0 + h) - A(x_0)$ setzt sich zusammen aus dem Flächeninhalt des in Abbildung 12.5 blau gekennzeichneten Rechtecks R und dem des roten Dreiecks D . Der Flächeninhalt von R ist $\alpha x_0 h$, der von D ist $\alpha h^2/2$. Es ergibt sich

$$\begin{aligned} \frac{1}{h} \left[\int_0^{x_0+h} f(x) \, dx - \int_0^{x_0} f(x) \, dx \right] &= \frac{1}{h} \left[\alpha x_0 h + \frac{\alpha h^2}{2} \right] \\ &\xrightarrow{h \rightarrow 0} \alpha x_0 = f(x_0). \end{aligned}$$

Die Ableitung des Flächeninhalts $A(x_0)$ nach x_0 entspricht dem Funktionswert $f(x_0)$.

Präzise sieht die Situation wie folgt aus:

Definition 12.2.1

*Es seien $F, f: [a, b] \rightarrow \mathbb{R}$. Ist F differenzierbar auf $[a, b]$ und gilt $F'(x) = f(x)$ für alle $x \in [a, b]$, so heißt F eine **Stammfunktion** von f .*

Bemerkungen.

- i) Wie üblich ist in den Randpunkten des Intervalls einseitige Differenzierbarkeit gemeint.
- ii) Mit F sind auch alle Funktionen $F_C = F(x) + C$, $C \in \mathbb{R}$, Stammfunktionen von f .
- iii) Die Differenz zweier Stammfunktionen ist konstant.

Satz 12.2.1 (Hauptsatz der Differential- und Integralrechnung)

Es sei $f: I = [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt:

i) Die Funktion

$$F(x) := \int_a^x f(t) dt$$

*ist eine **Stammfunktion von $f(x)$** .*

ii) Ist $F(x) \in C^1(I)$ eine Stammfunktion von $f(x)$, so gilt

$$\int_a^b f(t) dt = F(b) - F(a) =: \left[F(x) \right]_a^b =: F(x) \Big|_a^b .$$

Beweis.

Ad i). Zu $h \neq 0$, $x, x + h \in [a, b]$ betrachtet man den Differenzenquotienten

$$\Delta_h F(x) := \frac{1}{h} [F(x+h) - F(x)] , \quad F(x) = \int_a^x f(t) dt .$$

Zu zeigen ist für das fixierte x :

$$\Delta_h F(x) \xrightarrow{h \rightarrow 0} f(x) .$$

In der Tat gilt

$$\Delta_h F(x) = \frac{1}{h} \int_x^{x+h} f(t) dt ,$$

$$f(x) = \frac{1}{h} \int_x^{x+h} f(x) dt .$$

In der zweiten Gleichung ist zu beachten, dass $f(x)$ für fixiertes x eine Konstante ist und bzgl. der **Integrationsvariablen** t integriert wird. Es folgt

$$\begin{aligned} |\Delta_h F(x) - f(x)| &\leq \frac{1}{|h|} \left| \int_x^{x+h} (f(t) - f(x)) dt \right| \\ &\leq \frac{1}{|h|} |h| \sup \{ |f(t) - f(x)| : \\ &\quad t \in [a, b], |x - t| \leq |h| \} \\ &\xrightarrow{h \rightarrow 0} 0 , \end{aligned}$$

wobei im letzten Schritt die Stetigkeit von f auf $[a, b]$ ausgenutzt wurde. Damit ist der erste Teil des Satzes gezeigt.

Ad *ii*). Da die Differenz zweier Stammfunktionen konstant ist, ist nach *i*) **jede Stammfunktion F** von f von der Form

$$F(x) = \int_a^x f(t) dt + C , \quad C \in \mathbb{R} .$$

Insbesondere ist

$$F(b) = \int_a^b f(t) dt + C ,$$

$$F(a) = \int_a^a f(t) dt + C = C ,$$

d.h.

$$F(b) - F(a) = \int_a^b f(t) dt + C - C = \int_a^b f(t) dt ,$$

und der Hauptsatz ist bewiesen. \square

Bemerkungen.

- i) Die **Gesamtheit aller Stammfunktionen** einer stetigen Funktion f heißt das **unbestimmte Integral** der Funktion f . Notation:

$$\{F : F \text{ ist Stammfunktion von } f\} =: \int f(x) \, dx =: F(x) + C .$$

- ii) Das Aufsuchen einer Stammfunktion wird als **Integration** von f bezeichnet. Nach dem Hauptsatz macht die Differentiation die Integration wieder rückgängig.

Beispiele. Eine Tabelle zur Integration bekannter Funktionen wird in den Übungen besprochen. Als ein einfaches Beispiel sei hier die Funktion $f(x) = x^n$, $n \in \mathbb{N}$, betrachtet. Diese hat eine Stammfunktion

$$F(x) = \frac{x^{n+1}}{n+1} ,$$

woraus folgt

$$\int_a^b x^n \, dx = \frac{1}{n+1} [b^{n+1} - a^{n+1}] .$$

12.3 Integrationstechniken (partielle Integration; Substitution; Partialbruchzerlegung)

In diesem Paragraphen werden einige Integrationstechniken angesprochen, die die Integration in konkreten Fällen ermöglichen.

I. Einfache Integrationstechniken. Es seien f , g stetig auf $[a, b]$. Dann gilt (auf $[a, b]$)

$$\int cf(x) \, dx = c \int f(x) \, dx , \quad c \in \mathbb{R} ;$$

$$\int (f(x) + g(x)) \, dx = \int f(x) \, dx + \int g(x) \, dx ;$$

$$\int \frac{f'(x)}{f(x)} \, dx = \ln(|f(x)|) + C , \quad \text{falls } f \text{ keine Nullstellen hat ,}$$

wobei in der letzten Zeile auch f' als stetig angenommen wurde.

Beispiele.

i) Mit $a_j \in \mathbb{R}$, $j = 1, \dots, n$, ist

$$\int \left[\sum_{j=0}^n a_j x^j \right] dx = \sum_{j=0}^n a_j \int x^j dx = \sum_{j=0}^n a_j \frac{x^{j+1}}{j+1} + C .$$

ii) Es sei $f(x) = \tan(x)$, $x \in [a, b] \subset (-\pi/2, \pi/2)$. In diesem Intervall hat der Kosinus keine Nullstelle und es gilt

$$\begin{aligned} \int \tan(x) dx &= \int \frac{\sin(x)}{\cos(x)} dx = - \int \frac{(-\sin(x))}{\cos(x)} dx \\ &= -\ln(\cos(x)) + C . \end{aligned}$$

II. Partielle Integration. Hierbei handelt es sich um die Umkehrung der Produktregel.

Satz 12.3.1

Sind $f, g: [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar, so gilt

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx ,$$

$$\int_a^b f(x)g'(x) dx = \left[f(x)g(x) \right]_a^b - \int_a^b f'(x)g(x) dx .$$

Beispiele.

i) Mit $f(x) = x$ und $g'(x) = e^x$ folgt

$$\int x e^x dx = x e^x - \int e^x dx = x e^x - e^x + C .$$

ii) Mit $f(x) = \ln(x)$ und $g'(x) = 1$ ergibt sich auf $[a, b] \subset (0, \infty)$

$$\begin{aligned} \int \ln(x) \, dx &= \int \ln(x) 1 \, dx = \ln(x)x - \int \frac{1}{x} x \, dx \\ &= x \ln(x) - x + C . \end{aligned}$$

iii) Ist $f(x) = g'(x) = \cos(x)$, so sieht man

$$\begin{aligned} \int \cos^2(x) \, dx &= \int \cos(x) \cos(x) \, dx \\ &= \cos(x) \sin(x) + \int \sin^2(x) \, dx \\ &= \sin(x) \cos(x) + \int 1 \, dx - \int \cos^2(x) \, dx , \end{aligned}$$

also

$$\int \cos^2(x) \, dx = \frac{1}{2} [\sin(x) \cos(x) + x] + C .$$

III Substitution. Die Regel zur Substitution ist eine Folgerung aus der Kettenregel.

Satz 12.3.2

Es sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und $g: [c, d] \rightarrow [a, b]$ sei stetig differenzierbar. Dann gilt für alle $t_0, t_1 \in [c, d]$

$$\int_{g(t_0)}^{g(t_1)} f(x) \, dx = \int_{t_0}^{t_1} f(g(t))g'(t) \, dt .$$

Für die Substitutionsregel gibt es **zwei Lesarten**:

Erste Lesart. Hier ist $\int_{t_0}^{t_1} h(t) \, dt$ gesucht, wobei angenommen sei (f, g wie oben)

$$h(t) = f(g(t))g'(t) .$$

Dann gilt

$$\int_{t_0}^{t_1} h(t) \, dt = \int_{g(t_0)}^{g(t_1)} f(x) \, dx .$$

Kann dabei die rechte Seite explizit berechnet werden, so ist auch die linke Seite bekannt.

Beispiel. Gesucht sei

$$\int_0^1 (1+t^2)^n t \, dt .$$

Die Funktion $h(t) := (1+t^2)^n t$ lässt sich in der Form

$$h(t) = \frac{1}{2} (1+t^2)^n \frac{d}{dt} (1+t^2)$$

schreiben, also

$$2h(t) = f(g(t))g'(t)$$

mit $f(x) := x^n$ und $g(t) := 1+t^2$.

Es folgt

$$\begin{aligned} \int_0^1 (1+t^2)^n t \, dt &= \frac{1}{2} \int_0^1 f(g(t))g'(t) \, dt \\ &= \frac{1}{2} \int_1^2 x^n \, dx \\ &= \left. \frac{x^{n+1}}{2(n+1)} \right|_1^2 = \frac{2^{n+1} - 1}{2(n+1)} . \end{aligned}$$

Zweite Lesart. Man lese die Substitutionsregel von links nach rechts, d.h. gesucht ist $\int_{x_0}^{x_1} f(x) \, dx$ durch geeignete **Variablentransformation**.

Mit anderen Worten ist eine Transformation

$$x = g(t) , \quad t_0 \leq t \leq t_1 ,$$

gesucht, die die Eigenschaft

$$x_0 = g(t_0) \quad \text{und} \quad x_1 = g(t_1)$$

haben muss. Hierbei sei $g: [t_0, t_1] \rightarrow [x_0, x_1]$ (o.E. gelte $t_0 < t_1$ und $x_0 < x_1$) stetig differenzierbar und bijektiv (etwa $g'(t) > 0$ für alle $t \in [t_0, t_1]$).

Bezeichnet dann $\psi := g^{-1}$ die Inverse von g , so gilt mit $t_0 = \psi(x_0)$ und $t_1 = \psi(x_1)$

$$\int_{x_0}^{x_1} f(x) \, dx = \int_{\psi(x_0)}^{\psi(x_1)} f(g(t))g'(t) \, dt .$$

Bemerkung. Das **formale Vorgehen** sieht wie folgt aus: Man substituiert $x = g(t)$ in $\int_{x_0}^{x_1} f(x) dx$ und schreibt in der Leibnizschen Weise

$$g'(t) = \frac{dx}{dt}, \quad \text{also} \quad dx = g'(t) dt .$$

Die untere Integrationsgrenze $x_0 = g(t_0)$ wird zu $t_0 = \psi(x_0)$, die obere transformiert sich analog.

Beispiel. Gesucht sei zu fixiertem $r > 0$

$$\int_0^r \sqrt{r^2 - x^2} dx .$$

Es bietet sich die Substitution

$$x = g(t) := r \sin(t), \quad 0 \leq t \leq \frac{\pi}{2},$$

an, wobei $g'(t) = r \cos(t)$. Man beachtet ($\psi = g^{-1}$)

$$0 = g(0), \quad r = g(\pi/2), \quad \text{d.h.} \quad \psi(0) = 0, \quad \psi(r) = \frac{\pi}{2} .$$

Es folgt

$$\begin{aligned} \int_0^r \sqrt{r^2 - x^2} dx &= \int_0^{\pi/2} \sqrt{r^2 - r^2 \sin^2(t)} r \cos(t) dt \\ &= r^2 \int_0^{\pi/2} \cos^2(t) dt = \frac{\pi r^2}{4} . \end{aligned}$$

IV. Partialbruchzerlegung. Nun sollen rationale Funktionen der Form

$$R(x) = \frac{b_m x^m + \dots + b_1 x + b_0}{a_n x^n + \dots + a_1 x + a_0}$$

mit reellen Koeffizienten $a_0, \dots, a_n, b_0, \dots, b_m$ integriert werden.

Schritt 1. Ist $m \geq n$, so beginne man mit einer Polynomdivision

Beispiel. Ist

$$R(x) = \frac{x^3 + 3x^2 - 4x + 2}{x^3 - x^2 - x + 1},$$

so liefert die Rechnung

$$R(x) = \underbrace{1}_{P(x)} + \frac{4x^2 - 3x + 1}{\underbrace{x^3 - x^2 - x + 1}_{p(x)/q(x)}},$$

es ist also

$$R(x) = P(x) + \frac{p(x)}{q(x)},$$

wobei das Polynom $P(x)$ elementar zu integrieren ist und wobei $\text{grad } p < \text{grad } q$.

Schritt 2. Es bleiben echt gebrochene rationale Funktionen zu integrieren,

$$\int \frac{p(x)}{q(x)} dx =? \quad \text{mit } \text{grad } p < \text{grad } q =: n.$$

Nun zerfällt $q(x)$ im Reellen zwar nicht in Linearfaktoren, nach dem Fundamentalsatz der Algebra kann $q(x)$ aber geschrieben werden als

$$q(x) = c \prod_{\nu=1}^k (x - \lambda_{\nu})^{r_{\nu}} \prod_{\mu=1}^l q_{\mu}(x)^{s_{\mu}},$$

wobei k, l, r_{ν}, s_{μ} natürliche Zahlen sind, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, $\lambda_{\nu} \neq \lambda_{\tilde{\nu}}$ für $\nu \neq \tilde{\nu}$, und wobei

$$q_{\mu}(x) = x^2 + 2b_{\mu}x + c_{\mu}, \quad q_{\mu} \neq q_{\tilde{\mu}} \text{ für } \mu \neq \tilde{\mu},$$

irreduzible quadratische Polynome sind, also $b_{\mu}^2 < c_{\mu}$ (ohne reelle Nullstellen).

Mit anderen Worten wird q geschrieben als Produkt von Linearfaktoren (mit Vielfachheiten) und irreduziblen quadratischen Polynomen (ebenfalls mit Vielfachheiten).

Es folgt die Existenz von (eindeutig bestimmten) reellen Zahlen $A_{\rho\nu}$, $B_{\sigma\mu}$, $C_{\sigma\mu}$ derart, dass

$$\frac{p(x)}{q(x)} = \sum_{\nu=1}^k \sum_{\rho=1}^{r_{\nu}} \frac{A_{\rho\nu}}{(x - \lambda_{\nu})^{\rho}} + \sum_{\mu=1}^l \sum_{\sigma=1}^{s_{\mu}} \frac{B_{\sigma\mu} + C_{\sigma\mu}x}{(x^2 + 2b_{\mu}x + c_{\mu})^{\sigma}},$$

genannt die Partialbruchzerlegung der rationalen Funktion p/q .

Mittels dieser Partialbruchzerlegung ist die Integration rationaler Funktionen zurückgeführt auf Integrale der Form (Warum ist das richtig?)

$$\int \frac{1}{(x - \lambda)^\rho} dx ;$$

$$\int \frac{2x + b}{(x^2 + bx + c)^\sigma} dx ;$$

$$\int \frac{1}{(x^2 + 2bx + c)^\sigma} dx , \quad b^2 < c .$$

Die Integration der ersten beiden Typen ist bereits bekannt, der dritte wird in den Übungen besprochen.

Beispiele.

i) Es sei

$$\frac{p(x)}{q(x)} = \frac{x^2 + 1}{x^3 - 2x^2 - 5x + 6} .$$

Dann hat q drei verschiedene reelle Nullstellen $x_1 = 1$, $x_2 = -2$, $x_3 = 3$ der Vielfachheit 1,

$$q(x) = (x - 1)(x + 2)(x - 3) .$$

Der **Ansatz** lautet in diesem Beispiel

$$\frac{p(x)}{q(x)} = \frac{A_1}{x - 1} + \frac{A_2}{x + 2} + \frac{A_3}{x - 3} .$$

Er liefert

$$\begin{aligned} \frac{x^2 + 1}{(x - 1)(x + 2)(x - 3)} &= \frac{p(x)}{q(x)} \\ &= \frac{A_1(x + 2)(x - 3) + A_2(x - 1)(x - 3) + A_3(x - 1)(x + 2)}{(x - 1)(x + 2)(x - 3)} \\ &= \frac{x^2(A_1 + A_2 + A_3) + x(-A_1 - 4A_2 + A_3)}{(x - 1)(x + 2)(x - 3)} \\ &\quad + \frac{(-6A_1 + 3A_2 - 2A_3)}{(x - 1)(x + 2)(x - 3)} . \end{aligned}$$

Ein **Koeffizientenvergleich** zeigt

$$\begin{aligned} A_1 + A_2 + A_3 &= 1, \\ -A_1 - 4A_2 + A_3 &= 0, \\ -6A_1 + 3A_2 - 2A_3 &= 1, \end{aligned}$$

mit der Lösung

$$A_1 = -\frac{1}{3}, \quad A_2 = \frac{1}{3}, \quad A_3 = 1.$$

Insgesamt ist

$$\begin{aligned} &\int \frac{x^2 + 1}{x^3 - 2x^2 - 5x + 6} dx \\ &= -\frac{1}{3} \int \frac{1}{x-1} dx + \frac{1}{3} \int \frac{1}{x+2} dx + \int \frac{1}{x-3} dx \\ &= -\frac{1}{3} \ln|x-1| + \frac{1}{3} \ln(|x+2|) + \ln(|x-3|) + C. \end{aligned}$$

ii) Es sei nun

$$\frac{p(x)}{q(x)} = \frac{1-x}{x^2(x^2+1)}.$$

Hier ist $x^2 + 1$ ein irreduzibles quadratisches Polynom und $x_1 = 0$ ist zweifache Nullstelle von q . Dementsprechend lautet der Ansatz

$$\begin{aligned} \frac{1-x}{x^2(x^2+1)} &= \frac{p(x)}{q(x)} \\ &= \sum_{\rho=1}^2 \frac{A_\rho}{x^\rho} + \frac{B+Cx}{x^2+1} \\ &= \frac{A_1}{x} + \frac{A_2}{x^2} + \frac{B+Cx}{x^2+1}. \end{aligned}$$

Es ergibt sich

$$1-x = A_1x(x^2+1) + A_2(x^2+1) + Bx^2 + Cx^3,$$

der Koeffizientenvergleich lautet

$$\begin{aligned} A_1 + C &= 0, \\ A_2 + B &= 0, \\ A_1 &= -1, \\ A_2 &= 1. \end{aligned}$$

Mit der Lösung

$$A_1 = -1, \quad A_2 = 1, \quad B = -1, \quad C = 1$$

ist in diesem Beispiel gezeigt

$$\begin{aligned} &\int \frac{1-x}{x^2(x^2+1)} dx \\ &= -\int \frac{1}{x} dx + \int \frac{1}{x^2} dx + \int \frac{x-1}{x^2+1} dx \\ &= -\ln(|x|) - \frac{1}{x} + \frac{1}{2} \ln(x^2+1) - \arctan(x) + C. \end{aligned}$$

12.4 Uneigentliche Integrale (lokal integrierbare Funktionen; Konvergenzkriterien)

Bisher wurden für die Integration stets zwei Einschränkungen angenommen:

- i) $I = [a, b]$, d.h. I ist abgeschlossen und beschränkt.
- ii) Auch die Funktion $f: I \rightarrow \mathbb{R}$ ist beschränkt.

In diesem Abschnitt geht es darum, ob auch Integrale wie

$$\int_0^\infty e^{-x^2} dx \quad \text{oder} \quad \int_0^1 \frac{dx}{\sqrt{1-x^2}}$$

sinnvoll definiert werden können.

Fall I. Unbeschränktes Integrationsintervall. Es sei $I = [a, \infty)$ und $f \in \mathcal{R}([a, b])$ für alle b mit $a < b < \infty$. Die Funktion f heißt dann **lokal integrierbar**.

Es existiere also für alle $b \in I$

$$F(b) := \int_a^b f(x) \, dx .$$

Definition 12.4.1

i) Falls der Grenzwert $\lim_{b \rightarrow \infty} F(b)$ existiert, so heißt dieser das **uneigentliche Integral** von f über $[a, \infty)$, man sagt, das uneigentliche Integral existiert oder **konvergiert**. Notation:

$$\int_a^\infty f(x) \, dx := \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx .$$

Andernfalls heißt $\int_a^\infty f(x) \, dx$ **divergent**.

ii) Das uneigentliche Integral $\int_a^\infty f(x) \, dx$ heißt **absolut konvergent**, falls $\int_a^\infty |f(x)| \, dx$ konvergiert.

Wie üblich ist die Konvergenz mit Hilfe geeigneter Kriterien zu überprüfen.

Satz 12.4.1 (Konvergenzkriterien)

Es sei $f: I = [a, \infty) \rightarrow \mathbb{R}$ lokal integrierbar.

i) Das uneigentliche Integral $\int_a^\infty f(x) \, dx$ existiert **genau dann**, wenn zu jedem $\varepsilon > 0$ ein $\xi \geq a$ existiert mit

$$\left| \int_b^{b'} f(x) \, dx \right| < \varepsilon \quad \text{für alle } b, b' > \xi .$$

ii) Absolute Konvergenz des uneigentlichen Integrals impliziert Konvergenz.

iii) Es gilt das **Majorantenkriterium**: Ist für alle $x \in I$ die Ungleichung $|f(x)| \leq g(x)$ richtig und ist $\int_a^\infty g(x) \, dx$ konvergent, so ist $\int_a^\infty f(x) \, dx$ absolut konvergent.

iv) Gilt umgekehrt $0 \leq g(x) \leq f(x)$ für alle $x \in I$ und divergiert das uneigentliche Integral $\int_a^\infty g(x) \, dx$, so divergiert auch $\int_a^\infty f(x) \, dx$.

Bemerkungen.

i) Uneigentliche Integrale der Form $\int_{-\infty}^b f(x) \, dx$ sind analog zu diskutieren.

ii) Ein uneigentliches Integral der Form $\int_{-\infty}^\infty f(x) \, dx$ konvergiert, wenn für ein beliebiges $a \in \mathbb{R}$ sowohl $\int_a^\infty f(x) \, dx$ als auch $\int_{-\infty}^a f(x) \, dx$ konvergiert. Es ist dann

$$\int_{-\infty}^\infty f(x) \, dx := \int_a^\infty f(x) \, dx + \int_{-\infty}^a f(x) \, dx .$$

Beispiele.

i) Es ist für alle $1 < b < \infty$

$$\begin{aligned} \int_1^b \frac{dx}{x^\alpha} &= \begin{cases} \frac{1}{1-\alpha} \frac{1}{x^{\alpha-1}} \Big|_1^b & \text{falls } \alpha \neq 1, \\ \ln(|x|) \Big|_1^b & \text{falls } \alpha = 1, \end{cases} \\ &= \begin{cases} \frac{b^{1-\alpha} - 1}{1-\alpha} & \text{falls } \alpha \neq 1, \\ \ln(b) & \text{falls } \alpha = 1. \end{cases} \end{aligned}$$

Demnach ist $\int_1^\infty \frac{dx}{x^\alpha}$ **divergent für $\alpha \leq 1$** . Es ist aber

$$\int_1^\infty \frac{dx}{x^\alpha} = \frac{1}{\alpha - 1} \quad \text{für } \alpha > 1.$$

ii) Betrachtet sei das uneigentliche Integral

$$\int_0^\infty \frac{\sin(x)}{x} dx,$$

wobei es im Nullpunkt keine Schwierigkeiten gibt, da $\sin(x)/x$ bekanntlich stetig in 0 fortgesetzt werden kann. Dieses uneigentliche Integral ist **konvergent, aber nicht absolut konvergent**.

Zum Konvergenzbeweis beachtet man, dass für beliebige $0 < b < b'$ gilt (partielle Integration)

$$\int_b^{b'} \frac{\sin(x)}{x} dx = -\frac{\cos(x)}{x} \Big|_b^{b'} - \int_b^{b'} \frac{\cos(x)}{x^2} dx,$$

somit kann abgeschätzt werden

$$\begin{aligned} \left| \int_b^{b'} \frac{\sin(x)}{x} dx \right| &\leq \frac{1}{b} + \frac{1}{b'} + \int_b^{b'} \frac{1}{x^2} dx \\ &< \frac{2}{b} + \frac{1}{b} \xrightarrow{b \rightarrow \infty} 0. \end{aligned}$$

Nach dem ersten Konvergenzkriterium aus Satz 12.4.1 konvergiert das Integral.²

Es gilt jedoch für alle $k \in \mathbb{N}$:

$$\begin{aligned} \int_0^{k\pi} \left| \frac{\sin(x)}{x} \right| dx &= \sum_{j=1}^k \int_{(j-1)\pi}^{j\pi} \left| \frac{\sin(x)}{x} \right| dx \\ &\geq \sum_{j=1}^k \frac{1}{j\pi} \int_{(j-1)\pi}^{\pi} |\sin(x)| dx \\ &= \frac{2}{\pi} \sum_{j=1}^k \frac{1}{j} \\ &\rightarrow \infty \quad \text{falls } k \rightarrow \infty, \end{aligned}$$

²Eulers Formel (1781) besagt $\int_0^\infty \frac{\sin(x)}{x} dx = \frac{\pi}{2}$.

das Integral konvergiert wie behauptet nicht absolut.

iii) Zur Analyse des Integrals $\int_0^\infty e^{-x^2} dx$ beachtet man zunächst, dass aus $(x-1)^2 \geq 0$ folgt $x^2 \geq 2x-1$, also für alle $x > 0$

$$e^{-x^2} \leq e^{1-2x} =: g(x).$$

Für alle $0 < b$ ist

$$\int_0^b g(x) dx = \int_0^b e^{1-2x} dx = -\frac{1}{2} e^{1-2x} \Big|_0^b = \frac{e}{2} (1 - e^{-2b}),$$

die Funktion g ist eine konvergente Majorante, das dritte Kriterium aus Satz 12.4.1 impliziert (absolute) Konvergenz von $\int_0^\infty e^{-x^2} dx$.

Fall II. Unbeschränkte Funktionen. Es sei nun $f: [a, b) \rightarrow \mathbb{R}$ bzw. $(a, b] \rightarrow \mathbb{R}$ auf allen kompakten Teilintervallen von $[a, b)$ (bzw. von $(a, b]$) integrierbar, aber nicht notwendigerweise für “ $x \rightarrow b^-$ ” (bzw. für “ $x \rightarrow a^+$ ”) beschränkt.

Definition 12.4.2

Es sei f wie oben. Falls der Grenzwert

$$\lim_{\xi \rightarrow b^-} \int_a^\xi f(x) dx \quad \text{bzw.} \quad \lim_{\xi \rightarrow a^+} \int_\xi^b f(x) dx$$

*existiert, so heißt dieser Grenzwert das **uneigentliche Integral** $\int_a^b f(x) dx$, das Integral **konvergiert**. Das Integral heißt **absolut konvergent**, falls $\int_a^b |f(x)| dx$ konvergiert.*

Bemerkung. Im Fall II gelten analoge Konvergenzkriterien wie im ersten Fall (insbesondere das Majorantenkriterium).

Beispiele.

i) Für $0 < \xi < 1$ ist

$$\int_\xi^1 \frac{dx}{\sqrt{x}} = \left[2\sqrt{x} \right]_\xi^1 = 2 - 2\sqrt{\xi} \xrightarrow{\xi \rightarrow 0^+} 2,$$

demnach ist $\int_0^1 \frac{dx}{\sqrt{x}} = 2$, insbesondere ist das Integral konvergent.

ii) Analog folgt die Konvergenz von

$$\int_0^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha} \quad \text{für } 0 < \alpha < 1 .$$

12.5 Parameterabhängige Integrale (Gamma-Funktion; partielle Differenzierbarkeit)

Zur Motivation betrachte man ein prominentes

Beispiel. Die **Gamma-Funktion** $\Gamma: (0, \infty) \rightarrow \mathbb{R}$ ist definiert via

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt ,$$

d.h. das Argument x der Gamma-Funktion ist ein Parameter in dem Integranden, über den zu integrieren ist.

Bemerkung. Die Gamma-Funktion erfüllt die **Funktionalgleichung**

$$\Gamma(x+1) = x\Gamma(x) , \quad x > 0 ,$$

und es folgt für alle $n \in \mathbb{N}$

$$\Gamma(n) = (n-1)! .$$

Frage: Hängt die Gamma-Funktion beispielsweise stetig oder differenzierbar von x ab?

Allgemein sieht die Situation wie folgt aus: Man betrachte das parameterabhängige (hier eigentliche) Integral über $I_y := [c, d]$

$$F(x) := \int_c^d f(x, y) dy , \quad x \in I_x = [a, b] . \quad (*)$$

Dabei sei $f: I_x \times I_y \rightarrow \mathbb{R}$ für jedes **fixierte** $x \in I_x$ als Funktion von y über I_y integrierbar.

Satz 12.5.1

- i) Ist $f(x, y)$ stetig auf $I_x \times I_y$, so existiert das Integral (*) für alle $x \in I_x$ und $F(x)$ ist stetig auf I_x .
- ii) Ist zusätzlich $f(x, y)$ nach x stetig partiell differenzierbar (vgl. Definition 12.5.1), so ist $F(x)$ auf I_x stetig differenzierbar und es gilt

$$F'(x) = \int_c^d \frac{\partial f}{\partial x}(x, y) dy \quad \text{für alle } x \in I_x .$$

Nachzuliefern bleibt

Definition 12.5.1

Die Funktion $f(x, y)$ (wie oben) heißt nach x stetig partiell differenzierbar, falls $f(x, y)$ für jedes fixierte $y \in I_y$ nach x differenzierbar ist (also als Funktion nur abhängig von x) und diese Ableitung (Bezeichnung $\frac{\partial f}{\partial x}(x, y)$) eine stetige Funktion auf $I_x \times I_y$ ist.

Bemerkung. Für uneigentliche parameterabhängige Integrale gelten analoge Aussagen, falls die auftretenden Integrale gleichmäßig konvergieren. Beispielsweise ist

$$\Gamma'(x) = \int_0^{\infty} e^{-t} t^{x-1} \ln(t) dt .$$

Beispiel. Es sei $f: I_x \times I_y \rightarrow \mathbb{R}$, $I_x = [-1, 1]$, $I_y = [0, 1]$,

$$f(x, y) = e^{xy} .$$

Dann ist für $x \in I_x$

$$F(x) := \int_0^1 f(x, y) dy = \int_0^1 e^{xy} dy = \begin{cases} 1 & \text{für } x = 0 , \\ \frac{e^x - 1}{x} & \text{für } x \neq 0 . \end{cases}$$

Es ist $f(x, y)$ stetig auf $I_x \times I_y$ (nachrechnen) und es ist $\frac{\partial f}{\partial x}(x, y) = ye^{xy}$ stetig auf $I_x \times I_y$ (nachrechnen), also für $x \neq 0$

$$F'(x) = \int_0^1 ye^{xy} dy = \frac{e^{xy}}{x^2}(xy - 1) \Big|_0^1 = \frac{xe^x - e^x + 1}{x^2}.$$

Übung. Man berechne

$$\lim_{x \rightarrow 0} \frac{xe^x - e^x + 1}{x^2}.$$

12.6 Numerische Integration (Newton-Cotes Formeln; Newton-Cotes Summenformeln; Integrationsverfahren von Romberg)

Die erste Idee zur numerischen Approximation von

$$I(f) = \int_a^b f(x) dx$$

ist es, den Integranden f durch ein Interpolationspolynom p_n zu ersetzen und

$$\int_a^b p_n(x) dx$$

als Näherung zu berechnen.

Im einfachsten Fall approximiert man $\int_a^b f(x) dx$ durch die “Fläche unter der Verbindungsstrecke zwischen $(a, f(a))$ und $(b, f(b))$ ” (vgl. Abbildung 12.6). Dies entspricht der unten diskutierten [Trapezregel](#).

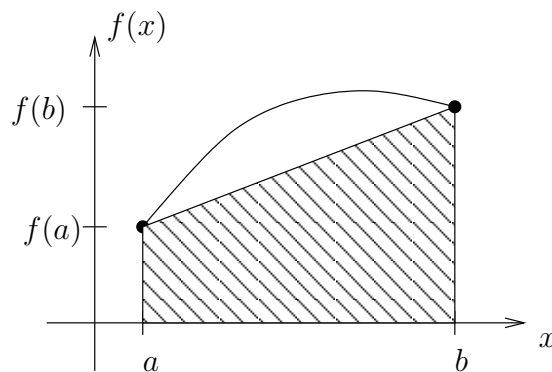


Abbildung 12.6: Zur numerischen Integration.

Es seien nun

$$a \leq x_0 < x_1 < \cdots < x_n \leq b$$

gegeben und p_n bezeichne das (eindeutig bestimmte) Interpolationspolynom vom Grad $\leq n$ mit

$$f(x_i) = p_n(x_i), \quad i = 0, \dots, n.$$

An dieser Stelle sei an die Lagrangesche Darstellung des Interpolationspolynoms (vgl. Kapitel 3.2) erinnert:

$$p_n(x) = \sum_{i=0}^n f(x_i) L_i^{(n)}(x),$$

$$L_i^{(n)}(x) = \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k}.$$

Entsprechend obiger Idee setzt man

$$I_n(f) := \int_a^b p_n(x) \, dx = \sum_{i=0}^n f(x_i) \left[\int_a^b L_i^{(n)}(x) \, dx \right]$$

$$= \sum_{i=0}^n f(x_i) \alpha_i^{(n)}.$$

Dabei heißen die

$$\alpha_i^{(n)} := \int_a^b L_i^{(n)}(x) \, dx$$

Quadraturgewichte.

Sie sind **unabhängig von f** und nur **abhängig von den gewählten Stützstellen**. Wird speziell $f \equiv 1$ gewählt, so folgt

$$\sum_{i=0}^n \alpha_i^{(n)} = b - a.$$

Beispiele. Im Fall **äquidistanter Stützstellen** ergeben sich die **Newton-Cotes Quadraturformeln**:

i) Es sei für $n \in \mathbb{N}$ und für $i = 0, 1, \dots, n$

$$x_i = a + ih, \quad h = \frac{b - a}{n}.$$

Hier ist $x_0 = a$, $x_n = b$, d.h. die Intervallgrenzen sind Teilpunkte und man spricht von den **abgeschlossenen Newton-Cotes Formeln**.

ii) Es sei für $n \in \mathbb{N}_0$ und für $i = 0, 1, \dots, n$

$$x_i = a + (i + 1)h, \quad h = \frac{b - a}{n + 2}.$$

In diesem Fall ist $x_0 > a$ und $x_n < b$, d.h. die Intervallgrenzen sind keine Teilpunkte und man spricht von den [offenen Newton-Cotes Formeln](#).

Die konkrete Berechnung der Quadraturgewichte $\alpha_i^{(n)}$ sei hier am Beispiel der abgeschlossenen Newton-Cotes Formeln vorgestellt. Dabei soll gleichzeitig zu Größen $\sigma_i^{(n)}$, $i = 0, \dots, n$, übergegangen werden, die nicht mehr vom Intervall $[a, b]$, sondern lediglich noch von i und n abhängen.

Mit der Transformation $t: [a, b] \mapsto [0, n]$,

$$t(x) = \frac{n}{b - a}(x - a) = \frac{1}{h}(x - a),$$

folgt

$$\begin{aligned} L_i^{(n)}(x) &= \prod_{k=0, k \neq i}^n \frac{x - x_k}{x_i - x_k} = \frac{a + t(x)h - (a + kh)}{a + ih - (a + kh)} \\ &= \prod_{k=0, k \neq i}^n \frac{t(x) - k}{i - k}. \end{aligned}$$

Für die Quadraturgewichte ergibt sich für $i = 0, \dots, n$ mittels der Substitution $\tau = t(x)$

$$\alpha_i^{(n)} = \int_a^b L_i^{(n)}(x) dx = h \int_0^n \prod_{k=0, k \neq i}^n \frac{\tau - k}{i - k} d\tau,$$

mit anderen Worten

$$\alpha_i^{(n)} = h\sigma_i^{(n)}, \quad \sigma_i^{(n)} := \int_0^n \prod_{k=0, k \neq i}^n \frac{t - k}{i - k} dt.$$

Hierbei ist

$$\sum_{i=0}^n \sigma_i^{(n)} = \frac{1}{h} \sum_{i=0}^n \alpha_i^{(n)} = \frac{1}{h}(b - a) = n$$

und die $\sigma_i^{(n)}$ hängen wie angestrebt nur von i und n ab. Sie werden **einmal berechnet und dann tabelliert**:

Abgeschlossene Newton-Cotes Formeln.

| n | h | $\frac{1}{n}\sigma_i^{(n)}, i = 0, 1, \dots, n$ | Name |
|-----|-----------------|---|----------------------|
| 1 | $b - a$ | $\frac{1}{2} \quad \frac{1}{2}$ | Trapez-Regel |
| 2 | $\frac{b-a}{2}$ | $\frac{1}{6} \quad \frac{4}{6} \quad \frac{1}{6}$ | Simpson-Regel |
| 3 | $\frac{b-a}{3}$ | $\frac{1}{8} \quad \frac{3}{8} \quad \frac{3}{8} \quad \frac{1}{8}$ | $\frac{3}{8}$ -Regel |
| 4 | $\frac{b-a}{4}$ | $\frac{7}{90} \quad \frac{32}{90} \quad \frac{12}{90} \quad \frac{32}{90} \quad \frac{7}{90}$ | Milner-Regel |
| 5 | $\frac{b-a}{5}$ | $\frac{19}{288} \quad \frac{75}{288} \quad \frac{50}{288} \quad \frac{50}{288} \quad \frac{75}{288} \quad \frac{19}{288}$ | |
| 6 | $\frac{b-a}{6}$ | $\frac{41}{840} \quad \frac{216}{840} \quad \frac{27}{840} \quad \frac{272}{840} \quad \frac{27}{840} \quad \frac{216}{840} \quad \frac{41}{840}$ | Weddle-Regel |

$$n = 1 \quad \text{Trapez-Regel} \quad I_1(f) = \frac{h}{2}(f(a) + f(b))$$

$$n = 2 \quad \text{Simpson-Regel} \quad I_2(f) = \frac{h}{3}(f(a) + 4f(\frac{b-a}{2}) + f(b))$$

$$n = 3 \quad \frac{3}{8}\text{-Regel} \quad I_3(f) = \frac{3}{8}h(f_0 + 3f_1 + 3f_2 + f_3)$$

$$n = 4 \quad \text{Milner-Regel} \quad I_4(f) = \frac{2}{45}h(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4)$$

$$n = 5 \quad I_5(f) = \frac{5}{288}h(19f_0 + 75f_1 + 50f_2 + 50f_3 + 75f_4 + 19f_5)$$

$$n = 6 \quad \text{Weddle-Regel} \quad I_6(f) = \frac{1}{140}h(41f_0 + 216f_1 + 27f_2 + 272f_3 + 27f_4 + 216f_5 + 41f_6).$$

Hierbei ist $f_i := f(a + ih)$, $i = 0, \dots, n$.

Analoge Argumente führen auf

Offene Newton-Cotes Formeln.

$$n = 0: J_0(f) = (b - a)f\left(\frac{a+b}{2}\right) \quad (\text{Mittelpunktsformel}),$$

$$h = \frac{b-a}{2},$$

$$n = 1: J_1(f) = \frac{b-a}{2}(f(a+h) + f(b-h)),$$

$$h = \frac{b-a}{3},$$

$$n = 2: J_2(f) = \frac{b-a}{3}(2f(a+h) - f\left(\frac{a+b}{2}\right) + 2f(b-h)),$$

$$h = \frac{b-a}{4},$$

$$n = 3: J_3(f) = \frac{b-a}{24}(11f(a+h) + f(a+2h) + f(b-2h)$$

$$+ 11f(b-h)),$$

$$h = \frac{b-a}{5}.$$

Beispiel. Zu berechnen sei

$$\int_0^1 \frac{x}{1+x^2} dx.$$

Der exakte Wert ist (auf 9 Stellen)

$$\int_0^1 \frac{x}{1+x^2} dx = \frac{1}{2} \ln(1+x^2) \Big|_0^1 = 0.346573590.$$

Die Trapez-Regel liefert in diesem Beispiel

$$I_1(f) = \frac{1}{2}[f(0) + f(1)] = 0.25.$$

Die Simpson-Regel ergibt

$$I_2(f) = \frac{1}{3} \left[f(0) + 4f\left(\frac{1}{2}\right) + f(1) \right] = \frac{21}{60} = 0.35.$$

Nun wurde bereits bei der Polynominterpolation diskutiert, dass es im Allgemeinen nicht günstig ist, mit einer großen Anzahl von Stützstellen zu arbeiten. Es besteht also nicht die Hoffnung, ein Integral durch Newton-Cotes Formeln mit wachsender Anzahl n von Stützstellen immer besser zu approximieren.

Idee. Stattdessen zerlegt man das Intervall $[a, b]$ in N Teilintervalle $[x_{i-1}, x_i]$,

$$x_i = a + i \frac{b-a}{N} =: a + ih, \quad i = 0, 1, \dots, N.$$

Auf diesen Teilintervallen wendet man dann eine Newton-Cotes Formel für kleines n an (vgl. Abbildung 12.6):

$$I_n^\Sigma(f) := \sum_{i=1}^N I_{n|[x_{i-1}, x_i]}(f)$$

ist eine geeignete Näherung, wobei für $i = 1, \dots, n$ die Größe $I_{n|[x_{i-1}, x_i]}$ das Integral über $[x_{i-1}, x_i]$ des Interpolationspolynoms p_n zu $(n+1)$ äquidistanten Stützstellen in $[x_{i-1}, x_i]$ bezeichne.

Es ergeben sich die **Newton-Cotes Summenformeln**.

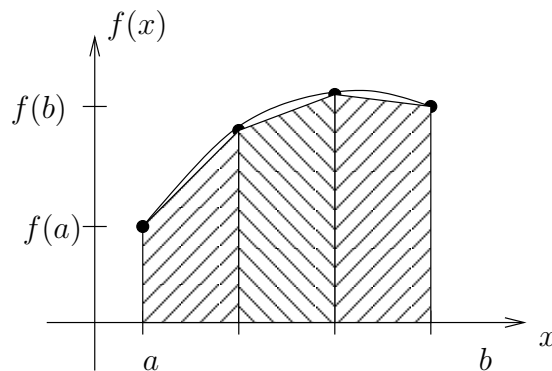


Abbildung 12.7: Zu den Newton-Cotes Summenformeln.

Beispielsweise lautet die **Trapez-Summenregel**: Ist $N \in \mathbb{N}$, $h = (b-a)/N$ und ist $x_i = a + ih$ für $i = 0, 1, \dots, N$, so ist

$$\begin{aligned} T(f, h) &:= I_1^\Sigma(f) = \sum_{i=1}^N \frac{h}{2} [f(x_{i-1}) + f(x_i)] \\ &= h \left[\frac{1}{2} f(x_0) + f(x_1) + f(x_2) + \dots + f(x_{N-1}) + \frac{1}{2} f(x_N) \right] \\ &= h \left[\frac{1}{2} f(a) + f(a+h) + f(a+2h) + \dots \right. \\ &\quad \left. + f(b-h) + \frac{1}{2} f(b) \right]. \end{aligned}$$

Bemerkung. Bei der Trapezsummenregel ist $h = x_i - x_{i-1}$ in Übereinstimmung mit obiger Tabelle.

Beispiel. Eine Näherung von $\int_0^1 \frac{x}{1+x^2} dx$ über die Trapez-Summenregel mit $N = 4$ ist

$$\begin{aligned} T(f, h) &= \frac{1}{4} \left[\frac{1}{2} f(0) + f\left(\frac{1}{4}\right) + f\left(\frac{1}{2}\right) + f\left(\frac{3}{4}\right) + \frac{1}{2} f(1) \right] \\ &= \frac{1}{4} \left[0 + \frac{1}{4} \frac{16}{17} + \frac{1}{2} \frac{4}{5} + \frac{3}{4} \frac{16}{25} + \frac{1}{4} \right] \\ &= 0.341323529 . \end{aligned}$$

Bemerkung. Die Idee des [Integrationsverfahrens von Romberg](#) ist es, $T(f, h_i)$ zu einer absteigenden Folge von Stützweiten $h_i, i = 1, \dots, m$, zu berechnen und anschließend zu diesen Daten eine Extrapolation zum Limes $h \rightarrow 0$ (vgl. Kapitel 11.3) durchzuführen. Dabei kann wieder mit “Polynomen in h^2 ” gearbeitet werden.

Kapitel 13

Der Satz von Taylor, Fehlerabschätzungen

13.1 Taylor-Formel und Taylor-Reihe (Taylor-Polynom; Restglied; Integraldarstellung des Restgliedes; Lagrangesche Rest- gliedformel; die Klasse C^∞ ; reell analytische Funktionen)

Wurde in Kapitel 11 eine Funktion nahe eines fixierten Punktes mit einer (affin) linearen Funktion approximiert, so ist die Approximation einer glatten Funktion mit Polynomen (nicht zu verwechseln mit der Polynominterpolation!) bzw. deren Darstellung als Potenzreihe Gegenstand dieses Paragraphen.

Die entscheidende Frage bei der Approximation bzw. beim Übergang zu einer Potenzreihe ist die nach der **Größe des Fehlers**.

Beobachtung. Betrachtet sei ein Polynom vom Grad n ,

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n .$$

Für $k = 0, 1, \dots, n$ ist die k -te Ableitung

$$p^{(k)}(0) = k!a_k ,$$

mit anderen Worten: Es gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(0) x^k .$$

Zu beliebigem $x_0 \in \mathbb{R}$ kann das Polynom ebenso geschrieben werden in der Form

$$p(x) = \sum_{k=0}^n b_k (x - x_0)^k ,$$

und völlig analog zu oben gilt

$$p(x) = \sum_{k=0}^n \frac{1}{k!} p^{(k)}(x_0) (x - x_0)^k .$$

Können “gutartige” Funktionen in ähnlicher Weise zumindest approximiert werden?

Definition 13.1.1 ¹

Es sei $f \in C^n([a, b])$ und $x_0 \in (a, b)$. Dann heißt

$$T_n(x; x_0) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das Taylor-Polynom n -ten Grades zum Entwicklungspunkt x_0 . Die Taylorsche Formel lautet

$$f(x) = T_n(x; x_0) + R_n(x - x_0) ,$$

wobei das Restglied $R_n(x - x_0)$ im Folgenden zu quantifizieren ist.

Bemerkung. Die Ableitungen des Taylor-Polynoms stimmen an der Stelle x_0 bis zur Ordnung n mit denen von f überein.

Eine Antwort auf die Güte der Approximation gibt

¹B. Taylor, 1685-1731; Cambridge.

Satz 13.1.1

Es sei $f \in C^{n+1}([a, b])$ und es seien $x_0, x = x_0 + h \in (a, b)$.
Dann gilt

$$R_n(h) = \int_0^1 \frac{1}{n!} (1-t)^n f^{(n+1)}(x_0 + th) h^{n+1} dt. \quad (*)$$

Zum Beweis benötigt man zunächst

Lemma 13.1.1

Für eine Funktion $\varphi \in C^{n+1}([0, 1])$ gilt

$$\varphi(1) = \sum_{k=0}^n \frac{1}{k!} \varphi^{(k)}(0) + \int_0^1 \frac{1}{n!} (1-t)^n \varphi^{(n+1)}(t) dt.$$

Beweis des Lemmas. Es wird sukzessive partiell integriert mit dem Ergebnis

$$\begin{aligned} & \int_0^1 \frac{1}{n!} (1-t)^n \varphi^{(n+1)}(t) dt \\ &= \varphi^{(n)}(t) \frac{1}{n!} (1-t)^n \Big|_0^1 + \int_0^1 \frac{1}{(n-1)!} (1-t)^{n-1} \varphi^{(n)}(t) dt \\ &= -\frac{1}{n!} \varphi^{(n)}(0) + \int_0^1 \frac{1}{(n-1)!} (1-t)^{n-1} \varphi^{(n)}(t) dt \\ &= \vdots \\ &= -\left[\frac{1}{n!} \varphi^{(n)}(0) + \dots + \frac{1}{1!} \varphi'(0) \right] + \int_0^1 \varphi'(t) dt. \end{aligned}$$

Aus $\int_0^1 \varphi'(t) dt = \varphi(1) - \varphi(0)$ folgt das Lemma. \square

Beweis von Satz 13.1.1. Für $0 \leq t \leq 1$ setzt man $\varphi(t) := f(x_0 + th)$ und es folgt

$$\varphi'(t) = hf'(x_0 + th), \quad \varphi''(t) = h^2 f''(x_0 + th), \quad \dots,$$

also für $k = 0, 1, \dots, n + 1$

$$\varphi^{(k)}(t) = h^k f^{(k)}(x_0 + th),$$

Lemma 13.1.1 liefert

$$\begin{aligned} f(x_0 + h) &= \varphi(1) \\ &= \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0) h^k + \int_0^1 \frac{1}{n!} (1-t)^n f^{(n+1)}(x_0 + th) h^{n+1} dt. \end{aligned}$$

Dies ist genau die Behauptung des Satzes. □

Es gibt eine Reihe von verschiedenen Darstellungen des Restgliedes, über eine einfache Substitution gelangt man zur [Integraldarstellung des Restgliedes](#) in ihrer üblichen Form, die am häufigsten benutzte Darstellung ist die [Lagrangesche Restgliedformel](#). Sie folgt mit Hilfe des Mittelwertsatzes der Integralrechnung, Satz 12.1.5.

Korollar 13.1.1

Es seien f , x und x_0 wie in Satz 13.1.1.

i) [Integraldarstellung des Restgliedes](#): Es ist

$$R_n(x - x_0) = \int_{x_0}^x \frac{1}{n!} (x - t)^n f^{(n+1)}(t) dt.$$

ii) [Lagrangesche Restgliedformel](#): Es ist für ein $\theta \in (0, 1)$

$$R_n(x - x_0) = \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta h) h^{n+1}.$$

Bemerkung. Eine “gute Approximation” der Funktion durch ein Taylor-Polynom kann in der Regel nur für “kleine h ” erwarten werden, d.h. in der Nähe des Entwicklungspunktes x_0 .

Umgekehrt ausgedrückt: Ist bei festem n der Punkt x sehr weit von x_0 entfernt, so kann das Restglied (der Fehler) sehr groß werden.

Beispiele.

i) Es sei $f(x) = e^x$. Dann ist

$$T_n(x; 0) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} = \sum_{k=0}^n \frac{x^k}{k!},$$

genau wie es die Definition als Potenzreihe erwarten lässt.

Für das Restglied gilt (für ein $\theta \in (0, 1)$)

$$\begin{aligned} |R_n(x)| &= \left| \frac{e^{\theta x}}{(n+1)!} \right| |x|^{n+1} \leq \frac{e^{\theta|x|}}{(n+1)!} |x|^{n+1} \\ &\leq \frac{e^{|x|}}{(n+1)!} |x|^{n+1}. \end{aligned}$$

Insbesondere ist für jedes fixierte $x \in \mathbb{R}$ bewiesen:

$$R_n(x) \xrightarrow{n \rightarrow \infty} 0.$$

ii) Ähnliches gilt für $\sin(x)$, $\cos(x)$ (Übungsaufgabe). Die Taylor-Polynome des Sinus (um $x_0 = 0$) sind in Abbildung 13.1 veranschaulicht.

iii) Es sei $\alpha \in \mathbb{R}$, $\alpha \neq 0$, fixiert und für $x > -1$ sei $f(x) = (1+x)^\alpha$. Dann gilt für alle $k \in \mathbb{N}$

$$\begin{aligned} f'(x) &= \alpha(1+x)^{\alpha-1}, \\ f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2}, \\ &\vdots \\ f^{(k)}(x) &= \alpha(\alpha-1)\cdots(\alpha-k+1)(1+x)^{\alpha-k}. \end{aligned}$$

Wird in Verallgemeinerung der Binomialkoeffizienten (Korollar 2.1.1) für alle $k \in \mathbb{N}$

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}, \quad \binom{\alpha}{0} = 1$$

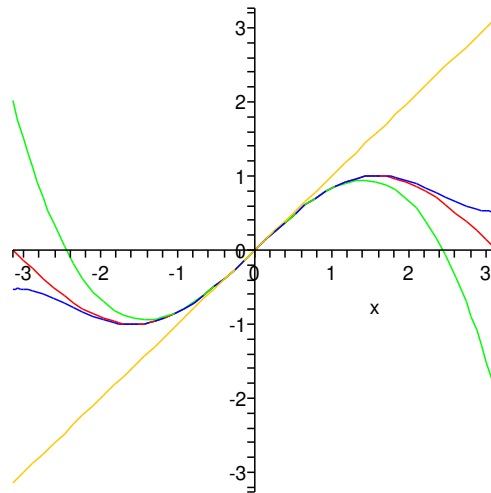


Abbildung 13.1: Taylor-Polynome des Sinus: Das Polynom $T_5(x; 0)$ (blau gekennzeichnet) schmiegt sich schon in einem weiten Bereich um die 0 sehr gut an die tatsächliche Funktion an.

gesetzt, so folgt

$$(1+x)^\alpha = \sum_{k=0}^n \binom{\alpha}{k} x^k + R_n(x).$$

Man vergleiche dies mit dem binomischen Lehrsatz (Satz 2.1.5).

Am Beispiel der Exponentialfunktion wurde bereits angedeutet, dass eine Funktion evtl. im Grenzwert $n \rightarrow \infty$ durch die Taylor-Polynome reproduziert wird. Es stellen sich die Fragen:

Konvergiert $T_n(x; x_0)$ im Grenzwert $n \rightarrow \infty$?
 Wenn ja, ist der Grenzwert gleich $f(x)$?

Ist f wie in Satz 13.1.1, so ist nach Voraussetzung $f^{(n+1)}$ stetig auf dem Kompaktum $[a, b]$, damit insbesondere beschränkt (warum?). D.h. es gibt eine Konstante K_n mit $f^{(n+1)}(x) \leq K_n$ für alle $x \in [a, b]$ und es

gilt

$$|R_n(x - x_0)| \leq \frac{K_n}{(n+1)!} |x - x_0|^{n+1} .$$

Dies bedeutet aber noch nicht zwingend, dass das Restglied für große n bei fixiertem x klein wird (K_n hängt von n ab!). Damit bleibt die Konvergenz von $T_n(x; x_0)$ unklar.

Cauchys Beispiel. (Vgl. Abbildung 13.2) Betrachtet sei die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} e^{-1/x^2} & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

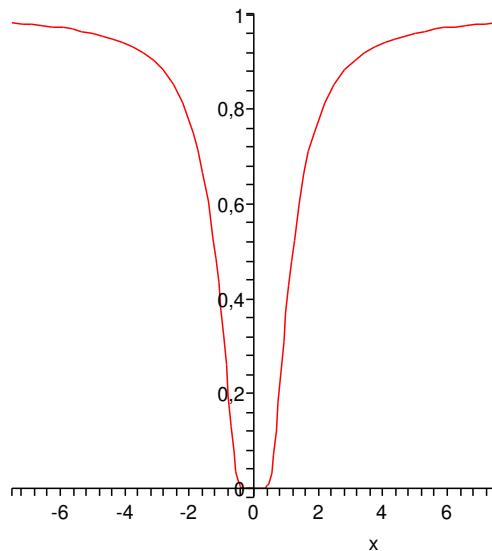


Abbildung 13.2: Die Funktion e^{-1/x^2} läuft so “flach” in den Nullpunkt, dass alle Taylor-Polynome um den Entwicklungspunkt 0 identisch verschwinden.

Es ist

$$\begin{aligned} f'(x) &= 2x^{-3}e^{-1/x^2}, \\ f''(x) &= (4x^{-6} - 6x^{-4})e^{-1/x^2}, \\ &\vdots \quad \quad \quad \vdots \end{aligned}$$

Induktiv sieht man, dass im Nullpunkt alle rechtsseitigen und linksseitigen Ableitungen existieren und gleich Null sind, die Funktion ist

beliebig oft differenzierbar (von der Klasse C^∞) und für alle $k \in \mathbb{N}_0$ gilt

$$f^{(k)}(0) = 0 .$$

Dementsprechend gilt für jedes $n \in \mathbb{N}$

$$T_n(x; 0) = \sum_{k=1}^n \frac{f^{(k)}(0)}{k!} x^k = 0 ,$$

alle Taylor-Polynome um 0 verschwinden identisch, die Funktion hingegen nicht und das Restglied kann für große n nicht klein werden.

Selbst wenn $T_n(x; x_0)$ im Grenzwert $n \rightarrow \infty$ konvergiert, kann nicht geschlossen werden, dass gilt $f(x) = \lim_{n \rightarrow \infty} T_n(x; x_0)$.

Satz 13.1.2

Es sei f in $[a, b]$ beliebig oft differenzierbar, es sei $x_0 \in (a, b)$ und an jeder Stelle $x \in (a, b)$ gelte

$$\lim_{n \rightarrow \infty} R_n(x - x_0) = 0 .$$

Dann ist für alle $x \in (a, b)$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k ,$$

die Funktion heißt auf (a, b) reell analytisch oder von der Klasse C^ω .

Bemerkungen.

- i)* Die Reihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$ heißt **Taylor-Reihe von f** . Nach Satz 13.1.2 konvergiert die Taylor-Reihe von f gegen f , falls das Restglied gegen Null konvergiert.
- ii)* Reell analytische Funktionen können also (lokal) als **Potenzreihen** dargestellt werden. Es sei an die Konvergenzkriterien von Potenzreihen aus Kapitel 5.2 erinnert.

iii) Die Potenzreihendarstellung einer Funktion ist, falls sie existiert, eindeutig bestimmt.

Beispiel. Es sei $-1/2 < x < 1/2$, $x_0 = 0$ und $f(x) = \ln(1+x)$. Es ist

$$\begin{aligned} f'(x) &= \frac{1}{1+x}, \\ f''(x) &= -\frac{1}{(1+x)^2}, \\ &\vdots \\ f^{(k)}(x) &= (-1)^{k-1}(k-1)! \frac{1}{(1+x)^k} \quad \text{für alle } k \in \mathbb{N}. \end{aligned}$$

Die Identität

$$\begin{aligned} \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots + (-1)^{n-1} \frac{1}{n}x^n + \dots \\ &= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{1}{k} x^k \end{aligned}$$

folgt aus

Satz 13.1.3

Es sei $f \in C^\infty([a, b])$ und mit reellen Konstanten M , $r > 0$ gelte für alle $x \in (a, b)$ und für alle $k \in \mathbb{N}_0$

$$|f^{(k)}(x)| \leq k! M r^{-k}.$$

Ist $\delta \in (0, r)$ und ist $x \in (a, b)$ derart, dass $|x - x_0| \leq \delta$ gilt, so folgt

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0) (x - x_0)^k.$$

Beweis. Es ist

$$R_n(h) = \frac{1}{(n+1)!} f^{(n+1)}(x_0 + \theta h) h^{n+1}, \quad h = x - x_0,$$

also für $|x - x_0| < \delta$ nach Voraussetzung

$$|R_n(h)| \leq M \left(\frac{\delta}{r}\right)^{n+1} \xrightarrow{n \rightarrow \infty} 0,$$

aus Satz 13.1.2 folgt die Behauptung. \square

Beispiel. Im obigen Beispiel $f(x) = \ln(1+x)$ ist für $|x| < 1/2$ und für alle $k \in \mathbb{N}$

$$|f^{(k)}(x)| \leq k!(1+x)^{-k} \leq k! \left(\frac{1}{2}\right)^{-k}.$$

Nach Satz 13.1.3 konvergiert die Taylor-Reihe um $x_0 = 0$ gegen f für $|x| < 1/2$.

Bemerkung. In der Tat kann gezeigt werden, dass die Taylor-Reihe von $\ln(1+x)$ um $x_0 = 0$ für $-1 < x \leq 1$ gegen $\ln(1+x)$ konvergiert.

Beispiel. Die Reihendarstellung etwa von $f(x) = 1/(1+x)$ um $x_0 = 1$ kann auf die geometrische Reihe zurückgeführt werden. Es ist für $|x-1| < 2$

$$\begin{aligned} \frac{1}{1+x} &= \frac{1}{2+(x-1)} = \frac{1}{2} \frac{1}{1+\frac{x-1}{2}} \\ &= \frac{1}{2} \frac{1}{1-\left[-\frac{x-1}{2}\right]} \\ &= \frac{1}{2} \sum_{k=0}^{\infty} \frac{(-1)^k}{2^k} (x-1)^k. \end{aligned}$$

Beispiele wie dieses werden noch häufig in der Funktionentheorie auftauchen.

13.2 Konditionierung einer numerischen Aufgabe

(Landausche Symbole; relative Konditionszahlen; Problemfehler)

Eine **numerische Aufgabe** ist die Berechnung von

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

bei einer gegebenen Funktion $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$ aus den **Ausgangsdaten** $x \in U$.

In der Praxis sind dabei in der Regel eine Vielzahl von Auswertungen vorzunehmen.

Einfaches Beispiel. Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $\underline{x} \mapsto y = x_1 + x_2$.

Was versteht man unter der Konditionierung einer numerischen Aufgabe?

Eine numerische Aufgabe heißt **gut konditioniert**, falls “kleine” Störungen der Ausgangsdaten nur “kleine” Störungen des Ergebnisses implizieren.

Beispiel. Die Eingabe in eine numerische Aufgabe sei einmal

$$\underline{\mathbf{b}} = \begin{pmatrix} 3.7041 \\ 1.8519 \end{pmatrix}.$$

Eine zweite Eingabe sei

$$\tilde{\underline{\mathbf{b}}} = \begin{pmatrix} 3.7040 \\ 1.8518 \end{pmatrix}.$$

Die numerische Aufgabe selbst laute: Man löse das Gleichungssystem

$$\begin{pmatrix} 2.4692 & 1.2345 \\ 1.2345 & 0.6172 \end{pmatrix} \underline{\mathbf{Ausgabedaten}} = \underline{\mathbf{Eingabedaten}},$$

wobei Eingabedaten durch obiges $\underline{\mathbf{b}}$, bzw. $\tilde{\underline{\mathbf{b}}}$ gegeben sei.

Die Ausgabedaten berechnen sich zu

$$\underline{\mathbf{y}} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix} 12344 \\ -6170 \end{pmatrix}.$$

Dabei ist zu betonen, dass die Ausgabedaten **exakt berechnet sind und keine numerischen Näherungen darstellen**.

Beobachtung. Störungen in der Größenordnung 10^{-4} in der Eingabe führen zu Störungen in der Größenordnung 10^4 in der Ausgabe. **Auch bei exakter Rechnung machen kleine Störungen in der Eingabe das Ergebnis völlig unsicher.** (Übung: Woran wird das in dem Beispiel wohl liegen?)

Bei der **differenziellen Fehleranalyse** hilft meist

Definition 13.2.1 (*Landausche² Symbole*)

Es sei $U \subset \mathbb{R}^m$ eine Umgebung der Null, $g: U \rightarrow \mathbb{R}^n$ und $h: U \rightarrow \mathbb{R}^l$.

i) Man schreibt

$$g(x) = O(h(x)) \quad \text{für } x \rightarrow 0 ,$$

falls $\varepsilon, C > 0$ existieren mit

$$|g(x)| \leq C|h(x)| \quad \text{für alle } x \in U \text{ mit } |x| < \varepsilon .$$

ii) Man schreibt

$$g(x) = o(h(x)) \quad \text{für } x \rightarrow 0 ,$$

falls es ein $\varepsilon > 0$ sowie eine Funktion $\varphi: U \rightarrow [0, \infty)$ gibt mit $\varphi(x) \rightarrow 0$ für $x \rightarrow 0$ und

$$|g(x)| \leq \varphi(x)|h(x)| \quad \text{für alle } x \in U \text{ mit } |x| < \varepsilon .$$

²E. Landau, 1877-1938; Berlin, Göttingen, Cambridge.

iii) Ist $f: U \rightarrow \mathbb{R}^n$ eine weitere Funktion, so setzt man ($o(h(x))$ analog)

$$f(x) = g(x) + O(h(x)) \quad :\Leftrightarrow \quad f(x) - g(x) = O(h(x)) .$$

Bemerkung. Analog können auch Grenzübergänge $x \rightarrow x_0$ für beliebiges $x_0 \in \mathbb{R}$ betrachtet werden.

Beispiele. Es ist

$$\begin{aligned} \sin(x) &= O(x) , \\ \cos(x) &= O(1) , \\ x^3 &= o(x^2) , \\ \cos^2(x) - 1 &= o(x) . \end{aligned}$$

Im einfachsten Fall, nämlich für $f: [a, b] \rightarrow \mathbb{R}$, sei die numerische Aufgabe der Berechnung von $y = f(x)$ näher analysiert. Dabei sei $f \in C^2([a, b])$ und $x, x + \Delta x \in (a, b)$, wobei Δx eine Störung der Eingabedaten (aufgrund von Rundungsfehlern etc.) bezeichne.

Aus dem Taylorschen Satz (Restglied nach Lagrange) folgt für die korrespondierende Störung Δy der Ausgabedaten

$$\begin{aligned} \Delta y &:= f(x + \Delta x) - f(x) \\ &= f'(x)\Delta x + R_1(\Delta x) , \end{aligned}$$

wobei

$$|R_1(\Delta x)| \leq \frac{1}{2} |\Delta x|^2 \sup_{0 < \theta < 1} f''(x + \theta \Delta x) .$$

Nach Voraussetzung ist f von der Klasse C^2 und für eine Konstante c gilt

$$\sup_{0 < \theta < 1} f''(x + \theta \Delta x) \leq c ,$$

es kann abgeschätzt werden

$$|R_1(\Delta x)| \leq \frac{c}{2} |\Delta x|^2 ,$$

d.h. für $|\Delta x| \rightarrow 0$ gilt

$$\Delta y = f'(x)\Delta x + O(|\Delta x|^2) .$$

Ist allgemein $f: \mathbb{R}^m \subset U \rightarrow \mathbb{R}^n$, so folgt analog für alle $j = 1, \dots, n$

$$\Delta y_j = \sum_{i=1}^m \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}})\Delta x_i + O(|\Delta \underline{\mathbf{x}}|^2) ,$$

wobei an die Definition 12.5.1 der partiellen Ableitungen erinnert sei und wobei

$$f(\underline{\mathbf{x}}) = \begin{pmatrix} f_1(\underline{\mathbf{x}}) \\ \vdots \\ f_n(\underline{\mathbf{x}}) \end{pmatrix} .$$

Schließlich bezeichne für $x_i \neq 0$, $i = 1, \dots, m$, und für $y_j \neq 0$, $j = 1, \dots, n$,

$$\varepsilon_{x_i} := \left| \frac{\Delta x_i}{x_i} \right| \quad (\text{relativer Fehler von } x_i) ,$$

$$\varepsilon_{y_j} := \left| \frac{\Delta y_j}{y_j} \right| \quad (\text{relativer Fehler von } y_j) ,$$

$$K_{ij}(\underline{\mathbf{x}}) := \left| \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}) \frac{x_i}{y_j} \right| \quad (\text{relative Konditionszahlen von } f \text{ in } \underline{\mathbf{x}}) ,$$

und es ist gezeigt

$$\begin{aligned} \varepsilon_{y_j} &\leq \sum_{i=1}^m \left| \frac{\partial f_j}{\partial x_i}(\underline{\mathbf{x}}) \frac{x_i}{y_j} \right| \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta x|^2) \\ &= \sum_{i=1}^m K_{ij}(\underline{\mathbf{x}})\varepsilon_{x_i} + O(|\Delta \underline{\mathbf{x}}|^2) . \end{aligned}$$

Interpretation. Die Zahl $K_{ij}(\underline{\mathbf{x}})$ gibt an, wie stark sich der relative Fehler ε_{x_i} der i^{ten} Komponente der Eingabgröße $\underline{\mathbf{x}}$ auf den relativen Fehler ε_{y_j} der j^{ten} Komponente der Ausgabgröße $\underline{\mathbf{y}}$ auswirkt.

Gibt es Indizes $1 \leq i \leq m$, $1 \leq j \leq n$ mit $K_{ij} \gg 1$, so spricht man von einem **schlecht konditionierten Problem**, andernfalls heißt das Problem **gut konditioniert**.

Der hier besprochene Fehler heißt der **Problemfehler**, der unabhängig vom konkret eingesetzten **Algorithmus** zur Lösung der Aufgabe ist.

Bemerkung. Die Bedingung $K_{ij} \gg 1$ bedeutet, dass K_{ij} wirklich um Größenordnungen größer als 1 ist, sind etwa alle K_{ij} ungefähr 2 (oder 3.165), so heißt das Problem nach wie vor gut konditioniert.

Beispiele.

i) Addition: Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = x_1 + x_2$. Dann gilt

$$K_1(\underline{\mathbf{x}}) = \left| \frac{x_1}{x_1 + x_2} \right|,$$

$$K_2(\underline{\mathbf{x}}) = \left| \frac{x_2}{x_1 + x_2} \right|.$$

Haben also x_1 und x_2 gleiches Vorzeichen, so ist das Problem gut konditioniert.

Gilt dagegen (**Auslöschung**)

$$x_1 \approx -x_2,$$

so ist das Problem schlecht konditioniert: **Kleine relative Eingabefehler können zu sehr großen relativen Fehlern im Ergebnis führen.**

ii) Multiplikation: Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $\underline{\mathbf{x}} \mapsto x_1 x_2$. Dann gilt

$$K_1 = 1 = K_2,$$

die Multiplikation ist **stets gut konditioniert**.

13.3 Stabilität eines Algorithmus (gutartiger Algorithmus)

Der im letzten Paragraphen diskutierte Problemfehler ist systematischer Natur und kann nicht abgewendet werden.

Zur konkreten Berechnung einer numerischen Aufgabe können in der Regel aber vielzählige unterschiedliche Algorithmen eingesetzt werden, die sich durch die Art und die Reihenfolge der auftretenden **Gleitpunktoperationen** unterscheiden. Dabei werden sich unterschiedliche (Rundungs-) Fehler im Ergebnis widerspiegeln.

Wie hängt der Fehler vom speziell eingesetzten Algorithmus ab?

Bezeichnungen.

i) Gegeben sei die numerische Aufgabe der Berechnung von $\underline{\mathbf{y}} = f(\underline{\mathbf{x}})$. Ein **Verfahren** (oder Algorithmus) zur (näherungsweise) Berechnung von $\underline{\mathbf{y}}$ aus $\underline{\mathbf{x}}$ ist eine **endliche** (manchmal auch abzählbar unendliche) Folge von “elementaren” Abbildungen $\varphi^{(k)}$, die nacheinander angewendet auf $\underline{\mathbf{x}}$ einen Näherungswert $\underline{\tilde{\mathbf{y}}}$ für $\underline{\mathbf{y}}$ liefern:

$$\underline{\mathbf{x}} = \underline{\mathbf{x}}^{(0)} \mapsto \varphi^{(1)}(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{x}}^{(1)} \mapsto \dots \mapsto \varphi^{(k)}(\underline{\mathbf{x}}^{(k-1)}) = \underline{\mathbf{x}}^{(k)} = \underline{\tilde{\mathbf{y}}}.$$

Im einfachsten Fall ist $\varphi^{(k)} \in \{+, -, \cdot, /\}$ für alle k .

ii) **Gutartiger Algorithmus** (oder **gut konditionierter Algorithmus**): Der Beitrag des Rundungsfehlers (durch die Maschinenoperationen des Algorithmus) zum **Gesamtfehler** ist höchstens **von der Größenordnung des Problemfehlers**.

iii) Die **numerische Mathematik** befasst sich mit der Suche nach gutartigen Lösungsverfahren.

Beispiel. Zu berechnen sei

$$y = f(\underline{\mathbf{x}}) = x_1^2 - x_2^2 = (x_1 + x_2)(x_1 - x_2).$$

Bevor konkrete Algorithmen auf ihre Gutartigkeit untersucht werden können, ist der Problemfehler abzuschätzen. Es ist

$$\begin{aligned} \varepsilon_y &= \left| \frac{\Delta y}{y} \right| \leq \sum_{i=1}^2 K_i(\underline{\mathbf{x}}) \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta \underline{\mathbf{x}}|^2) \\ &= \sum_{i=1}^2 \left| 2x_i \frac{x_i}{x_1^2 - x_2^2} \right| \left| \frac{\Delta x_i}{x_i} \right| + O(|\Delta \underline{\mathbf{x}}|^2) \\ &\leq 2 \left| \frac{x_1^2 + x_2^2}{x_1^2 - x_2^2} \right| \text{eps} + O(\text{eps}^2) \\ &= 2 \left| \frac{(x_1/x_2)^2 + 1}{(x_1/x_2)^2 - 1} \right| \text{eps} + O(\text{eps}^2). \end{aligned}$$

Dabei wurde (wie es im Folgenden auch getan wird) angenommen, dass relative Eingabefehler durch die Maschinengenauigkeit eps (vgl. Kapitel 4.3) abgeschätzt werden können.

Die Lösung der numerischen Aufgabe kann mit den folgenden beiden Algorithmen angegangen werden:

Algorithmus I.

$$\left. \begin{array}{l} x_1 \mapsto x_1 \odot x_1 \\ x_2 \mapsto x_2 \odot x_2 \end{array} \right\} \mapsto (x_1 \odot x_1) \ominus (x_2 \odot x_2) =: \tilde{y} .$$

Algorithmus II.

$$\left. \begin{array}{l} \underline{x} \mapsto x_1 \ominus x_2 \\ \underline{x} \mapsto x_1 \oplus x_2 \end{array} \right\} \mapsto (x_1 \ominus x_2) \odot (x_1 \oplus x_2) =: \tilde{y} .$$

Zum Vergleich der beiden Algorithmen sei daran erinnert (siehe Kapitel 4.3), dass die exakten Rechenoperationen durch Maschinenoperationen \oplus , \ominus , \odot , \oslash realisiert werden, die fehlerbehaftet sind. Es gilt beispielsweise (alle anderen Operationen analog)

$$x_1 \oplus x_2 = (x_1 + x_2)(1 + \varepsilon) ,$$

wobei ε zwar von x_1 , x_2 abhängt, der Absolutbetrag ist jedoch (für x_1 , x_2 , $x_1 + x_2 \in D$, was im Folgenden stets angenommen wird) durch eps beschränkt.

Zum Algorithmus I. Hier ergibt sich mit $|\varepsilon_1|, |\varepsilon_2|, |\varepsilon_3| < \text{eps}$ (wie oben angedeutet müssen die ε_i unterschieden werden)

$$\begin{aligned} \tilde{y} &= (x_1 \odot x_1) \ominus (x_2 \odot x_2) \\ &= [x_1^2(1 + \varepsilon_1) - x_2^2(1 + \varepsilon_2)](1 + \varepsilon_3) \\ &= \underbrace{x_1^2 - x_2^2}_{=y} + x_1^2\varepsilon_1 - x_2^2\varepsilon_2 + (x_1^2 - x_2^2)\varepsilon_3 + O(\text{eps}^2) . \end{aligned}$$

Der **relative Rundungsfehler durch die eingesetzten Maschinenoperationen** ist abgeschätzt durch

$$\begin{aligned} \left| \frac{\Delta y}{y} \right| &= \left| \frac{y - \tilde{y}}{y} \right| \\ &\leq \frac{x_1^2 + x_2^2 + |x_1^2 - x_2^2|}{|x_1^2 - x_2^2|} \text{eps} + O(\text{eps}^2) \\ &\leq \left[1 + \left| \frac{(x_1/x_2)^2 + 1}{(x_1/x_2)^2 - 1} \right| \right] \text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dabei ist $O(\text{eps}^2)$ sehr klein und kann vernachlässigt werden.

Hinzu kommt ein Term proportional zu eps, solche Terme sind stets gutartig, da der Fehler nie unterhalb der Maschinengenauigkeit abgeschätzt werden kann.

Der verbleibende Term entspricht (bis auf einen konstanten Faktor) genau dem Problemfehler. Insgesamt ist der durch den Algorithmus bedingte Beitrag zum Gesamtfehler von der Größenordnung des Problemfehlers, der Algorithmus ist gutartig.

Zum Algorithmus II. Hier ist mit $|\varepsilon_i| < \text{eps}$, $i = 1, 2, 3$,

$$\begin{aligned} \tilde{y} &= [(x_1 - x_2)(1 + \varepsilon_1) (x_1 + x_2)(1 + \varepsilon_2)](1 + \varepsilon_3) \\ &= y(1 + \varepsilon_1)(1 + \varepsilon_2)(1 + \varepsilon_3) \\ &= y(1 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3) + O(\text{eps}^2) , \end{aligned}$$

d.h.

$$\begin{aligned} \left| \frac{\Delta y}{y} \right| &= \left| \frac{y - \tilde{y}}{y} \right| = |\varepsilon_1 + \varepsilon_2 + \varepsilon_3| + O(\text{eps}^2) \\ &\leq 3\text{eps} + O(\text{eps}^2) . \end{aligned}$$

Dieser relative Fehler ist von der Größenordnung der Maschinengenauigkeit, der Algorithmus ist auf jeden Fall gutartig.

Für $x_1 \approx x_2$ ist der Fehler sogar extrem kleiner als der Problemfehler, der zweite Algorithmus ist deutlich zu bevorzugen.

Faustregeln.

- i)* Numerisch schlecht konditionierte Operationen sollten möglichst früh im Algorithmus ausgeführt werden (vgl. obige Beispiele).
- ii)* Bei der Lösung quadratischer Gleichungen sollten nicht beide Wurzeln aus der Lösungsformel berechnet werden (vgl. Übungen).
- iii)* Die Auswertung von Polynomen sollte mit Hilfe des Horner-Schemas erfolgen.

Zum Abschluss dieses Paragraphen sollen die obigen Betrachtungen anhand eines Zahlenbeispiels verdeutlicht werden.

Beispiel. Betrachtet sei wieder die numerische Aufgabe der Berechnung von $y = f(\mathbf{x}) = x_1^2 - x_2^2$.

Die Mantissenlänge einer konkreten Rechenmaschine sei $t = 3$ (was natürlich keine praxisnahe Annahme ist). Die Eingabedaten seien $x_1 = 0.38049$ und $x_2 = 0.37851$. Die Situation ist in Abbildung 13.3 verdeutlicht.

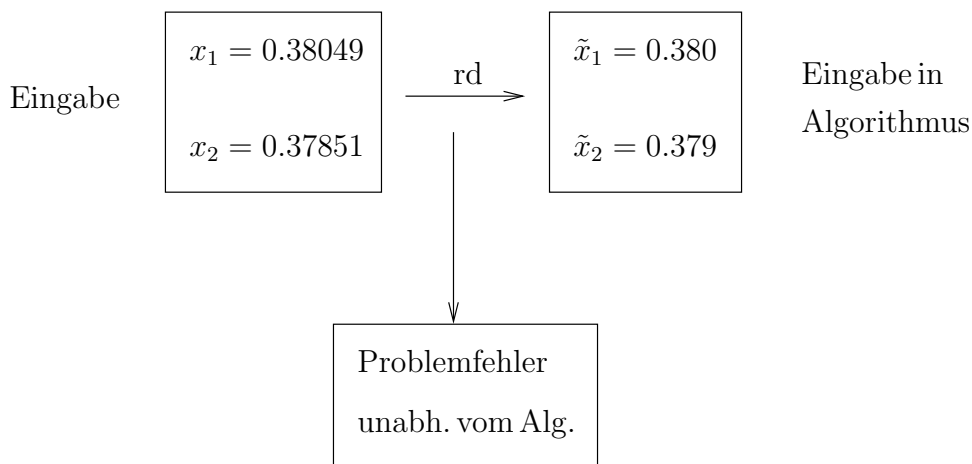


Abbildung 13.3: Zum Fehler bei der Berechnung einer numerischen Aufgabe.

Die **exakten** Eingabedaten x_1, x_2 werden von der Maschine gerundet, die relativen Fehler $\varepsilon_{x_1}, \varepsilon_{x_2} \approx 0.001$ bedingen einen relativen Problemfehler $\varepsilon_y \approx 0.5$. Hier wurde ε_y explizit mittels des Vergleichs von $0.380^2 - 0.379^2$ und $0.38049^2 - 0.37851^2$ berechnet.

Man beachte, dass sich die Konditionszahlen etwa zu 190 berechnen, die theoretisch berechnete obere Schranke für ε_y ist also in guter Übereinstimmung mit der expliziten Rechnung

$$\varepsilon_y \approx 2 \cdot 190 \cdot \underbrace{5 \cdot 10^{-3}}_{\text{eps}} \approx 2 .$$

Die Werte $\tilde{x}_1 = 0.380$, $\tilde{x}_2 = 0.379$ sind die Eingabedaten für den berechnenden Algorithmus.

Algorithmus I. Im ersten Schritt wird berechnet

$$\left. \begin{array}{l} \tilde{x}_1 \\ \tilde{x}_2 \end{array} \right\} \mapsto \left\{ \begin{array}{l} u = \tilde{x}_1 \odot \tilde{x}_1 = \text{rd}(\tilde{x}_1 \tilde{x}_1) = \text{rd}(0.1444) = 0.144 , \\ v = \tilde{x}_2 \odot \tilde{x}_2 = \text{rd}(\tilde{x}_2 \tilde{x}_2) = \text{rd}(0.143641) = 0.144 . \end{array} \right.$$

Dabei werden (wie bereits erwähnt) $\tilde{x}_1 \tilde{x}_1$ und $\tilde{x}_2 \tilde{x}_2$ zunächst intern mit erhöhter Stellenzahl berechnet und dann gerundet.

Es verbleibt die Operation

$$\left. \begin{array}{l} u \\ v \end{array} \right\} \mapsto u - v = 0 .$$

Die fehlerbehafteten Werte u und v dienen hier als Eingabe in den schlecht konditionierten Teil des Algorithmus. Insgesamt ist der durch den Algorithmus bedingte relative Rundungsfehler

$$\left| \frac{0 - (\tilde{x}_1^2 - \tilde{x}_2^2)}{\tilde{x}_1^2 - \tilde{x}_2^2} \right| = 1 .$$

Dieser ist zwar von der gleichen Größenordnung wie der Problemfehler (der Algorithmus ist gutartig), trotzdem ist er sehr groß.

Algorithmus II. Der zweite Algorithmus beginnt mit

$$\left. \begin{array}{l} \tilde{x}_1 \\ \tilde{x}_2 \end{array} \right\} \mapsto \left\{ \begin{array}{l} u = \tilde{x}_1 \ominus \tilde{x}_2 = \text{rd}(\tilde{x}_1 - \tilde{x}_2) = \text{rd}(0.001) = 0.001 , \\ v = \tilde{x}_1 \oplus \tilde{x}_2 = \text{rd}(\tilde{x}_1 + \tilde{x}_2) = \text{rd}(0.759) = 0.759 . \end{array} \right.$$

Der letzte Schritt ist hier

$$\left. \begin{array}{l} u \\ v \end{array} \right\} \mapsto u \odot v = \text{rd}(uv) = \text{rd}(0.759 \cdot 10^{-3}) = 0.759 \cdot 10^{-3} .$$

Im gewählten Zahlenbeispiel liefert der zweite Algorithmus sogar die exakte Lösung der vom Algorithmus zu berechnenden Aufgabe $y = \tilde{x}_1^2 - \tilde{x}_2^2$. Der wesentliche Vorteil des zweiten Algorithmus ist, dass als Eingabe für den schlecht konditionierten Teil die exakte Algorithmeingabe dient.

Bemerkung. Natürlich können auch im Verlauf von Algorithmus II (kleine) Fehler auftreten. Ist etwa

$$\tilde{x}_1 = 0.106 \cdot 10^2, \quad \tilde{x}_2 = 0.612,$$

so ist

$$\tilde{x}_1 \oplus \tilde{x}_2 = \text{rd}(0.11212 \cdot 10^2) = 0.112 \cdot 10^2.$$

13.4 Restglied bei der Polynominterpolation (Interpolationsfehler)

Die bisher angesprochenen numerischen Verfahren beruhen auf der Polynominterpolation. Deshalb soll als exemplarische Fehlerabschätzung für numerische Verfahren das [Restglied bei der Polynominterpolation](#) hergeleitet werden.

Die Form ist eng verwandt mit der Lagrangeschen Restgliedformel, der Beweis der Fehlerabschätzung kann jedoch elementar geführt werden.

In diesem Abschnitt sei $I = [a, b] \subset \mathbb{R}$, $f: I \rightarrow \mathbb{R}$ stetig, $x_0, x_1, \dots, x_n \in I$ seien paarweise verschieden und $p_n(x)$ sei das (eindeutig bestimmte) Interpolationspolynom zu

$$(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_n, f(x_n)),$$

d.h. für $i = 0, 1, \dots, n$ gelte

$$p_n(x_i) = f(x_i).$$

Untersucht wird die Güte der Interpolation, d.h. die Größe $\max_{x \in I} |f(x) - p_n(x)|$.

Satz 13.4.1 (Interpolationsfehler)

Mit obiger Notation sei $f \in C^{n+1}(I)$. Für $x \in I$ sei $I_x \subset I$ das kleinste Intervall, das die Punkte x, x_0, x_1, \dots, x_n enthält.

i) Dann gibt es ein $\eta \in I_x$ mit

$$f(x) - p_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\eta_x) \prod_{j=0}^n (x - x_j) .$$

ii) Für die **maximale Abweichung** folgt

$$\max_{x \in I} |f(x) - p_n(x)| \leq \frac{(b-a)^{n+1}}{(n+1)!} \max_{\eta \in I} |f^{(n+1)}(\eta)| .$$

Beweis. Ist $x \in \{x_0, \dots, x_n\}$, so ist die Aussage trivial.

Es sei nun $x \notin \{x_0, \dots, x_n\}$ und es sei

$$w_n(y) := \prod_{j=0}^n (y - x_j) ,$$

$$F(y) := f(y) - p_n(y) - K w_n(y) ,$$

wobei K die reelle Konstante

$$K = \frac{f(x) - p_n(x)}{w_n(x)} \quad (*)$$

bezeichne (nach Annahme ist $w_n(x) \neq 0$), es gilt also $F(x) = 0$.

Da F an den Stützstellen ebenfalls verschwindet, hat F mindestens $(n+2)$ paarweise verschiedene Nullstellen, und aus dem Satz von Rolle (Satz 11.2.2) folgt

$$F' \text{ hat mindestens } (n+1) \text{ Nullstellen in } I_x .$$

$$\vdots$$

$$F^{(n+1)} \text{ hat mindestens eine Nullstelle } \eta_x \in I_x .$$

Dabei hängt η_x über K von x ab. Es gilt (zur Erinnerung: x ist fixiert und es wird nach y abgeleitet)

$$0 = F^{(n+1)}(\eta_x) = f^{(n+1)}(\eta_x) - \underbrace{p^{(n+1)}(\eta_x)}_{=0} - \underbrace{Kw_n^{(n+1)}(\eta_x)}_{=K(n+1)!},$$

also

$$K = \frac{f^{(n+1)}(\eta_x)}{(n+1)!}.$$

Zusammen mit (*) ist der Satz gezeigt. \square

Beispiel. Für eine Funktion f der Klasse $C^2(I)$ seien die Funktionswerte $f(x_0)$, $f(x_0 + h)$, x_0 , $x_0 + h \in I$ bekannt.

Für $x_0 < x < x_0 + h$ soll $f(x)$ durch das lineare Interpolationspolynom

$$p_1(x) = f(x_0) \frac{x_0 + h - x}{h} + f(x_0 + h) \frac{x - x_0}{h}$$

approximiert werden.

Satz 13.4.1, i), besagt

$$|f(x) - p_1(x)| \leq \frac{1}{2!} \max_{y \in [x_0, x_0+h]} |f''(y)| \max_{y \in [x_0, x_0+h]} (x_0 + h - y)(y - x_0).$$

Eine elementare Kurvendiskussion belegt

$$\max_{y \in [x_0, x_0+h]} (x_0 + h - y)(y - x_0) = \frac{h^2}{4}$$

und der Interpolationsfehler ist abgeschätzt via

$$\max_{y \in [x_0, x_0+h]} |f(y) - p_1(y)| \leq \frac{h^2}{8} \max_{y \in [x_0, x_0+h]} |f''(y)|.$$



Michael Bildhauer

Höhere Mathematik für Ingenieure

– Ein Viersemestriger Grundkurs –

– **Band 2** –

Universität des Saarlandes

F.R. 6.1 Mathematik

Postfach 151150

D-66041 Saarbrücken

Teil IV

Gewöhnliche
Differentialgleichungen (Teil 1)

Kapitel 14

Gewöhnliche lineare Differentialgleichungen

14.1 Einführung (explizite, implizite, lineare, homogene, inhomogene, partielle Differentialgleichung; Ordnung einer Differentialgleichung; System von Differentialgleichungen; Anfangswertproblem)

Eine **gewöhnliche Differentialgleichung** ist eine Gleichung, in welcher eine unabhängige reelle Variable, eine gesuchte Funktion sowie deren Ableitungen auftreten.

Beispiel. Es sei

$$y' + 2xy = 0 . \quad (*)$$

In diesem Fall ist

- $x \in \mathbb{R}$ die unabhängige Variable;
- $y = y(x)$ die gesuchte Funktion;
- y' deren Ableitung erster Ordnung.

Eine auf (einem verallgemeinerten) Intervall I differenzierbare Funktion $y(x)$ heißt **Lösung** von $(*)$ auf dem Intervall I , falls $(*)$ **für alle $x \in I$** gilt. Im obigen Beispiel verifiziert man leicht, dass die Funktion $y(x) = e^{-x^2}$ die Gleichung auf $(-\infty, \infty)$ löst, es ist

$$\frac{d}{dx}e^{-x^2} + 2xe^{-x^2} \equiv 0 .$$

Beispiele aus den Anwendungen.

- i) Man betrachte eine punktförmige Masse m an einer Feder (vgl. Abbildung 14.1). Dabei besagt das **Hookesche Gesetz**, dass die Rückstellkraft der Feder proportional zu deren Auslenkung ist, die Proportionalitätskonstante heißt die Federkonstante c :

$$my''(t) = -cy(t) .$$

Hier symbolisiert t die unabhängige Variable (die Zeit), $y(t)$ ist die Auslenkung der Masse zum Zeitpunkt t (die gesuchte Lösung).

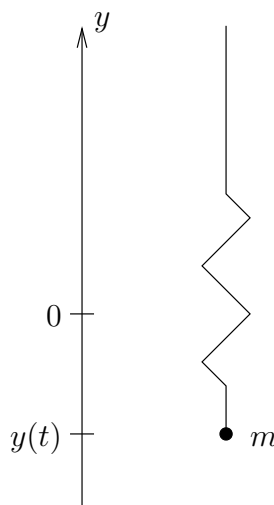


Abbildung 14.1: Zum Hookeschen Gesetz.

Mit Reibungsterm geht die Differentialgleichung über in die **Bewegungsgleichung** ($k \geq 0$)

$$my''(t) + ky'(t) + cy(t) = 0 .$$

Zusätzlich sind **Anfangsbedingungen** in das Problem miteinzubeziehen: Ist t_0 der Anfangszeitpunkt, so hängt das Problem noch ab von

$$\begin{aligned} y(t_0) &= y_0 && \text{mit einer Anfangsauslenkung } y_0 , \\ y'(t_0) &= v_0 && \text{mit einer Anfangsgeschwindigkeit } v_0 . \end{aligned}$$

- ii) Man betrachte einen Wechselstromkreis mit Spule und Ohmschen Widerstand in Reihenschaltung (vgl. Abbildung 14.2).

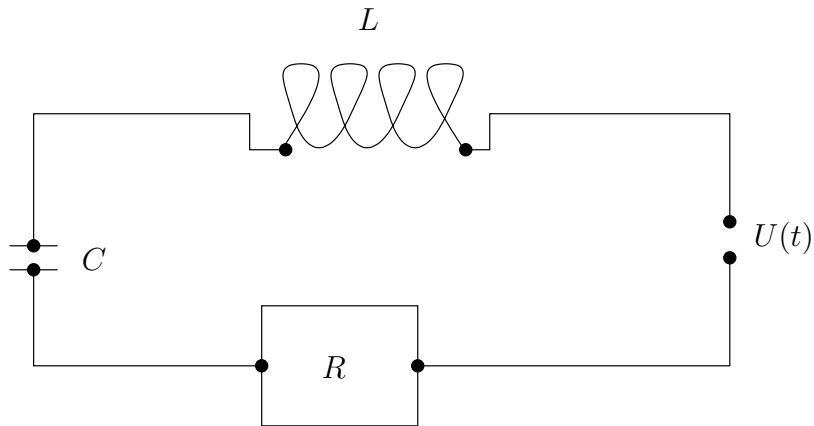


Abbildung 14.2: Ein Wechselstromkreis mit Spule und Ohmschen Widerstand in Reihenschaltung.

Eine feste Spannung U sei vorgegeben und zunächst sei der Schalter geöffnet. Schließt man den Schalter zum Zeitpunkt $t_0 = 0$, so gilt für die Stromstärke I in Abhängigkeit von der Zeit t

$$U = RI(t) + LI'(t)$$

mit der Anfangsbedingung $I(0) = 0$.

iii) Beim radioaktiven Zerfall ist die Änderung der Masse $m(t)$ einer radioaktiven Substanz proportional zu dieser Masse, mit einer Zerfallskonstanten $k > 0$ gilt

$$m'(t) = -km(t) .$$

iv) Es gibt unzählige weitere Beispiele (Abkühlung eines Körpers, Durchbiegung eines Balkens ...), bei denen man jeweils auf eine Differentialgleichung geführt wird, die das Problem in geeigneter Weise beschreibt.

Definition 14.1.1

Es sei $n \in \mathbb{N}$ fixiert, $U \subset \mathbb{R}^{n+2}$ und $F: U \rightarrow \mathbb{R}$. Eine Beziehung der Form

$$F(x, y, y', \dots, y^{(n)}) , \quad (**)$$

in der neben der (unabhängigen) Variablen x und der (gesuchten) Funktion $y(x)$ Ableitungen von y bis zur Ordnung n auftreten, heißt **gewöhnliche Differentialgleichung der Ordnung n** . Eine auf einem (verallgemeinerten) Intervall I n -mal differenzierbare Funktion $y(x)$ heißt **Lösung** der Differentialgleichung (**), falls $(x, y(x), y'(x), \dots, y^{(n)}(x)) \in U$ für alle $x \in I$ und falls (**) für alle $x \in I$ erfüllt ist.

Bemerkungen.

- i) Die höchste auftretende Ableitung legt die Ordnung der Differentialgleichung fest.
- ii) Eine Differentialgleichung der Form (**) heißt **implizite Differentialgleichung**. Eine **explizite Differentialgleichung** hat die spezielle Form

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) .$$

- iii) Eine gewöhnliche **lineare Differentialgleichung** n^{ter} Ordnung ist von der Form

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = r(x) ,$$

wobei die skalaren Funktionen $a_0(x), \dots, a_n(x), r(x)$ auf einem (verallgemeinerten) Intervall I definiert seien. Ist $r(x) \equiv 0$, so heißt die Differentialgleichung **homogen**, andernfalls heißt sie **inhomogen**.

- iv) Bei mehreren Differentialgleichungen für mehrere gesuchte Funktionen spricht man von einem **System von Differentialgleichungen**.

Beispiel. Es sei $n \in \mathbb{N}$ und

$$A(x) = \begin{pmatrix} a_{11}(x) & \dots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \dots & a_{nn}(x) \end{pmatrix} .$$

Dann betrachte man das **System erster Ordnung**

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} + \underline{\mathbf{r}}(x) ,$$

d.h. gesucht ist eine Funktion $\underline{\mathbf{y}}: I \rightarrow \mathbb{R}^n$ mit

$$\begin{aligned} y_1'(x) &= a_{11}(x)y_1(x) + a_{12}(x)y_2(x) + \dots + a_{1n}(x)y_n(x) + r_1(x) \\ &\vdots \\ y_n'(x) &= a_{n1}(x)y_1(x) + a_{n2}(x)y_2(x) + \dots + a_{nn}(x)y_n(x) + r_n(x) . \end{aligned}$$

- v) Bei **partiellen Differentialgleichungen** treten partielle Ableitungen einer Funktion mehrerer Veränderlicher auf. Partielle Differentialgleichungen werden an dieser Stelle aber nicht weiter diskutiert.
- vi) Auf $I = (-\infty, \infty)$ sei folgende Differentialgleichung zweiter Ordnung betrachtet:

$$y'' + y = 0 .$$

Lösungen sind $y_1(x) = \cos(x)$ und $y_2(x) = \sin(x)$. Ebenso löst mit beliebigen Konstanten $c_1, c_2 \in \mathbb{R}$ die Funktion

$$y(x) = c_1 \cos(x) + c_2 \sin(x)$$

die Differentialgleichung. Die Konstanten sind (wie oben bereits angedeutet) dann festgelegt, wenn ein **Anfangswertproblem** betrachtet wird: In der Tat, sind für ein $x_0 \in I$ Werte $y(x_0) = y_0$ und $y'(x_0) = y_0^{(1)}$ vorgegeben, etwa $y(0) = 0$, $y'(0) = 1$, so folgt

$$\begin{aligned} 0 &= y(0) = c_1 \cos(0) + c_2 \sin(0) = c_1 , \\ 1 &= y'(0) = -c_1 \sin(0) + c_2 \cos(0) = c_2 , \end{aligned}$$

man wird auf die **eindeutig bestimmte** Lösung $y(t) = \sin(t)$ des Anfangswertproblems geführt. Das bedeutet tatsächlich im Fall der schwingenden Masse an einer Feder: Zusätzlich zur Bewegungsgleichung muss die Anfangsauslenkung und die Anfangsgeschwindigkeit bekannt sein.

Im Allgemeinen muss eine Differentialgleichung gewissen Bedingungen genügen, damit die Existenz und die Eindeutigkeit einer Lösung des Anfangswertproblems sichergestellt ist. Diese Bedingungen sind im hier diskutierten Fall linearer Differentialgleichungen jedoch stets erfüllt, so dass dieses Problem erst an späterer Stelle angesprochen wird. Zunächst genügt

Satz 14.1.1

In der linearen gewöhnlichen Differentialgleichung n^{ter} Ordnung

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = r(x)$$

mit den Anfangsbedingungen

$$y(x_0) = y_0, \quad y'(x_0) = y_0^{(1)}, \quad \dots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}$$

*seien die Funktionen $a_i(x)$, $r(x)$, $i = 0, \dots, n - 1$, stetig auf dem (verallgemeinerten) Intervall I und es sei $x_0 \in I$. Dann **existiert auf I genau eine Lösung des Anfangswertproblems**, d.h. eine Funktion $y: \mathbb{R} \rightarrow \mathbb{R}$, die auf I der Differentialgleichung genügt sowie den Anfangsbedingungen.*

Beweis. Der Beweis kann “elementar” geführt werden, folgt aber auch aus einem später zu beweisenden Existenz- und Eindeutigkeitsatz für allgemeinere Anfangswertprobleme. \square

Bemerkung. In der Regel wird o.E. der Koeffizient a_n auf eins normiert.

14.2 Homogene lineare Differentialgleichungen

(Fundamentalsystem; allgemeine Lösung; Wronski-Determinante)

In diesem Paragraphen werden gewöhnliche lineare Differentialgleichungen der Form (o.E. $a_n = 1$)

$$L[y] := y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0 \quad (*)$$

diskutiert, wobei die $a_i: [a, b] \rightarrow \mathbb{R}$, $i = 0, \dots, n-1$, stetig auf $[a, b]$ seien. In $(*)$ verschwindet also die rechte Seite $r(x)$ identisch, es handelt sich um eine homogene Differentialgleichung. Nach Satz 14.1.1 hat zwar das Anfangswertproblem eine eindeutige Lösung, es gilt jedoch:

Beobachtung. Sind $y_1(x), \dots, y_k(x)$ Lösungen der homogenen Differentialgleichung $(*)$, so ist auch (vgl. Bemerkung *vi*) aus 14.1)

$$c_1y_1(x) + c_2y_2(x) + \cdots + c_ky_k(x)$$

mit beliebigen Konstanten $c_i \in \mathbb{R}$, $i = 1, \dots, k$, Lösung von $L[y] = 0$ (deshalb heißt die Differentialgleichung **linear**).

Frage: Kann aus der Eindeutigkeit der Lösung des Anfangswertproblems ein Rückschluss auf die Gesamtheit aller Lösungen von $(*)$ gezogen werden?

Beispiel. Es sei wieder exemplarisch die Differentialgleichung

$$y'' + y = 0$$

betrachtet. Zwei bekannte Lösungen sind $y_1(x) = \cos(x)$, $y_2(x) = \sin(x)$. Es sei nun $g(x)$ irgendeine weitere Lösung der Differentialgleichung. Um zu Informationen über g zu gelangen, betrachtet man das Anfangswertproblem mit

$$y(0) = y_0, \quad y'(0) = y_0^{(1)},$$

wobei $y_0, y_0^{(1)} \in \mathbb{R}$ beliebig vorgegeben seien. Für beliebige $c_1, c_2 \in \mathbb{R}$ ist

$$y(x) = c_1 \cos(x) + c_2 \sin(x) + g(x)$$

eine Lösung der Differentialgleichung, das Anfangswertproblem wird gelöst, falls

$$\begin{aligned} y(0) &= c_1 + g(0) = y_0 \quad , \\ y'(0) &= c_2 + g'(0) = y_0^{(1)} \quad . \end{aligned}$$

Damit sind die Konstanten c_1, c_2 bestimmt, nämlich

$$\begin{aligned} c_1 &= y_0 - g(0) \quad , \\ c_2 &= y_0^{(1)} - g'(0) \quad , \end{aligned}$$

die Funktion

$$y(x) = (y_0 - g(0)) \cos(x) + (y_0^{(1)} - g'(0)) \sin(x) + g(x)$$

ist die eindeutig bestimmte Lösung des Anfangswertproblems. Diese eindeutig bestimmte Lösung kann aber auch geschrieben werden als (Probe!)

$$y(x) = y_0 \cos(x) + y_0^{(1)} \sin(x) \quad ,$$

es folgt

$$\begin{aligned} g(x) &= (y_0 - y_0 + g(0)) \cos(x) + (y_0^{(1)} - y_0^{(1)} + g'(0)) \sin(x) \\ &= g(0) \cos(x) + g'(0) \sin(x) \quad . \end{aligned}$$

Zusammenfassend sei festgehalten: Eine **beliebige Lösung** $g(x)$ der Differentialgleichung kann geschrieben werden als **Linearkombination** des Kosinus und des Sinus. In Anlehnung an Kapitel 6.1 sei hier nochmals definiert

Definition 14.2.1

*Die Funktionen y_1, \dots, y_k seien auf einem (verallgemeinerten) Intervall I definiert. Sie heißen **linear unabhängig auf I** , falls aus*

$$c_1 y_1(x) + c_2 y_2(x) + \dots + c_k y_k(x) = 0 \quad \text{für alle } x \in I \quad ,$$

$c_i \in \mathbb{R}, i = 1, \dots, k$, stets folgt

$$c_1 = c_2 = \dots = c_k = 0 \quad .$$

*Sonst heißen die Funktionen **linear abhängig auf I** .*

Bemerkungen.

- i) Der Begriff der linearen Unabhängigkeit war in Definition 6.1.2 auf Vektorräumen, also linearen Räumen eingeführt worden. Formal muss in diesem Abschnitt immer der lineare Raum der Lösungen der Differentialgleichung (*) betrachtet werden.
- ii) Damit hängt zusammen, dass der Begriff der linearen Unabhängigkeit (im Sinne von Definition 14.2.1, d.h. ohne Bezug auf einen Lösungsraum) von dem betrachteten Intervall abhängt (\rightsquigarrow Übungen).

Beispiel. Es sei

$$y_1(x) = \sin(x), \quad y_2(x) = \cos(x), \quad I = (-\infty, \infty).$$

Ist für alle $x \in I$

$$c_1 \sin(x) + c_2 \cos(x) = 0,$$

so ist dies insbesondere für $x = 0$ und für $x = \frac{\pi}{2}$ richtig:

$$\begin{aligned} c_1 \sin(0) + c_2 \cos(0) = 0 &\Rightarrow c_2 = 0, \\ c_1 \sin(\pi/2) + c_2 \cos(\pi/2) = 0 &\Rightarrow c_1 = 0. \end{aligned}$$

Demnach sind $\sin(x)$ und $\cos(x)$ linear unabhängige Funktionen auf $I = (-\infty, \infty)$.

Analog zum obigen Beispiel folgt aus Satz 14.1.1 allgemein

Satz 14.2.1

Die Funktionen $a_i(x)$, $i = 1, \dots, n-1$, seien stetig auf I . Dann besitzt der Differentialgleichung (*), d.h. $L[y] = 0$, **genau n linear unabhängige Lösungen** y_1, y_2, \dots, y_n auf I . **Jede Lösung** von $L[y] = 0$ auf I ist von der Form

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x) \quad \text{für alle } x \in I,$$

wobei die c_i , $i = 1, \dots, n$, reelle Konstanten bezeichnen.

Bemerkungen.

- i) Es heißen n linear unabhängige Lösungen der Differentialgleichung (*) auf I ein **Fundamentalsystem** von Lösungen der Differentialgleichung.
- ii) Die **allgemeine Lösung** $y(x)$ der Differentialgleichung besteht aus allen Linearkombinationen des Fundamentalsystems,

$$y(x) = \sum_{i=1}^n c_i y_i(x), \quad c_i \in \mathbb{R}, \quad i = 1, \dots, n, \quad x \in I.$$

- iii) Die allgemeine Lösung von $y'' + y = 0$ ist demnach

$$y(x) = c_1 \cos(x) + c_2 \sin(x).$$

Im Allgemeinen wird es jedoch nicht immer einfach sein, die lineare Unabhängigkeit von Lösungen zu verifizieren.

Kriterien für die lineare Unabhängigkeit von Lösungen?

Satz 14.2.2¹

Es sind n Lösungen y_1, \dots, y_n der Differentialgleichung () auf I **genau dann** linear unabhängig, wenn (für mindestens) ein $x \in I$ die **Wronski-Determinante***

$$D = \begin{vmatrix} y_1(x) & y_2(x) & \dots & y_n(x) \\ y_1'(x) & y_2'(x) & \dots & y_n'(x) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \dots & y_n^{(n-1)}(x) \end{vmatrix}$$

von Null verschieden ist.

¹J. M. Hoene-Wronski, 1778-1853; Paris.

Bemerkungen.

- i) Ist die Wronski-Determinante für ein $x \in I$ von Null verschieden, so ist sie für alle $x \in I$ von Null verschieden.
- ii) Die Voraussetzung “Lösung von $L[y] = 0$ ” ist wesentlich für den Satz. Sind beispielsweise $\eta_1(x) = x^3$, $\eta_2(x) = |x|^3$ und ist $I = (-1, 1)$, so rechnet man leicht nach, dass die beiden Funktionen linear unabhängig auf I sind, die Wronski-Determinante aber verschwindet (\rightsquigarrow Übungen). Die Funktionen können also nicht beide Lösung ein und derselben homogen linearen Differentialgleichung (hier zweiter Ordnung) sein.

Beispiel. Man betrachte die Lösungen $y_1(x) = \cos(x)$, $y_2(x) = \sin(x)$ von $y'' + y = 0$. Die Wronski-Determinante ist

$$D = \begin{vmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{vmatrix} = 1 \neq 0,$$

wie bereits gezeigt bilden also y_1 und y_2 ein Fundamentalsystem.

Bemerkung. Ist

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + \cdots + c_n y_n(x), \quad c_1, \dots, c_n \in \mathbb{R},$$

die allgemeine Lösung der linearen homogenen Differentialgleichung $L[y] = 0$ und sind

$$y(x_0) = y_0, \quad y'(x_0) = y_0^{(1)}, \quad \dots, \quad y^{(n-1)}(x_0) = y_0^{(n-1)}$$

die Anfangsdaten eines Anfangswertproblems, so erhält man die eindeutige Lösung des Anfangswertproblems durch geeignete Wahl der Konstanten. Dazu ist das folgende lineare Gleichungssystem in den c_i , $i = 1, \dots, n$ zu lösen

$$\begin{aligned} y_0 &= c_1 y_1(x_0) + c_2 y_2(x_0) + \cdots + c_n y_n(x_0) \\ y_0^{(1)} &= c_1 y_1'(x_0) + c_2 y_2'(x_0) + \cdots + c_n y_n'(x_0) \\ &\vdots \\ y_0^{(n-1)} &= c_1 y_1^{(n-1)}(x_0) + c_2 y_2^{(n-1)}(x_0) + \cdots + c_n y_n^{(n-1)}(x_0). \end{aligned}$$

Dieses Gleichungssystem ist eindeutig lösbar, da die Gleichungsdeterminante gerade die von Null verschiedene Wronski-Determinante ist.

14.3 Homogene lineare Differentialgleichungen mit konstanten Koeffizienten (charakteristische Gleichung; Reduktion der Ordnung; charakteristisches Polynom)

In diesem Abschnitt wird der Spezialfall homogener (gewöhnlicher) linearer Differentialgleichungen diskutiert, bei den die Koeffizienten nicht von x abhängen:

$$L[y] = y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0 \quad \text{auf } I, \quad (*)$$

wobei a_0, a_1, \dots, a_{n-1} (wieder ist o.E. $a_n = 1$ angenommen), reelle Konstanten seien. In diesem Spezialfall kann die allgemeine Lösung explizit konstruiert werden.

Beispiele.

i) Es sei $I = (-\infty, \infty)$ und

$$L[y] = y'' + y' - 2y.$$

Zur Konstruktion der allgemeinen Lösung macht man den **Ansatz**:

$$y = e^{\lambda x}, \quad \lambda \in \mathbb{R}.$$

Setzt man dies in die Gleichung ein, so ergibt sich

$$e^{\lambda x}(\lambda^2 + \lambda - 2) = 0,$$

man wird auf die **charakteristische Gleichung**

$$\lambda^2 + \lambda - 2 = 0$$

geführt. Diese hat die beiden Wurzeln

$$\lambda_1 = 1, \quad \lambda_2 = -2$$

und als Fundamentalsystem (lineare Unabhängigkeit nachrechnen!) ergibt sich

$$y_1 = e^x, \quad y_2 = e^{-2x}.$$

Die allgemeine Lösung der Differentialgleichung lautet

$$y(x) = c_1e^x + c_2e^{-2x}, \quad c_1, c_2 \in \mathbb{R}.$$

ii) Es sei wieder $I = (-\infty, \infty)$ und

$$L[y] = y'' + 2y' + y = 0 .$$

Wie oben erhält man die charakteristische Gleichung

$$\lambda^2 + 2\lambda + 1 = 0 ,$$

die nur eine (zweifache) Wurzel $\lambda_1 = -1$ hat. Somit liefert der obige Ansatz lediglich eine Lösung, nämlich

$$y_1(x) = e^{-x} .$$

Um eine weitere, linear unabhängige Lösung zu finden, bedient man sich einer **Reduktion der Ordnung**: Es wird der Ansatz

$$y_2(x) = u(x)e^{-x}$$

gemacht, wobei die Funktion u geeignet zu bestimmen ist. Die Berechnung von $L[y_2]$ liefert

$$\begin{aligned} L[y_2] &= u''(x)e^{-x} - 2u'(x)e^{-x} + u(x)e^{-x} + 2u'(x)e^{-x} \\ &\quad - 2u(x)e^{-x} + u(x)e^{-x} \\ &= u''(x)e^{-x} . \end{aligned}$$

Somit ist $L[y_2] = 0$, falls gilt $u''(x) \equiv 0$, also $u = k_1x + k_2$ mit reellen Konstanten k_1, k_2 . Insbesondere ist $(k_1 = 1, k_2 = 0)$

$$y_2(x) = xe^{-x}$$

eine weitere von y_1 linear unabhängige (nachrechnen!) Lösung der Differentialgleichung, die allgemeine Lösung lautet

$$y(x) = c_1e^{-x} + c_2xe^{-x} , \quad c_1, c_2 \in \mathbb{R} .$$

iii) Auf $I = (-\infty, \infty)$ sei nun

$$L[y] = y'' - 2y' + 5y = 0 ,$$

die charakteristische Gleichung

$$\lambda^2 - 2\lambda + 5 = 0$$

hat in diesem Fall nur die komplexen Wurzeln

$$\lambda_1 = 1 + 2i, \quad \lambda_2 = 1 - 2i.$$

Man betrachtet zunächst **formal** die **komplexe Lösung**

$$y(x) = K_1 e^{(1+2i)x} + K_2 e^{(1-2i)x}, \quad K_1, K_2 \in \mathbb{C}.$$

Schreibt man $K_1 = A_1 + iB_1$, $K_2 = A_2 + iB_2$, so ergibt sich

$$\begin{aligned} y(x) &= (A_1 + iB_1)e^x(\cos(2x) + i\sin(2x)) \\ &\quad + (A_2 + iB_2)e^x(\cos(2x) - i\sin(2x)) \\ &= (A_1 + A_2)e^x \cos(2x) + (B_2 - B_1)e^x \sin(2x) \\ &\quad + i(B_1 + B_2)e^x \cos(2x) + i(A_1 - A_2)e^x \sin(2x). \end{aligned}$$

Nun erinnert man sich daran, dass $y(x)$ **für alle** $K_1, K_2 \in \mathbb{C}$ eine komplexe Lösung war, insbesondere auch im Fall

$$A_1 - A_2 = 0, \quad B_1 + B_2 = 0.$$

In diesem Fall ist y aber in der Tat reellwertig und (setze $c_1 = 2A_1$, $c_2 = -2B_1$) die allgemeine (reelle) Lösung lautet

$$y(x) = c_1 e^x \cos(2x) + c_2 e^x \sin(2x), \quad c_1, c_2 \in \mathbb{R}.$$

Fasst man die in den drei Beispielen diskutierten Aspekte zusammen, so ergibt sich im allgemeinen Fall

Satz 14.3.1

*Es seien a_j , $j = 0, \dots, n-1$, reelle **konstante Koeffizienten** der homogenen Differentialgleichung*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y + a_0y = 0$$

und es sei

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$$

*das zugehörige **charakteristische Polynom**. Dann gilt:*

i) Ist λ_k eine *r-fache reelle Nullstelle* von $p(\lambda)$, so sind die r Funktionen

$$e^{\lambda_k x}, \quad x e^{\lambda_k x}, \quad \dots, \quad x^{r-1} e^{\lambda_k x}$$

Lösungen der homogenen Differentialgleichung.

ii) Sind $\lambda_k = \sigma_k + i\tau_k$ und $\bar{\lambda}_k = \sigma_k - i\tau_k$ ein *Paar von konjugiert komplexen r-fachen Nullstellen* von $p(\lambda)$, so sind die $2r$ Funktionen

$$x^m e^{\sigma_k x} \cos(\tau_k x) \quad \text{und} \quad x^m e^{\sigma_k x} \sin(\tau_k x),$$

$m = 0, \dots, r - 1$, Lösungen der homogenen Differentialgleichung. *Insgesamt* erhält man so n linear unabhängige Lösungen, also ein *Fundamentalsystem von Lösungen*.

Bemerkung. In Kapitel 7.1 wurde bereits festgestellt, dass mit einer Nullstelle λ eines Polynoms mit reellen Koeffizienten auch $\bar{\lambda}$ eine Nullstelle des Polynoms ist.

Beispiele.

i) Man diskutiere den freien Schwingkreis, d.h. mit Konstanten $L > 0$, $C > 0$, $R \geq 0$ die homogene lineare Differentialgleichung (\rightsquigarrow Übungen)

$$LI''(t) + RI'(t) + \frac{1}{C}I(t) = 0.$$

ii) Es sei die Differentialgleichung

$$y''' - y = 0$$

betrachtet. Das charakteristische Polynom lautet

$$p(\lambda) = \lambda^3 - 1,$$

die Nullstellen sind

$$\lambda_1 = 1, \quad \lambda_{2,3} = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}.$$

Ein Fundamentalsystem von Lösungen ist

$$e^x, \quad e^{-\frac{1}{2}x} \cos\left(\frac{\sqrt{3}}{2}x\right), \quad e^{-\frac{1}{2}x} \sin\left(\frac{\sqrt{3}}{2}x\right),$$

die allgemeine Lösung ist

$$y(x) = c_1 e^x + c_2 e^{-\frac{1}{2}x} \cos\left(\frac{\sqrt{3}}{2}x\right) + c_3 e^{-\frac{1}{2}x} \sin\left(\frac{\sqrt{3}}{2}x\right),$$

$c_1, c_2, c_3 \in \mathbb{R}$ (Probe!).

14.4 Inhomogene lineare Differentialgleichungen

(spezielle, allgemeine Lösung der inhomogenen linearen Differentialgleichung; Variation der Konstanten; Resonanz)

Nun sei ein rechte Seite zugelassen, d.h. betrachtet wird die inhomogene (gewöhnliche) lineare Differentialgleichung

$$L[y] := y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0y = r(x). \quad (*)$$

Man beachte, dass die Koeffizienten nicht als konstant vorausgesetzt sind, es gelte aber wieder, dass $a_i : \mathbb{R} \rightarrow \mathbb{R}$ stetig ist auf einem (verallgemeinerten) Intervall I für $i = 1, \dots, n-1$. Gleiches gelte für $r(x)$.

Satz 14.4.1

Ist y_s irgendeine spezielle Lösung der inhomogenen Differentialgleichung () und ist y_h die allgemeine Lösung der homogenen Differentialgleichung $L[y] = 0$, so ist*

$$y(x) = y_h(x) + y_s(x), \quad x \in I,$$

die allgemeine Lösung der inhomogenen Differentialgleichung.

Beweis. Übung: Man betrachte die Differenz zweier spezieller Lösungen. \square

Bemerkungen.

- i) Man beachte die enge Verwandtschaft zu Satz 8.2.1 über lineare Gleichungssysteme.
- ii) Bilden die Lösungen y_1, \dots, y_n der homogenen Differentialgleichung $L[y] = 0$ ein Fundamentalsystem auf I und ist y_s eine spezielle Lösung der inhomogenen Differentialgleichung $L[y] = r$, so lautet die **allgemeine Lösung der inhomogenen Differentialgleichung**

$$y(x) = c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x) + y_s(x), \quad x \in I,$$
 $c_i \in \mathbb{R}$ für $i = 1, \dots, n$.
- iii) Aus der allgemeinen Lösung kann analog zum homogenen Fall die eindeutige Lösung eines Anfangswertproblems berechnet werden.

Beispiel. Auf $I = (-\infty, \infty)$ sei die inhomogene Differentialgleichung

$$y'' + y = x^2$$

betrachtet. Es ist $y_s(x) = x^2 - 2$ eine spezielle Lösung (Probe!), die allgemeine Lösung ist

$$y(x) = c_1 \cos(x) + c_2 \sin(x) + x^2 - 2, \quad c_1, c_2 \in \mathbb{R}.$$

Mittels Nachrechnen verifiziert man leicht, dass in den Beispiel $y_s = x^2 - 2$ eine spezielle Lösung ist. Die entscheidende Frage ist jedoch:

Wie findet man eine spezielle Lösung der inhomogenen linearen Differentialgleichung?

Dazu startet man mit einem Fundamentalsystem y_1, \dots, y_n der **homogenen** Gleichung

$$L[y] = y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = 0.$$

Eine spezielle Lösung der **inhomogenen** Gleichung gewinnt man durch eine sogenannte **Variation der Konstanten**. In der allgemeinen Lösung der homogenen Gleichung

$$c_1 y_1(x) + c_2 y_2(x) + \cdots + c_n y_n(x)$$

werden die Konstanten durch Funktionen $c_i(x)$, $i = 1, \dots, n$, ersetzt, d.h. man macht den **Ansatz**:

$$y_s(x) = c_1(x)y_1(x) + c_2(x)y_2(x) + \cdots + c_n(x)y_n(x) .$$

Beispiel. Wie oben sei auf $I = (-\infty, \infty)$ die inhomogene Gleichung

$$y'' + y = x^2$$

betrachtet. Die allgemeine Lösung der homogenen Gleichung lautet

$$c_1 y_1(x) + c_2 y_2(x) = c_1 \cos(x) + c_2 \sin(x) , \quad c_1, c_2 \in \mathbb{R} .$$

Mit dem Ansatz

$$y_s(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

sind zunächst die Ableitungen von y_s zu berechnen. Es ist

$$y'_s(x) = c'_1(x)y_1(x) + c_1(x)y'_1(x) + c'_2(x)y_2(x) + c_2(x)y'_2(x) .$$

Nun wählt man die gesuchten Funktionen $c_1(x)$, $c_2(x)$ im ersten Schritt derart, dass für alle $x \in I$

$$c'_1(x)y_1(x) + c'_2(x)y_2(x) = 0 . \tag{1}$$

Mit dieser Wahl ist

$$y'_s(x) = c_1(x)y'_1(x) + c_2(x)y'_2(x)$$

und die zweite Ableitung berechnet sich zu

$$y''_s(x) = c'_1(x)y'_1(x) + c_1(x)y''_1(x) + c'_2(x)y'_2(x) + c_2(x)y''_2(x) .$$

Zusätzlich zu (1) wählt man nun die Funktionen $c_1(x)$, $c_2(x)$ so, dass

$$c'_1(x)y'_1(x) + c'_2(x)y'_2(x) = x^2 . \tag{2}$$

Das impliziert

$$y_s''(x) = c_1(x)y_1''(x) + c_2(x)y_2''(x) + x^2$$

und somit

$$y_s''(x) + y_s(x) = c_1(x)y_1''(x) + c_2(x)y_2''(x) + x^2 + c_1(x)y_1(x) + c_2(x)y_2(x) .$$

An dieser Stelle kann ausgenutzt werden, dass y_1 und y_2 Lösungen der homogenen Gleichung sind, d.h.

$$\begin{aligned} c_1(x)y_1''(x) + c_1(x)y_1(x) &= c_1(x)(y_1''(x) + y_1(x)) = 0 , \\ c_2(x)y_2''(x) + c_2(x)y_2(x) &= c_2(x)(y_2''(x) + y_2(x)) = 0 . \end{aligned}$$

Tatsächlich gilt also

$$y_s'' + y_s = x^2 ,$$

vorausgesetzt, dass die Bedingungen (1) und (2) erfüllt sind. Diese lauten im konkreten Beispiel

$$\begin{aligned} c_1'(x) \cos(x) + c_2'(x) \sin(x) &= 0 , \\ -c_1'(x) \sin(x) + c_2'(x) \cos(x) &= x^2 . \end{aligned}$$

Als Lösung findet man

$$\begin{aligned} c_1'(x) &= -x^2 \sin(x) , \\ c_2'(x) &= x^2 \cos(x) . \end{aligned}$$

Schließlich muss noch aufintegriert werden und man erhält (die Integrationskonstanten werden gleich Null gewählt)

$$\begin{aligned} c_1(x) &= x^2 \cos(x) - 2x \sin(x) - 2 \cos(x) , \\ c_2(x) &= x^2 \sin(x) + 2x \cos(x) - 2 \sin(x) . \end{aligned}$$

Eine spezielle Lösung

$$\begin{aligned} y_s(x) &= c_1(x) \cos(x) + c_2(x) \sin(x) \\ &= x^2 \cos^2(x) - 2x \sin(x) \cos(x) - 2 \cos^2(x) \\ &\quad + x^2 \sin^2(x) + 2x \cos(x) \sin(x) - 2 \sin^2(x) \\ &= x^2 - 2 \end{aligned}$$

ist gefunden, von der oben schon kontrolliert war, dass sie tatsächlich die inhomogene Differentialgleichung löst.

Wie im Beispiel führt im allgemeinen Fall obiger Ansatz auf das Gleichungssystem

$$\left. \begin{aligned} y_1(x)c_1'(x) + y_2(x)c_2'(x) + \cdots + y_n(x)c_n'(x) &= 0 \\ y_1'(x)c_1(x) + y_2'(x)c_2(x) + \cdots + y_n'(x)c_n(x) &= 0 \\ &\vdots \\ y_1^{(n-1)}(x)c_1'(x) + \cdots + y_n^{(n-1)}(x)c_n'(x) &= r(x) . \end{aligned} \right\} \quad (3)$$

Die y_i und deren Ableitungen sind bekannt, d.h. (3) ist ein inhomogenes lineares Gleichungssystem für die c_i' , $i = 1, \dots, n$. Die Gleichungsdeterminante ist die Wronski-Determinante, also von Null verschieden, und man erhält eine eindeutige Lösung $c_1'(x), \dots, c_n'(x)$. Durch Integration lassen sich schließlich die gesuchten Funktionen $c_i(x)$, $i = 1, \dots, n$, ermitteln.

Zusammenfassend kann gesagt werden: Prinzipiell können über die Methode der Variation der Konstanten immer spezielle Lösungen gefunden werden, die Rechnungen sind aber oft sehr aufwendig.

Gibt es in bestimmten Situationen einfachere Möglichkeiten, eine spezielle Lösung der inhomogenen Gleichung zu finden?

Dazu sei nun eine inhomogene Differentialgleichung mit **konstanten Koeffizienten**

$$L[y] = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = r(x) , \quad (4)$$

$a_i \in \mathbb{R}$, $i = 0, 1, \dots, n-1$, betrachtet, wobei angenommen sei, dass die rechte Seite $r(x)$ eine ganz besondere Struktur habe, nämlich

$$\left. \begin{aligned} r(x) &= (c_0 + c_1x + \cdots + c_kx^k)e^{\sigma x} \cos(\tau x) \\ &+ (d_0 + d_1x + \cdots + d_kx^k)e^{\sigma x} \sin(\tau x) , \end{aligned} \right\} \quad (5)$$

$c_i, d_i, \sigma, \tau \in \mathbb{R}$, $i = 0, 1, \dots, k$.

Beispiele.

$$r(x) = c_0 + c_1x + \cdots + c_kx^k ;$$

$$r(x) = c_0e^{\sigma x} ;$$

$$r(x) = c_0 \cos(\tau x) + d_0 \sin(\tau x) .$$

Liegt eine solche Struktur vor, so ist es naheliegend, einen der Struktur angepassten Ansatz zu machen. Dabei sind zwei Fälle zu unterscheiden, wie das Beispiel des Schwingkreises verdeutlicht:

Beispiel Schwingkreis. Mit einer Induktivität $L > 0$, einer Kapazität $C > 0$, einem Ohmschen Widerstand $R \geq 0$ und einer Wechselspannung $U(t)$, $U'(t) = a \cos(\omega t) + b \sin(\omega t)$, $\omega \in \mathbb{R}$, $\omega \neq 0$, werde der Schwingkreis aus Abbildung 14.3 betrachtet.

Die gesuchte Stromstärke erfüllt die inhomogene lineare Differentialgleichung mit konstanten Koeffizienten

$$LI''(t) + RI'(t) + \frac{1}{C}I(t) = a \cos(\omega t) + b \sin(\omega t) .$$

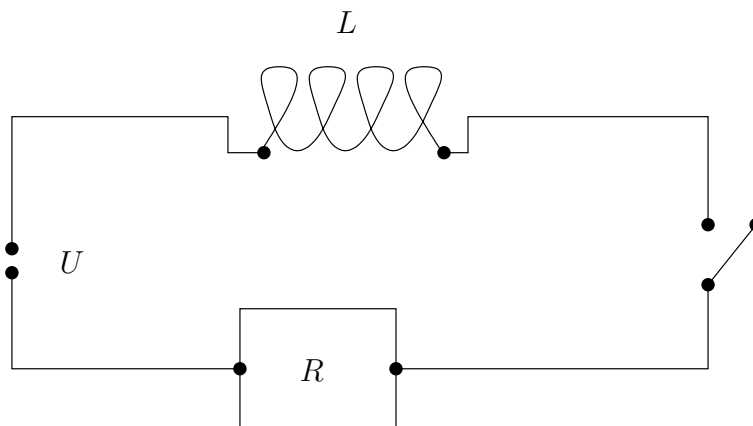


Abbildung 14.3: Ein Schwingkreis.

Der freie Schwingkreis (ohne angelegte Spannung) schwingt mit einer Frequenz ω_0 , die Schwingung wird durch den Ohmschen Widerstand (im Fall $R > 0$) im Laufe der Zeit abklingen.

Wird der Schwingkreis wie oben zusätzlich durch eine angelegte Spannung angeregt (Frequenz ω), so ist zu unterscheiden, ob die anregende Frequenz gleich ω_0 ist oder nicht: Stimmen die Frequenzen nicht

überein, so wird sich im Laufe der Zeit das System genau wie die anregende Spannung verhalten (die freie Schwingung ebbt ab), der Ansatz zur Lösung der Gleichung sollte die gleiche Struktur wie die rechte Seite haben.

Stimmen die beiden Frequenzen jedoch überein, so wird sich das System (ohne Ohmschen Widerstand) “aufschaukeln”, es liegt der **Resonanzfall** vor. Die anwachsende Amplitude der Schwingung spiegelt sich in einer zusätzlichen x -Potenz im Ansatz wider.

Formal sieht das wie folgt aus:

Fall I. In (5) ist $\sigma + i\tau$ **nicht Wurzel** des zu $L[y] = 0$ gehörigen charakteristischen Polynoms. Der Ansatz in diesem Fall lautet:

$$y_s(x) = (K_0 + K_1x + \cdots + K_kx^k)e^{\sigma x} \cos(\tau x) \\ + (L_0 + L_1x + \cdots + L_kx^k)e^{\sigma x} \sin(\tau x),$$

$K_i, L_i \in \mathbb{R}, i = 0, 1, \dots, k$.

Fall II. In (5) ist $\sigma + i\tau$ eine Wurzel der Vielfachheit m des zu $L[y] = 0$ gehörigen charakteristischen Polynoms. Der Ansatz im zweiten Fall lautet

$$y_s(x) = x^m(K_0 + K_1x + \cdots + K_kx^k)e^{\sigma x} \cos(\tau x) \\ + x^m(L_0 + L_1x + \cdots + L_kx^k)e^{\sigma x} \sin(\tau x),$$

$K_i, L_i \in \mathbb{R}, i = 0, 1, \dots, k$.

In beiden Fällen liefert das Einsetzen in die Differentialgleichung zusammen mit einem Koeffizientenvergleich die gesuchten $K_i, L_i, i = 0, 1, \dots, k$.

Beispiel Schwingkreis. Die Nullstellen des charakteristischen Polynoms erfüllen die charakteristische Gleichung

$$\lambda^2 L + \lambda R + \frac{1}{C} = 0,$$

die Lösungen sind

$$\lambda_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}.$$

Fall A. Es sei $R > 0$. Mit obiger Notation ist $\sigma + i\tau = i\omega$ rein imaginär, im Fall $R > 0$ ist also stets der erste Ansatz zu machen:

$$I_s(t) = K_0 \cos(\omega t) + L_0 \sin(\omega t) .$$

Die Berechnung von $I'_s(t)$, $I''_s(t)$, das Einsetzen in die Differentialgleichung und ein Koeffizientenvergleich ergibt die gesuchten Konstanten K_0 und L_0 (Übung). Ist $I_h(t)$ die allgemeine Lösung der homogenen Differentialgleichung, so ergibt sich als allgemeine Lösung der inhomogenen Differentialgleichung

$$I(t) = I_h(t) + K_0 \cos(\omega t) + L_0 \sin(\omega t) ,$$

$I_h(t)$ ist die **freie (abklingende) Schwingung**,

$$K_0 \cos(\omega t) + L_0 \sin(\omega t)$$

ist die **Dauerschwingung**, die durch die Anregung erzeugt wird.

Fall B. Im zweiten Fall ist $R = 0$, die Wurzeln des charakteristischen Polynoms sind

$$\lambda_{1,2} = \pm i\omega_0 \quad \text{mit} \quad \omega_0 = \frac{1}{\sqrt{LC}} .$$

Wie bereits erwähnt ist $\sigma + i\tau = i\omega$ und es ist eine erneute Fallunterscheidung notwendig.

Fall B1. Im Fall $\omega \neq \omega_0$ führt der gleiche Ansatz wie oben zum Ziel.

Fall B2. Es sei $\omega = \omega_0$, physikalisch interpretiert bedeutet das wie bereits oben gesagt, dass die Anregungsfrequenz des Schwingkreises gleich der Eigenfrequenz des freien Systems ist, man ist im Resonanzfall.

Der Ansatz sieht nun wie folgt aus:

$$I_s(t) = t[K_0 \cos(\omega t) + L_0 \sin(\omega t)] ,$$

und man berechnet

$$I'_s(t) = K_0 \cos(\omega t) + L_0 \sin(\omega t) - tK_0\omega \sin(\omega t) + tL_0\omega \cos(\omega t)$$

sowie

$$\begin{aligned}
 I_s''(t) &= -K_0\omega \sin(\omega t) + L_0\omega \cos(\omega t) \\
 &\quad -K_0\omega \sin(\omega t) - tK_0\omega^2 \cos(\omega t) \\
 &\quad +L_0\omega \cos(\omega t) - tL_0\omega^2 \sin(\omega t) \\
 &= -2K_0\omega \sin(\omega t) + 2L_0\omega \cos(\omega t) \\
 &\quad -t[K_0\omega^2 \cos(\omega t) + L_0\omega^2 \sin(\omega t)] .
 \end{aligned}$$

In die Differentialgleichung

$$LI'' + \frac{1}{C}I = a \cos(\omega t) + b \sin(\omega t)$$

eingesetzt, ergibt das

$$\begin{aligned}
 &-2LK_0\omega \sin(\omega t) + 2LL_0\omega \sin(\omega t) \\
 &\quad -tL[K_0\omega^2 \cos(\omega t) + L_0\omega^2 \sin(\omega t)] \\
 &\quad +\frac{1}{C}t[K_0 \cos(\omega t) + L_0 \sin(\omega t)] \\
 &= a \cos(\omega t) + b \sin(\omega t) .
 \end{aligned}$$

Hierbei ist insbesondere zu beachten, dass die Gleichheit nur gelten kann, wenn sich die beiden Terme der Form $t[\dots]$ gegenseitig aufheben. In der Tat gilt genau im hier betrachteten Resonanzfall $\omega = \omega_0 = 1/\sqrt{LC}$

$$-tLK_0\omega^2 + t\frac{K_0}{C} = 0$$

wegen

$$-L\omega^2 + \frac{1}{C} = -L\frac{1}{LC} + \frac{1}{C} = 0 .$$

Ebenso erkennt man

$$-tLL_0\omega^2 + t\frac{L_0}{C} = 0 ,$$

der obige Ansatz ist also genau auf den Resonanzfall zugeschnitten.

Für die Konstanten K_0 , L_0 erhält man mittels eines Koeffizientenvergleiches

$$K_0 = -\frac{b}{2L\omega_0} = -\frac{b}{2}\sqrt{\frac{C}{L}},$$
$$L_0 = \frac{a}{2L\omega_0} = \frac{a}{2}\sqrt{\frac{C}{L}}.$$

Mit dieser Wahl ist y_s eine spezielle Lösung der inhomogenen Gleichung, die allgemeine Lösung ist mit weiteren Konstanten $M_1, M_2 \in \mathbb{R}$

$$I(t) = (M_1 + tK_0) \cos(\omega_0 t) + (M_2 + tL_0) \sin(\omega_0 t),$$

die Amplitude der Lösung wird mit fortschreitender Zeit immer größer, es entwickelt sich eine [Resonanzkatastrophe](#).

Ausblick. Der Ansatz $y = e^{\lambda x}$ hat im Fall einer Differentialgleichung mit konstanten Koeffizienten zum Ziel, nämlich der allgemeinen Lösung der Differentialgleichung, geführt. Bevor Systeme von Differentialgleichungen mit konstanten Koeffizienten untersucht werden können, müssen zunächst weitere Bausteine aus der linearen Algebra zur Verfügung gestellt werden.

Teil V

Lineare Algebra (Teil 2)

Kapitel 15

Spektraltheorie quadratischer Matrizen

15.1 Eigenwerte und Eigenvektoren (Eigenraum; charakteristisches Polynom einer Matrix; algebraische und geometrische Vielfachheit eines Eigenwertes; ähnliche Matrizen)

Es sei $A \in M(n, n, \mathbb{R})$ eine quadratische $n \times n$ Matrix. Will man das lineare Differentialgleichungssystem mit konstanten Koeffizienten erster Ordnung

$$\underline{y}' = A\underline{y} ,$$

d.h. die n Gleichungen in den n gesuchten Funktionen y_1, \dots, y_n ,

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ &\vdots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n , \end{aligned}$$

studieren, so liegt es nahe, ähnlich wie in Kapitel 14.3 den Ansatz zu machen

$$\underline{y}(x) = \underline{v}e^{\lambda x} , \quad \underline{v} \in \mathbb{R}^n , \quad \lambda \in \mathbb{R} .$$

Bemerkung. Hier ist \underline{y} eine Funktion $I \rightarrow \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein (verallgemeinertes) Intervall bezeichnet.

Eingesetzt in die Differentialgleichung ergibt der Ansatz

$$\underline{v}\lambda e^{\lambda x} = A\underline{v}e^{\lambda x} ,$$

d.h. $\underline{\mathbf{v}}$ muss der Beziehung

$$A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}$$

genügen. Auch die Umkehrung ist richtig: Gilt diese Beziehung, so ist $\underline{\mathbf{y}}(x) = \underline{\mathbf{v}}e^{\lambda x}$ eine Lösung des obigen Differentialgleichungssystems.

Definition 15.1.1

Es heißt $\lambda \in \mathbb{C}$ *Eigenwert* der Matrix $A \in M(n, n, \mathbb{C})$, falls $\underline{\mathbf{v}} \in \mathbb{C}^n$, $\underline{\mathbf{v}} \neq \underline{\mathbf{0}}$, existiert mit

$$A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}.$$

Der Vektor $\underline{\mathbf{v}}$ heißt in diesem Fall ein *Eigenvektor* von A zum Eigenwert λ .

Bemerkungen.

i) Im Allgemeinen kann selbst für reelle Matrizen nicht erwartet werden, dass reelle Eigenwerte existieren. Ist etwa

$$A = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \in M(n, n, \mathbb{R}), \quad \varphi \in \mathbb{R},$$

so kann kein reeller Eigenwert existieren (warum?). Es ist aber

$$\begin{aligned} A \begin{pmatrix} 1 \\ -i \end{pmatrix} &= \begin{pmatrix} \cos(\varphi) + i \sin(\varphi) \\ \sin(\varphi) - i \cos(\varphi) \end{pmatrix} \\ &= (\cos(\varphi) + i \sin(\varphi)) \begin{pmatrix} 1 \\ -i \end{pmatrix} = e^{i\varphi} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \end{aligned}$$

Der Vektor $\begin{pmatrix} 1 \\ -i \end{pmatrix}$ ist also ein Eigenvektor zum Eigenwert $\lambda =$

$e^{i\varphi}$. Analog sieht man, dass der Vektor $\begin{pmatrix} 1 \\ i \end{pmatrix}$ ein Eigenvektor zum Eigenwert $e^{-i\varphi}$ ist.

ii) Mit einem Vektor $\underline{\mathbf{v}}$ ist auch jedes Vielfache von $\underline{\mathbf{v}}$ ein Eigenvektor zum Eigenwert λ , die Menge

$$E_\lambda := \{\underline{\mathbf{v}} \in \mathbb{C}^n : A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}}\}$$

ist ein Unterraum des \mathbb{C}^n , genannt der **Eigenraum** zum Eigenwert λ .

Als evidente (aber wichtige) Eigenschaft sei zunächst festgehalten:

Satz 15.1.1

Sind $\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(k)} \in \mathbb{C}^n$ Eigenvektoren zu *paarweise verschiedenen* Eigenwerten $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ einer Matrix $A \in M(n, n, \mathbb{C})$, so sind sie *linear unabhängig*.

Beweis. Der Beweis sei hier nur im Fall $k = 2$ geführt (der allgemeine Fall ist mit Induktion zu zeigen). Es sei also $\underline{\mathbf{v}}^{(1)}$ Eigenvektor von A zum Eigenwert λ_1 und es sei $\underline{\mathbf{v}}^{(2)}$ Eigenvektor zum Eigenwert $\lambda_2 \neq \lambda_1$:

$$A\underline{\mathbf{v}}^{(1)} = \lambda_1\underline{\mathbf{v}}^{(1)}, \quad A\underline{\mathbf{v}}^{(2)} = \lambda_2\underline{\mathbf{v}}^{(2)}.$$

Wären $\underline{\mathbf{v}}^{(1)}$ und $\underline{\mathbf{v}}^{(2)}$ linear abhängig, so gäbe es ein $\alpha \in \mathbb{C}$ mit $\underline{\mathbf{v}}^{(2)} = \alpha\underline{\mathbf{v}}^{(1)}$. Dies würde aber bedeuten

$$A\underline{\mathbf{v}}^{(2)} = A\alpha\underline{\mathbf{v}}^{(1)} = \alpha A\underline{\mathbf{v}}^{(1)} = \alpha\lambda_1\underline{\mathbf{v}}^{(1)} = \lambda_1\alpha\underline{\mathbf{v}}^{(1)} = \lambda_1\underline{\mathbf{v}}^{(2)}.$$

Aus der Voraussetzung $\lambda_1 \neq \lambda_2$ folgt ein Widerspruch. □

Nun existiert ein Eigenvektor $\underline{\mathbf{v}} \in \mathbb{C}^n$ zu einem Eigenwert $\lambda \in \mathbb{C}$ nach Definition genau dann, wenn

$$A\underline{\mathbf{v}} = \lambda\underline{\mathbf{v}},$$

also genau dann, wenn ($I_n \in M(n, n, \mathbb{C})$ Einheitsmatrix)

$$(A - \lambda I_n)\underline{\mathbf{v}} = \underline{\mathbf{0}},$$

also genau dann, wenn

$$\underline{\mathbf{v}} \in \text{kern}(A - \lambda I_n),$$

also genau dann, wenn (vgl. Korollar 8.6.1) (zur Erinnerung: Der Nullvektor ist nach Definition kein Eigenvektor)

$$\det(A - \lambda I_n) = 0.$$

Satz 15.1.2

Es ist $\lambda \in \mathbb{C}$ genau dann ein Eigenwert einer Matrix $A \in M(n, n, \mathbb{C})$, wenn $\det(A - \lambda I_n) = 0$ gilt. Dabei heißt

$$\chi_A(\lambda) := \det(A - \lambda I_n)$$

das *charakteristische Polynom* der Matrix A .

Beispiel. Es sei

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in M(n, n, \mathbb{R}).$$

Das charakteristische Polynom ist

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda I_n) = \begin{vmatrix} 2 - \lambda & 0 & 0 \\ 0 & 2 - \lambda & 1 \\ 0 & 0 & 1 - \lambda \end{vmatrix} \\ &= (2 - \lambda)^2(1 - \lambda). \end{aligned}$$

Die Nullstellen des charakteristischen Polynoms, d.h. die Eigenwerte von A sind $\lambda_1 = 2$ und $\lambda_2 = 1$. Für die Eigenvektoren zum Eigenwert λ_i , $i = 1, 2$, gilt

$$(A - \lambda_i I_3)\underline{\mathbf{v}} = \begin{pmatrix} 2 - \lambda_i & 0 & 0 \\ 0 & 2 - \lambda_i & 1 \\ 0 & 0 & 1 - \lambda_i \end{pmatrix} \underline{\mathbf{v}} = \underline{\mathbf{0}}.$$

Es ergibt sich

$$E_{\lambda_1=2} = \{\underline{\mathbf{v}} \in \mathbb{C}^3 : v_3 = 0\}$$

und

$$E_{\lambda_2=1} = \{t\underline{\mathbf{w}} : t \in \mathbb{C}\}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Das charakteristische Polynom einer Matrix $A \in M(n, n, \mathbb{C})$ ist ein Polynom vom Grad n , nach dem Fundamentalsatz der Algebra (Satz 7.1.1) zerfällt es in Linearfaktoren, wobei $a_n = (-1)^n$ der höchste Koeffizient ist (warum?),

$$\chi_A(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \dots (\lambda - \lambda_r)^{k_r}.$$

Dabei ist $\mathbb{N} \ni r \leq n$, $k_1 + k_2 + \dots + k_r = n$ und für $i = 1, \dots, r$ sind die λ_i die Eigenwerte von A mit den **algebraischen Vielfachheiten** (als Nullstellen des charakteristischen Polynoms) k_i . Werden also die **Eigenwerte mit ihrer Vielfachheit gezählt, so hat A genau n Eigenwerte**. Im obigen Beispiel ist $\lambda_1 = 2$ mit der Vielfachheit 2 zu zählen, $\lambda_2 = 1$ mit der Vielfachheit 1. In dem Beispiel stimmt die algebraische Vielfachheit mit der Dimension des zugehörigen Eigenraums überein – E_{λ_1} ist dort eine Ebene, E_{λ_2} eine Gerade im \mathbb{C}^3 (zur Veranschaulichung stelle man sich in dem Beispiel den Eigenraum als Teilraum des \mathbb{R}^3 vor). Das ist jedoch nicht immer der Fall.

Beispiel. Es sei

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in M(2, 2, \mathbb{R}).$$

Es ist

$$\chi_A(\lambda) = (1 - \lambda)^2,$$

$\lambda_1 = 1$ ist die einzige Nullstelle des charakteristischen Polynoms und hat die algebraische Vielfachheit 2. Der Eigenraum ist jedoch nur ein-dimensional, es ist

$$E_{\lambda_1=1} = \{t\underline{\mathbf{w}}, t \in \mathbb{C}\}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

es gibt keine weiteren von \underline{w} linear unabhängigen Eigenvektoren. Die geometrische Vielfachheit, das ist die Dimension des Eigenraums, ist 1 und damit kleiner als die algebraische Vielfachheit.

Definition 15.1.2

Es sei $\lambda \in \mathbb{C}$ ein Eigenwert der Matrix $A \in M(n, n, \mathbb{C})$. Dann heißt

$$1 \leq g(\lambda) = n - \operatorname{rg}(A - \lambda I_n)$$

die geometrische Vielfachheit des Eigenwertes λ .

Bemerkungen.

- i)* Was ist ein k -dimensionaler Unterraum des \mathbb{C}^n , $k \leq n$?
- ii)* In den letzten beiden Beispielen kann ebensogut reell gerechnet werden.
- iii)* Nach der Dimensionsformel (Satz 8.2.4) ist

$$n - \operatorname{rg}(A - \lambda I_n) = \dim(\operatorname{kern}(A - \lambda I_n)),$$

die Definition liefert also genau die Dimension des Eigenraumes.

- iv)* Warum ist $g(\lambda) \geq 1$?

Nach Kapitel 9.2 kann eine Matrix $A \in M(n, n, \mathbb{C})$ aufgefasst werden als die Darstellung einer linearen Abbildung $L: \mathbb{C}^n \rightarrow \mathbb{C}^n$ bzgl. einer Basis \mathcal{A} des \mathbb{C}^n . Bei einem Basiswechsel transformiert sich die darstellende Matrix zu

$$B = S^{-1}AS$$

mit einer regulären Matrix S . Man nennt Matrizen, die durch eine solche Transformation auseinander hervorgehen **ähnlich**.

Haben ähnliche Matrizen das gleiche charakteristische Polynom?

Satz 15.1.3

*Die charakteristischen Polynome (und damit die Eigenwerte und ihre algebraischen Vielfachheiten) **ähnlicher Matrizen sind gleich.***

Beweis. Es seien $A, S \in M(n, n, \mathbb{C})$, S sei regulär und es sei $B = S^{-1}AS$. Aus dem Determinantenmultiplikationssatz (Satz 8.6.1) folgt

$$\begin{aligned} \det(A - \lambda I_n) &= \det(SS^{-1}(A - \lambda I_n)) \\ &= \det(S)\det(S^{-1}(A - \lambda I_n)) \\ &= \det(S^{-1}A - \lambda I_n S^{-1})\det(S) \\ &= \det(S^{-1}AS - \lambda I_n), \end{aligned}$$

also genau die Behauptung. □

Bemerkung. Dementsprechend kann auch von **Eigenwerten einer linearen Abbildung** gesprochen werden.

Ist nun \underline{v} ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{C}$ einer Matrix $A \in M(n, n, \mathbb{C})$ und ist $S \in M(n, n, \mathbb{C})$ regulär, so ist äquivalent dazu

$$(S^{-1}AS)(S^{-1}\underline{v}) = \lambda(S^{-1}\underline{v}),$$

also dass $S^{-1}\underline{v}$ Eigenvektor der Matrix $B = S^{-1}AS$ zum Eigenwert λ ist (man vergleiche die Koordinatentransformation aus Satz 9.2.1).

Korollar 15.1.1

- i) Bei Ähnlichkeitstransformationen bleiben auch die **geometrischen Vielfachheiten unverändert.***
- ii) Die geometrische Vielfachheit eines Eigenwertes ist **stets kleiner als oder gleich der algebraischen Vielfachheit.***

Beweis. Der erste Teil des Korollars ist bereits gezeigt (warum?).

Zum Beweis der zweiten Aussage sei $\lambda_0 \in \mathbb{C}$ Eigenwert einer Matrix $A \in M(n, n, \mathbb{C})$ der geometrischen Vielfachheit $k \leq n$. Dann existiert eine Basis $(\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(k)})$ des Eigenraums E_{λ_0} .

Die Matrix A repräsentiert (bzgl. der kanonischen Basis) eine lineare Abbildung $L: \mathbb{C}^n \rightarrow \mathbb{C}^n$ mit

$$L(\underline{\mathbf{v}}^{(i)}) = \lambda_0 \underline{\mathbf{v}}^{(i)}, \quad i = 1, \dots, k.$$

Nun wird die Basis des Eigenraums zu einer Basis $\mathcal{B} = (\underline{\mathbf{v}}^{(1)}, \dots, \underline{\mathbf{v}}^{(k)}, \underline{\mathbf{v}}^{(k+1)}, \dots, \underline{\mathbf{v}}^{(n)})$ des \mathbb{C}^n ergänzt und die Matrixdarstellung B der linearen Abbildung bzgl. dieser Basis betrachtet. Nach Abschnitt 9.2 geht B aus A durch eine Ähnlichkeitstransformation hervor, λ_0 ist auch Eigenwert von B mit geometrischer Vielfachheit k (vgl. i). Ist $1 \leq i \leq k$, so stehen in der i^{ten} Spalte von B die Koordinaten von $L(\underline{\mathbf{v}}^{(i)})$ bzgl. der Basis \mathcal{B} , es ist also

$$B = \begin{pmatrix} \lambda_0 & 0 & 0 & \dots & 0 & b_{1(k+1)} & \dots & b_{1n} \\ 0 & \lambda_0 & 0 & \dots & 0 & b_{2(k+1)} & \dots & b_{2n} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_0 & b_{k(k+1)} & \dots & b_{kn} \\ 0 & 0 & 0 & \dots & 0 & b_{(k+1)(k+1)} & \dots & b_{(k+1)m} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & b_{n(k+1)} & \dots & b_{nn} \end{pmatrix}.$$

Nach dem Laplaceschen Entwicklungssatz (Satz 8.6.2) folgt unmittelbar (wie?), dass die algebraische Vielfachheit von λ_0 mindestens k sein muss. \square

15.2 Diagonalisierbare Matrizen (Hauptachsentransformation)

Verfolgt man den Beweis von Korollar 15.1.1 (d.h. betrachtet man die Gestalt der Matrix B in Beweis des Korollars), so stellt sich die Frage nach der [Diagonalisierbarkeit](#) quadratischer Matrizen. Ist

$A \in M(n, n, \mathbb{C})$, so hofft man, A (unter geeigneten Voraussetzungen) mit Hilfe einer regulären Matrix $S \in M(n, n, \mathbb{C})$ zu transformieren auf

$$B = S^{-1}AS \quad \text{mit} \quad B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Daraus könnte man nach dem oben Gesagten beispielsweise direkt schließen, dass $\lambda_1, \dots, \lambda_n$ (nicht notwendig verschieden) die Eigenwerte von A sind.

Beobachtung. Gibt es eine solche Transformation, so schreibt man S als Tupel von Spaltenvektoren,

$$S = (\underline{\mathbf{v}}^{(1)} \quad \underline{\mathbf{v}}^{(2)} \quad \dots \quad \underline{\mathbf{v}}^{(n)}).$$

Es ist $B = S^{-1}AS$ äquivalent zu $SB = AS$, also

$$(\underline{\mathbf{v}}^{(1)} \quad \underline{\mathbf{v}}^{(2)} \quad \dots \quad \underline{\mathbf{v}}^{(n)}) \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix} = A(\underline{\mathbf{v}}^{(1)} \quad \underline{\mathbf{v}}^{(2)} \quad \dots \quad \underline{\mathbf{v}}^{(n)})$$

und man erkennt

$$(\lambda_1 \underline{\mathbf{v}}^{(1)} \quad \lambda_2 \underline{\mathbf{v}}^{(2)} \quad \dots \quad \lambda_n \underline{\mathbf{v}}^{(n)}) = (A \underline{\mathbf{v}}^{(1)} \quad A \underline{\mathbf{v}}^{(2)} \quad \dots \quad A \underline{\mathbf{v}}^{(n)}).$$

Die Spaltenvektoren der Transformationsmatrix sind Eigenvektoren von A . Man erkennt weiter: Ist $L: \mathbb{C}^n \rightarrow \mathbb{C}^n$ die lineare Abbildung $\underline{\mathbf{x}} \mapsto A \underline{\mathbf{x}}$ (Darstellung bzgl. der kanonischen Basis), und betrachtet man diese Abbildung in den Koordinaten bzgl. der Eigenvektoren, d.h. relativ zu der Basis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}, \dots, \underline{\mathbf{v}}^{(n)})$, so ist L in jeder

Komponente einfach eine Streckung um den entsprechenden Eigenwert.

Es ist aber nicht jede Matrix diagonalisierbar, aus dem bisher Gesagten kann lediglich entnommen werden:

Satz 15.2.1

- i) Eine Matrix $A \in M(n, n, \mathbb{C})$ lässt sich genau dann auf Diagonalform transformieren, wenn sie n linear unabhängige Eigenvektoren hat. Die Spaltenvektoren der Transformationsmatrix sind genau obige Eigenvektoren, in der Diagonalmatrix stehen die zugehörigen Eigenwerte (nicht notwendig verschieden).*
- ii) Eine Matrix $A \in M(n, n, \mathbb{C})$ ist genau dann diagonalisierbar, wenn für jeden Eigenwert die algebraische Vielfachheit gleich der geometrischen Vielfachheit ist.*

Man sieht es in der Regel einer Matrix A nicht ohne Weiteres an, ob sie eines der Kriterien zur Diagonalisierbarkeit erfüllt.

Impliziert eine spezielle Struktur einer Matrix deren Diagonalisierbarkeit?

Zur Beantwortung dieser Frage sei folgender Satz vorausgeschickt:

Satz 15.2.2

Die Matrix $A \in M(n, n, \mathbb{R})$ habe *reelle Eintragungen*.

- i) Ist $\lambda = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, ein *komplexer Eigenwert* von A , so ist auch *die konjugiert komplexe Zahl $\bar{\lambda}$ Eigenwert* von A .
- ii) Ist $\underline{\mathbf{v}} = \underline{\mathbf{a}} + i\underline{\mathbf{b}} \in \mathbb{C}^n$, $\underline{\mathbf{a}}, \underline{\mathbf{b}} \in \mathbb{R}^n$, ein *komplexer Eigenvektor* zum Eigenwert λ von A , so ist $\underline{\bar{\mathbf{v}}} := \underline{\mathbf{a}} - i\underline{\mathbf{b}}$ *Eigenvektor zum Eigenwert $\bar{\lambda}$* von A .

Beweis. Die erste Aussage ist wieder eine Konsequenz aus der Tatsache, dass für ein reelles Polynom mit $\lambda \in \mathbb{C}$ auch $\bar{\lambda}$ eine Nullstelle ist (vgl. Kapitel 7.1). Zum Beweis der zweiten Aussage beobachtet man

$$A(\underline{\mathbf{a}} - i\underline{\mathbf{b}}) = \overline{A(\underline{\mathbf{a}} + i\underline{\mathbf{b}})} = \overline{\lambda \underline{\mathbf{a}} + i\lambda \underline{\mathbf{b}}} = \bar{\lambda} \underline{\mathbf{a}} - i\bar{\lambda} \underline{\mathbf{b}} = \bar{\lambda}(\underline{\mathbf{a}} - i\underline{\mathbf{b}}),$$

wobei ausgenutzt wurde, dass A , $\underline{\mathbf{a}}$ und $\underline{\mathbf{b}}$ nur reelle Eintragungen haben. \square

Korollar 15.2.1

Es sei $A \in M(n, n, \mathbb{R})$ eine *symmetrische Matrix mit reellen Eintragungen*. Dann gilt

- i) Die *Eigenwerte von A sind reell* und damit auch die *Eigenvektoren*.
- ii) *Eigenvektoren zu verschiedenen Eigenwerten stehen senkrecht aufeinander*.

Beweis. Es sei $A \in M(n, n, \mathbb{R})$ symmetrisch, d.h. $A = A^T$ und es sei λ ein Eigenwert von A , $\underline{\mathbf{v}} \neq \underline{\mathbf{0}}$ ein zugehöriger Eigenvektor. Es gilt also

$$\begin{aligned} A\underline{\mathbf{v}} &= \lambda \underline{\mathbf{v}} \quad \text{und damit} \quad \underline{\mathbf{v}}^T A^T = \lambda \underline{\mathbf{v}}^T, \\ A\underline{\bar{\mathbf{v}}} &= \bar{\lambda} \underline{\bar{\mathbf{v}}}. \end{aligned}$$

Es folgt aus der Symmetrie von A

$$\begin{aligned}\lambda \langle \underline{\mathbf{v}}, \underline{\mathbf{v}} \rangle &= \lambda \underline{\mathbf{v}}^T \underline{\mathbf{v}} = \underline{\mathbf{v}}^T A^T \underline{\mathbf{v}} = \underline{\mathbf{v}}^T A \underline{\mathbf{v}} = \underline{\mathbf{v}}^T \bar{\lambda} \underline{\mathbf{v}} \\ &= \bar{\lambda} \langle \underline{\mathbf{v}}, \underline{\mathbf{v}} \rangle.\end{aligned}$$

Also ist $\lambda = \bar{\lambda}$, der Eigenwert ist reell.

Sind nun $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ Eigenvektoren zu $\lambda \in \mathbb{R}$ bzw. $\mu \in \mathbb{R}$, $\lambda \neq \mu$, so gilt

$$\begin{aligned}(\lambda - \mu) \underline{\mathbf{v}}^T \underline{\mathbf{w}} &= \lambda \underline{\mathbf{v}}^T \underline{\mathbf{w}} - \underline{\mathbf{v}}^T \mu \underline{\mathbf{w}} = \underline{\mathbf{v}}^T A^T \underline{\mathbf{w}} - \underline{\mathbf{v}}^T A \underline{\mathbf{w}} \\ &= \underline{\mathbf{v}}^T A \underline{\mathbf{w}} - \underline{\mathbf{v}}^T A \underline{\mathbf{w}} = 0,\end{aligned}$$

es sind $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ orthogonal. □

Bemerkung. Am Beispiel einer reellen Drehmatrix sieht man, dass ein Analogon für nicht-symmetrische Matrizen nicht richtig sein kann.

Als Hauptergebnis dieses Abschnittes folgt schließlich

Satz 15.2.3 (*Hauptachsentransformation*)

Eine symmetrische reellwertige Matrix $A \in M(n, n, \mathbb{R})$ ist diagonalisierbar.

Bemerkungen.

- i) Die Transformationsmatrix ist eine Matrix S , deren Spaltenvektoren aus einem Orthonormalsystem von Eigenvektoren besteht. Die Matrix ist orthonormal, es gilt $S^T = S^{-1}$. Es folgt $\det S = \pm 1$, durch Spaltenvertauschung kann $\det S = 1$ erreicht werden, **S ist dann eine Drehmatrix.**
- ii) Ein analoges Resultat gilt für Hermitesche Matrizen $A \in M(n, n, \mathbb{C})$, d.h. (vgl. Kapitel 8.1) Matrizen mit der Eigenschaft $A = A^* := \bar{A}^T$.
- iii) Der Name Hauptachsentransformation bekommt in nächsten Abschnitt eine geometrische Interpretation.

Beweisidee. Zu einem reellen (vgl. Korollar 15.2.1) Eigenwert λ_1 wähle man einen Eigenvektor $\underline{\mathbf{v}}$ der Länge eins und ergänze ihn zu einer Orthonormalbasis $(\underline{\mathbf{v}}, \underline{\mathbf{w}}^{(2)}, \dots, \underline{\mathbf{w}}^{(n)})$. Ist M die Matrix mit diesen Spaltenvektoren, so gilt

$$\begin{aligned}
 M^T A M &= \begin{pmatrix} \underline{\mathbf{v}}^T \\ (\underline{\mathbf{w}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{w}}^{(n)})^T \end{pmatrix} A (\underline{\mathbf{v}} \ \underline{\mathbf{w}}^{(2)} \ \dots \ \underline{\mathbf{w}}^{(n)}) \\
 &= \begin{pmatrix} \underline{\mathbf{v}}^T \\ (\underline{\mathbf{w}}^{(2)})^T \\ \vdots \\ (\underline{\mathbf{w}}^{(n)})^T \end{pmatrix} (\lambda \underline{\mathbf{v}} \ A \underline{\mathbf{w}}^{(2)} \ \dots \ A \underline{\mathbf{w}}^{(n)}) \\
 &= \begin{pmatrix} \lambda \underline{\mathbf{v}}^T \underline{\mathbf{v}} & \underline{\mathbf{v}}^T A \underline{\mathbf{w}}^{(2)} & \dots & \underline{\mathbf{v}}^T A \underline{\mathbf{w}}^{(n)} \\ \lambda (\underline{\mathbf{w}}^{(2)})^T \underline{\mathbf{v}} & (\underline{\mathbf{w}}^{(2)})^T A \underline{\mathbf{w}}^{(2)} & \dots & (\underline{\mathbf{w}}^{(2)})^T A \underline{\mathbf{w}}^{(n)} \\ \vdots & \vdots & & \vdots \\ \lambda (\underline{\mathbf{w}}^{(n)})^T \underline{\mathbf{v}} & (\underline{\mathbf{w}}^{(n)})^T A \underline{\mathbf{w}}^{(2)} & \dots & (\underline{\mathbf{w}}^{(n)})^T A \underline{\mathbf{w}}^{(n)} \end{pmatrix}.
 \end{aligned}$$

Da $\underline{\mathbf{v}}$ auf die Länge eins normiert ist und senkrecht auf $\underline{\mathbf{w}}^{(i)}$, $i = 2, \dots, n$, steht, ist der oberste Eintrag in der ersten Spalte λ , alle anderen verschwinden. In der ersten Zeile beachtet man, dass aufgrund der Symmetrie von A für $i = 2, \dots, n$ gilt

$$\underline{\mathbf{v}}^T A \underline{\mathbf{w}}^{(i)} = (A^T \underline{\mathbf{v}})^T \underline{\mathbf{w}}^{(i)} = (A \underline{\mathbf{v}})^T \underline{\mathbf{w}}^{(i)} = \lambda \underline{\mathbf{v}}^T \underline{\mathbf{w}}^{(i)} = 0,$$

wobei in der letzten Gleichheit wieder die Orthogonalität ausgenutzt wurde. Es ist gezeigt

$$M^T A M = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & (\underline{\mathbf{w}}^{(2)})^T A \underline{\mathbf{w}}^{(2)} & \dots & (\underline{\mathbf{w}}^{(2)})^T A \underline{\mathbf{w}}^{(n)} \\ \vdots & \vdots & & \vdots \\ 0 & (\underline{\mathbf{w}}^{(n)})^T A \underline{\mathbf{w}}^{(2)} & \dots & (\underline{\mathbf{w}}^{(n)})^T A \underline{\mathbf{w}}^{(n)} \end{pmatrix},$$

wobei die verbleibende $(n-1) \times (n-1)$ Restmatrix aufgrund der Symmetrie selbst wieder symmetrisch ist. Induktiv kann der Satz bewiesen werden, wenn die Orthogonalität von Eigenvektoren zu verschiedenen Eigenwerten (Korollar 15.2.1, *ii*) berücksichtigt wird und wenn für Eigenräume mit Dimension größer als 1 eine Orthonormalbasis aus Eigenvektoren gewählt wird (vgl. das Gram-Schmidtsche Verfahren, Kapitel 6.2). \square

Beispiel. Man betrachte die symmetrische Matrix

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \in M(3, 3, \mathbb{R}).$$

Die Eigenwerte von A berechnen sich zu $\lambda_1 = 1$, $\lambda_2 = 3$ und $\lambda_3 = -4$. Zu λ_1 und λ_2 findet man die auf Einheitslänge normierten Eigenvektoren

$$\underline{\mathbf{v}}^{(1)} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \frac{1}{\sqrt{14}} \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}.$$

Einen (normierten) Eigenvektor zum dritten Eigenwert erhält man direkt aus der Aussage, dass Eigenvektoren zu verschiedenen Eigenwerten senkrecht aufeinander stehen (Korollar 15.2.1)

$$\underline{\mathbf{v}}^{(3)} = \underline{\mathbf{v}}^{(1)} \times \underline{\mathbf{v}}^{(2)} = \frac{1}{\sqrt{35}} \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix}.$$

Ist $S = (\underline{\mathbf{v}}^{(1)} \ \underline{\mathbf{v}}^{(2)} \ \underline{\mathbf{v}}^{(3)})$, so berechne man $S^T A S$.

15.3 Anwendung auf Quadriken (Normalformen von Quadriken)

Mit Hilfe der Hauptachsentransformation sollen nun Lösungsmengen quadratischer Gleichungen in n Unabhängigen, sogenannte **Quadriken**, geometrisch charakterisiert werden. Es sei also ein quadratisches Polynom $q(\underline{\mathbf{x}}): \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben,

$$\begin{aligned} q(\underline{\mathbf{x}}) &= \underline{\mathbf{x}}^T A \underline{\mathbf{x}} + \underline{\mathbf{b}}^T \underline{\mathbf{x}} + c \\ &= \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c. \end{aligned}$$

Hierbei sind $A = (a_{ij}) \in M(n, n, \mathbb{R})$, $\underline{\mathbf{b}} \in \mathbb{R}^n$ und $c \in \mathbb{R}$ vorgegeben. Wegen

$$a_{ij} x_i x_j + a_{ji} x_j x_i = \left(\frac{a_{ij} + a_{ji}}{2} \right) x_i x_j + \left(\frac{a_{ij} + a_{ji}}{2} \right) x_j x_i$$

bleibt die Gleichung erhalten, wenn A durch $(A + A^T)/2$ ersetzt wird. Mit anderen Worten kann o.E. angenommen werden, dass A symmetrisch ist, $A = A^T$. Untersucht werden soll die geometrische Struktur der Nullstellenmenge, also von

$$\{ \underline{\mathbf{x}} \in \mathbb{R}^n : q(\underline{\mathbf{x}}) = 0 \}.$$

Die Vorgehensweise soll exemplarisch anhand von zwei Beispielen vorgestellt werden:

Ein zweidimensionales Beispiel. Es sei $n = 2$,

$$A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} \in M(2, 2, \mathbb{R}), \quad \underline{\mathbf{b}} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad c = -6,$$

und es sei

$$\underline{\mathbf{x}}^T A \underline{\mathbf{x}} + \underline{\mathbf{b}}^T \underline{\mathbf{x}} + c = 0, \quad \underline{\mathbf{x}} \in \mathbb{R}^2.$$

Die Gleichung kann auch geschrieben werden in der Form

$$2x_1^2 - x_2^2 + 4x_1x_2 - 2x_1 + x_2 - 6 = 0.$$

In dieser Form ist keinesfalls ersichtlich, wie die Lösungsmenge aussieht, die Gleichung muss zunächst geeignet **transformiert** werden.

Dazu wird im ersten Schritt A diagonalisiert, d.h. zu bestimmen sind die **Eigenwerte und eine Orthonormalbasis von Eigenvektoren**. Dies ist nach Satz 15.2.3 möglich, da A symmetrisch ist. In dem Beispiel ist das charakteristische Polynom

$$\chi_A(\lambda) = (2 - \lambda)(-1 - \lambda) - 4 = \lambda^2 - \lambda - 6 ,$$

es existieren die beiden (einfachen) Nullstellen $\lambda_1 = 3$ und $\lambda_2 = -2$, eine Orthonormalbasis von Eigenvektoren ist gegeben durch

$$\underline{\mathbf{v}}^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} , \quad \underline{\mathbf{v}}^{(2)} = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix} .$$

Die Matrix bestehend aus diesen Spaltenvektoren,

$$S = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} ,$$

bringt A über die Transformation (es gilt $S^{-1} = S^T$)

$$S^T A S = B = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

auf Diagonalgestalt. Die **Koordinaten von $\underline{\mathbf{x}}$ bzgl. der Basis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)})$** sind nach Satz 9.2.1 gegeben durch

$$\underline{\mathbf{y}} = S^T \underline{\mathbf{x}} , \quad \text{also} \quad \underline{\mathbf{x}} = S \underline{\mathbf{y}} ,$$

und es gilt

$$\underline{\mathbf{x}}^T A \underline{\mathbf{x}} = \underline{\mathbf{y}}^T S^T A S \underline{\mathbf{y}} = \underline{\mathbf{y}}^T B \underline{\mathbf{y}} .$$

Zudem ist

$$\underline{\mathbf{b}}^T \underline{\mathbf{x}} = \underline{\mathbf{b}}^T S \underline{\mathbf{y}} =: \underline{\mathbf{d}}^T \underline{\mathbf{y}} = \frac{1}{\sqrt{5}} (-3 \ 4) \underline{\mathbf{y}} ,$$

die Gleichung transformiert sich zu

$$\underline{\mathbf{y}}^T B \underline{\mathbf{y}} + \underline{\mathbf{d}}^T \underline{\mathbf{y}} - 6 = 0 ,$$

oder ausführlich

$$3y_1^2 - 2y_2^2 - \frac{3}{\sqrt{5}}y_1 + \frac{4}{\sqrt{5}}y_2 - 6 = 0 .$$

Im nächsten Schritt transformiert man die Gleichung mittels einer quadratischen Ergänzung auf

$$3\left[y_1^2 - \frac{1}{2\sqrt{5}}\right]^2 - 3\left[\frac{1}{2\sqrt{5}}\right]^2 - 2\left[y_2 - \frac{1}{\sqrt{5}}\right]^2 + 2\left[\frac{1}{\sqrt{5}}\right]^2 - 6 = 0,$$

also

$$3\left[y_1^2 - \frac{1}{2\sqrt{5}}\right]^2 - 2\left[y_2 - \frac{1}{\sqrt{5}}\right]^2 - \frac{23}{4} = 0.$$

Setzt man schließlich

$$z_1 := y_1 - \frac{1}{2\sqrt{5}}, \quad z_2 := y_2 - \frac{1}{\sqrt{5}}, \quad a := \sqrt{\frac{23}{12}}, \quad b := \sqrt{\frac{23}{8}},$$

so transformiert sich die Gleichung auf die **Normalform**

$$\frac{z_1^2}{a^2} - \frac{z_2^2}{b^2} - 1 = 0,$$

die Lösungsmenge ist eine **Hyperbel**.

Geometrische Interpretation. (vgl. die Abbildungen 15.1–15.3)

i) Ausgehend von einem Lösungsvektor \underline{x} (seiner Darstellung bzgl. der Standardbasis) wird im ersten Schritt eine Hauptachsentransformation durchgeführt. Der transformierte Koordinatenvektor \underline{y} gibt die Koordinaten bzgl. der neuen Orthonormalbasis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)})$ aus Eigenvektoren von A an. Dabei ist die Reihenfolge von $\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)}$ so gewählt, dass $\det S = 1$ gilt, die Basis $(\underline{\mathbf{v}}^{(1)}, \underline{\mathbf{v}}^{(2)})$ geht durch eine Rotation aus der kanonischen Basis hervor.

Die neuen Koordinatenachsen bzgl. derer das Problem angegangen wird sind in Abbildung 15.1 blau angedeutet.

ii) Die quadratische Ergänzung im zweiten Schritt zusammen mit der Einführung von z_1, z_2 entspricht einer Translation des Ursprunges des Koordinatensystems in den Punkt

$$\underline{\mathbf{O}} = \frac{1}{2\sqrt{5}}\underline{\mathbf{v}}^{(1)} + \frac{1}{\sqrt{5}}\underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1/2 \end{pmatrix}.$$

Ist nämlich

$$\underline{\mathbf{x}} = y_1\underline{\mathbf{v}}^{(1)} + y_2\underline{\mathbf{v}}^{(2)} = \underline{\mathbf{O}} + z_1\underline{\mathbf{v}}^{(1)} + z_2\underline{\mathbf{v}}^{(2)},$$

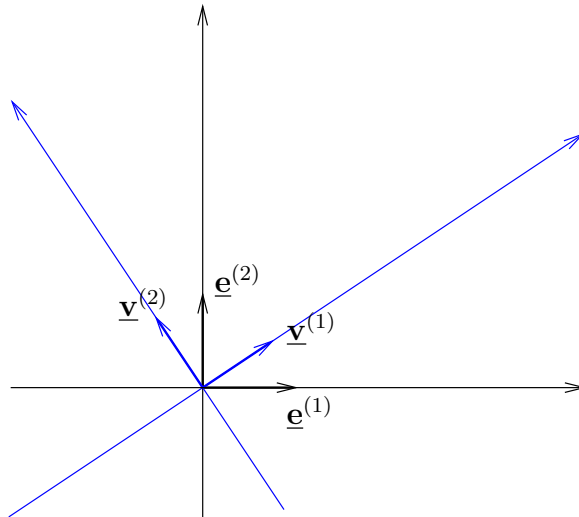


Abbildung 15.1: Die Hauptachsentransformation entspricht einer Rotation der Basisvektoren.

und ist (Koordinaten bzgl. der Basis $(\underline{v}^{(1)}, \underline{v}^{(2)})$)

$$\underline{\mathbf{O}} = \frac{1}{2\sqrt{5}}\underline{v}^{(1)} + \frac{1}{\sqrt{5}}\underline{v}^{(2)},$$

so ist

$$z_1 = y_1 - \frac{1}{2\sqrt{5}}, \quad z_2 = y_2 - \frac{1}{\sqrt{5}}.$$

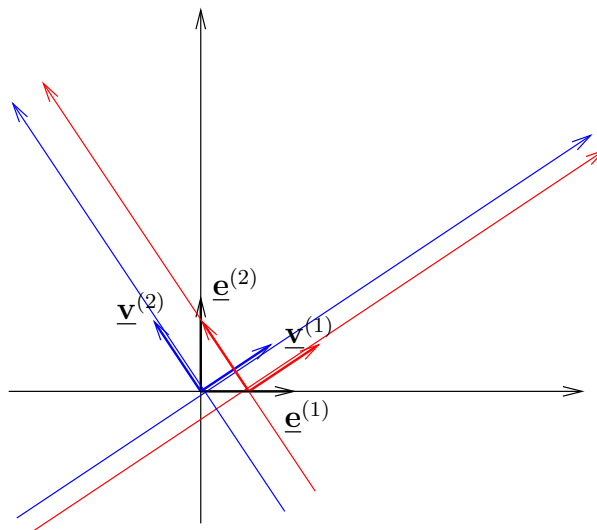


Abbildung 15.2: Zur Verschiebung des Ursprungs des Koordinatensystems.

Die Lösungsmenge ist eine Hyperbel, deren Achsen durch das in

Abbildung 15.2 rot angedeutete Koordinatensystem gegeben sind. (vgl. auch Abbildung 15.3).

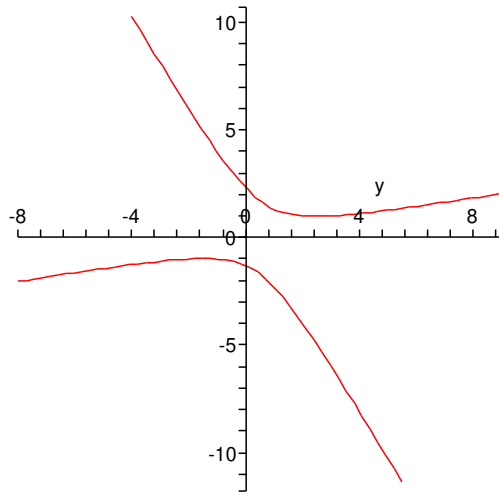


Abbildung 15.3: Die Lösungsmenge ist eine Hyperbel.

Ein dreidimensionales Beispiel. Es sei $n = 2$,

$$A = \begin{pmatrix} 5 & -2 & -4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{pmatrix} \in M(3, 3, \mathbb{R}), \quad \underline{\mathbf{b}} = \begin{pmatrix} 16 \\ 8 \\ -20 \end{pmatrix}, \quad c = -19,$$

und es sei

$$\underline{\mathbf{x}}^T A \underline{\mathbf{x}} + \underline{\mathbf{b}}^T \underline{\mathbf{x}} + c = 0, \quad \underline{\mathbf{x}} \in \mathbb{R}^3.$$

Wieder wird zunächst eine Hauptachsentransformation durchgeführt, über eine elementare Rechnung findet man das charakteristische Polynom

$$\chi_A(\lambda) = -\lambda(\lambda - 9)^2,$$

die Eigenwerte sind $\lambda_1 = 9$ und $\lambda_2 = 0$.

Bemerkung. Das Beispiel unterscheidet sich in zwei Punkten vom letzten Beispiel: Einerseits ist λ_1 ein Eigenwert der Vielfachheit 2, andererseits hat die Matrix A nicht vollen Rang, da der zweite Eigenwert

0 ist.

Berechnet man zwei (Vielfachheit 2!) linear unabhängige Eigenvektoren zum Eigenwert λ_1 , so erhält man etwa

$$\underline{\tilde{\mathbf{v}}}^{(1)} = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \underline{\tilde{\mathbf{v}}}^{(2)} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

Aus diesen beiden linear unabhängigen Eigenvektoren zum Eigenwert λ_1 sind mit dem Gram-Schmidtschen Verfahren (vgl. Kapitel 6.2) zwei orthogonale Eigenvektoren der Länge eins zu konstruieren, man erhält

$$\underline{\mathbf{v}}^{(1)} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \frac{1}{3} \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}.$$

Wie im Beispiel am Ende des letzten Paragraphen findet man einen orthogonalen Eigenvektor mit Einheitslänge zum verbleibenden Eigenwert über das Vektorprodukt,

$$\underline{\mathbf{v}}^{(3)} = \underline{\mathbf{v}}^{(1)} \times \underline{\mathbf{v}}^{(2)} = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Als Transformationsmatrix ergibt sich

$$S = \frac{1}{3} \begin{pmatrix} 1 & -2 & 2 \\ 2 & 2 & 1 \\ -2 & 1 & 2 \end{pmatrix}.$$

Die entsprechende Drehung des Koordinatensystems führt auf

$$9y_1^2 + 9y_2^2 + 24y_1 - 12y_2 + 19 = 0,$$

wieder schließt sich eine quadratische Ergänzung an und mit $z_1 := y_1 + 4/3$, $z_2 := y_2 - 2/3$ lautet die Normalform

$$9z_1^2 + 9z_2^2 - 1 = 0.$$

Es handelt sich um einen Kreiszyylinder mit der z_3 -Achse als Symmetrieachse.

Bemerkung. Die Tatsache $\operatorname{rg} A = 2 < 3 = n$ spiegelt sich darin wider, dass keine Komponente z_3 in der Normalform auftaucht.

Übung. Man schlage in der Literatur die Normalformen der Quadriken im \mathbb{R}^2 und im \mathbb{R}^3 nach.

15.4 Die Jordansche Normalform (Hauptvektoren; Kettenbedingung; Jordan-Kästchen)

Nach Satz 15.2.1 ist eine Matrix genau dann diagonalisierbar, wenn die algebraische und die geometrische Vielfachheit jedes Eigenwertes übereinstimmen. In diesem Paragraphen soll anhand eines typischen Beispiels die Idee vermittelt werden, wie andernfalls vorzugehen ist.

Beispiel. Eine nicht diagonalisierbare Matrix ist

$$A = \begin{pmatrix} 2 & -4 & 2 \\ 0 & -2 & -2 \\ 0 & 8 & 6 \end{pmatrix} \in M(3, 3, \mathbb{R}).$$

Das charakteristische Polynom von A berechnet sich nämlich mit Hilfe des Laplaceschen Entwicklungssatzes (Satz 8.6.2) zu

$$\chi_A(\lambda) = \det(A - \lambda I_3) = (2 - \lambda)[(-2 - \lambda)(6 - \lambda) + 16] = (2 - \lambda)^3.$$

Demnach ist $\lambda = 2$ die einzige Nullstelle des charakteristischen Polynoms und hat die algebraische Vielfachheit $a(\lambda = 2) = 3$. Zur Bestimmung der Eigenvektoren beobachtet man, dass aus

$$(A - \lambda I_3)\underline{\mathbf{v}} = \begin{pmatrix} 0 & -4 & 2 \\ 0 & -4 & -2 \\ 0 & 8 & 4 \end{pmatrix} \underline{\mathbf{v}} = \underline{\mathbf{0}}$$

folgt

$$\underline{\mathbf{v}} = t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Die geometrische Vielfachheit ist somit $g(\lambda = 2) = 1$ und wie behauptet ist A nicht diagonalisierbar.

Nun sei wie zu Beginn des Kapitels das lineare Differentialgleichungssystem erster Ordnung mit konstanten Koeffizienten

$$\underline{\mathbf{y}}' = A\underline{\mathbf{y}}$$

betrachtet. Der Ansatz aus Kapitel 15.1 liefert nur die Lösung

$$\underline{\mathbf{y}}^{(0)}(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} e^{2x}$$

sowie deren Vielfache. In der Hoffnung auf eine größere Lösungsmenge und motiviert durch die in Kapitel 14.2 diskutierte Reduktion der Ordnung, wird der Ansatz

$$\underline{\mathbf{y}}^{(1)}(x) = \underline{\mathbf{v}}^{(0)} x e^{2x} + \underline{\mathbf{v}}^{(1)} e^{2x}$$

gemacht, wobei $\underline{\mathbf{v}}^{(0)}$ einen Eigenvektor von A bezeichne — o.E. und

nur zur Vereinfachung der Rechnung in diesem Beispiel $\underline{\mathbf{v}}^{(0)} = \begin{pmatrix} 8 \\ 0 \\ 0 \end{pmatrix}$

— und $\underline{\mathbf{v}}^{(1)}$ geeignet zu bestimmen ist.

Falls $\underline{\mathbf{y}}^{(1)}$ das Differentialgleichungssystem löst, so gilt

$$\begin{aligned} \underline{\mathbf{v}}^{(0)} e^{2x} + \underline{\mathbf{v}}^{(0)} x 2e^{2x} + \underline{\mathbf{v}}^{(1)} 2e^{2x} &= \underline{\mathbf{y}}^{(1)'}(x) \\ &= A\underline{\mathbf{y}}^{(1)} \\ &= A\underline{\mathbf{v}}^{(0)} x e^{2x} + A\underline{\mathbf{v}}^{(1)} e^{2x} \\ &= 2\underline{\mathbf{v}}^{(0)} x e^{2x} + A\underline{\mathbf{v}}^{(1)} e^{2x}. \end{aligned}$$

Als Bedingung an $\underline{\mathbf{v}}^{(1)}$ ergibt sich

$$(A - 2I_3)\underline{\mathbf{v}}^{(1)} = \underline{\mathbf{v}}^{(0)}. \quad (*)$$

Umgekehrt gilt: Ist $(*)$ erfüllt, so löst $\underline{\mathbf{y}}^{(1)}$ das Differentialgleichungssystem.

Genügt $\underline{\mathbf{v}}^{(1)}$ der Bedingung $(*)$, so heißt $\underline{\mathbf{v}}^{(1)}$ ein **Hauptvektor erster Stufe** zum Eigenwert $\lambda = 2$.

Bemerkung. An der Rechnung erkennt man, dass der einfache Ansatz $\underline{\mathbf{y}}^{(1)} = \underline{\mathbf{v}}^{(0)}xe^{2x}$ nicht zum Ziel führen kann.

In dem Beispiel lautet das zu lösende Gleichungssystem

$$\begin{pmatrix} 0 & -4 & 2 \\ 0 & -4 & -2 \\ 0 & 8 & 4 \end{pmatrix} \begin{pmatrix} v_1^{(1)} \\ v_2^{(1)} \\ v_3^{(1)} \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 0 \end{pmatrix}.$$

Die Lösungen sind gegeben durch

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} t \\ -1 \\ 2 \end{pmatrix}, \quad t \in \mathbb{R},$$

o.E. wird $t = 0$ gewählt, d.h. im Folgenden ist

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.$$

In dem Beispiel ist die Differenz zwischen algebraischer und geometrischer Vielfachheit 2 (zum einzigen Eigenwert $\lambda = 2$ der geometrischen Vielfachheit 1), man sucht noch nach einem **Hauptvektor zweiter Stufe**, d.h. per definitionem

$$(A - 2I_3)\underline{\mathbf{v}}^{(2)} = \underline{\mathbf{v}}^{(1)}. \quad (**)$$

Mit der fixierten Wahl von $\underline{\mathbf{v}}^{(1)}$ führt das auf das Gleichungssystem

$$\begin{pmatrix} 0 & -4 & 2 \\ 0 & -4 & -2 \\ 0 & 8 & 4 \end{pmatrix} \begin{pmatrix} v_1^{(2)} \\ v_2^{(2)} \\ v_3^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}.$$

Ein Hauptvektor zweiter Stufe ist

$$\underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1/8 \\ 1/4 \end{pmatrix}.$$

Analog zu Kapitel 15.2 wird nun die Matrix S ,

$$S = (\underline{\mathbf{v}}^{(0)} \ \underline{\mathbf{v}}^{(1)} \ \underline{\mathbf{v}}^{(2)}) = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -1 & 1/8 \\ 0 & 2 & 1/4 \end{pmatrix}$$

gebildet. S besteht jetzt nicht mehr aus einer Orthonormalbasis aus Eigenvektoren, die ja nicht existiert, sondern aus einem Eigenvektor, einem Hauptvektor erster Stufe und einem Hauptvektor zweiter Stufe, die über eine **Kettenbedingung** (d.h. über die Bedingungen (*) und (**)) miteinander verknüpft sind. S ist dementsprechend nicht mehr zwingend orthonormal, dennoch ist S regulär und in Analogie zu Kapitel 15.2 wird die transformierte Matrix

$$J = S^{-1}AS$$

untersucht. Die konkrete Berechnung ergibt

$$S^{-1} = \begin{pmatrix} 1/8 & 0 & 0 \\ 0 & -1/2 & 1/4 \\ 0 & 4 & 2 \end{pmatrix}$$

sowie

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Die transformierte Matrix J ist zwar nicht mehr von Diagonalgestalt, hat aber dennoch eine besonders einfache Struktur. Allgemein nennt man eine Matrix der Form

$$J(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

ein [Jordan¹-Kästchen](#).

Es gilt:

Satz 15.4.1 (*Jordansche Normalform im einfachsten Fall*)

Es sei $A \in M(n, n, \mathbb{C})$ und λ sei der einzige Eigenwert der Matrix A . Es habe λ die geometrische Vielfachheit $g(\lambda) = 1$ und die algebraische Vielfachheit $a(\lambda) > 1$. Dann transformiert die Matrix S , die genau wie im obigen Beispiel aus einem Eigenvektor und aus Hauptvektoren (bis zur Stufe $a(\lambda) - 1$) besteht, die Matrix A auf ein Jordan-Kästchen $J(\lambda)$, d.h.

$$S^{-1}AS = J(\lambda) .$$

Bemerkung. In seiner allgemeinen Form ist der Satz von der Jordanschen Normalform wesentlich komplexer: Hat A mehrere Eigenwerte, so zerfällt die transformierte Matrix in verschiedene Jordan-Kästchen. Ist die geometrische Vielfachheit eines Eigenwertes größer als 1, so zerfallen diese im Allgemeinen in weitere Jordan-Kästchen, deren Dimension

¹C. Jordan, 1838-1922; Paris.

in der Regel erst nach der Berechnung der Hauptvektoren bekannt ist. Ist etwa $g(\lambda) = 2$ und $a(\lambda) = 4$, so ist zunächst nicht klar, ob man je einen Eigenvektor und ein zugehörigen Hauptvektor erster Stufe oder einen Eigenvektor und einen Eigenvektor mit einem Hauptvektor erster und einem Hauptvektor zweiter Stufe zu betrachten hat.

Teil VI

Gewöhnliche
Differentialgleichungen (Teil 2)

Kapitel 16

Systeme linearer gewöhnlicher Differentialgleichungen erster Ordnung

Zunächst sei als Beobachtung festgehalten, dass die Diskussion von Systemen gewöhnlicher Differentialgleichungen **erster Ordnung** die Betrachtungen aus Kapitel 14 verallgemeinert. Ist nämlich

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \quad (1)$$

eine **explizite Differentialgleichung n^{ter} Ordnung** und wird das zugehörige Anfangswertproblem mit

$$\begin{aligned} y(x_0) &= y_0 \\ y'(x_0) &= y_0^{(1)} \\ &\vdots \\ y^{(n-1)}(x_0) &= y_0^{(n-1)} \end{aligned} \quad (2)$$

betrachtet, so ist dies äquivalent zum Anfangswertproblem für das **System gewöhnlicher Differentialgleichungen erster Ordnung**

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_n' &= f(x, y_1, \dots, y_n) \end{aligned} \quad (3)$$

mit den Anfangsbedingungen

$$\begin{aligned} y_1(x_0) &= y_0 \\ y_2(x_0) &= y_0^{(1)} \\ &\vdots \\ y_n(x_0) &= y_0^{(n-1)}. \end{aligned} \quad (4)$$

Die Äquivalenz der Anfangswertprobleme (1), (2) und (3), (4) wird in den Übungen besprochen.

Als Gegenstück und Verallgemeinerung zu Kapitel 14 werden im Folgenden Systeme gewöhnlicher linearer Differentialgleichungen erster Ordnung untersucht.

16.1 Systeme linearer gewöhnlicher Differentialgleichungen erster Ordnung (homogenes, inhomogenes System; Fundamentalsystem; allgemeine Lösung; Wronski-Determinante; Reduktion der Ordnung; Variation der Konstanten; Fundamentalmatrix)

In diesem Abschnitt wird stets das lineare System erster Ordnung

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} + \underline{\mathbf{r}}(x) \quad (*)$$

betrachtet. Dabei ist $x \in I$, I ein verallgemeinertes Intervall und die gesuchte Lösung $\underline{\mathbf{y}}$ eine Funktion $\underline{\mathbf{y}}: I \rightarrow \mathbb{R}^n$, $n \geq 1$, d.h.

$$\underline{\mathbf{y}}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}.$$

Es ist $A(x) \in M(n, n, \mathbb{R})$ für jedes $x \in I$, also

$$A(x) = \begin{pmatrix} a_{11}(x) & \dots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \dots & a_{nn}(x) \end{pmatrix}.$$

Dabei sei ausdrücklich betont, dass die Koeffizienten a_{ij} , $1 \leq i, j \leq n$ nicht konstant sein müssen und **von x abhängen** können. Die Funktion $\underline{\mathbf{r}}(x)$ ist wie die gesuchte Funktion $\underline{\mathbf{y}}(x)$ eine vektorwertige Funktion $I \rightarrow \mathbb{R}^n$ und das System lautet ausführlich

$$\begin{aligned} y_1' &= a_{11}(x)y_1 + \cdots + a_{1n}(x)y_n + r_1(x) \\ &\vdots \\ y_n' &= a_{n1}(x)y_1 + \cdots + a_{nn}(x)y_n + r_n(x). \end{aligned}$$

Um den Lösungsraum von (*) zu charakterisieren, benötigt man den Begriff der **linearen Unabhängigkeit**, der wörtlich aus Kapitel 14 übernommen werden kann (vgl. Definition 14.2.1, die anschließenden Bemerkungen übertragen sich ebenfalls)

Definition 16.1.1

*Es sei $I \subset \mathbb{R}$ ein verallgemeinertes Intervall. Die Funktionen $\underline{\mathbf{y}}^{(1)}, \dots, \underline{\mathbf{y}}^{(k)}: I \rightarrow \mathbb{R}^n$ heißen **linear unabhängig auf I** , falls aus*

$$c_1 \underline{\mathbf{y}}^{(1)}(x) + c_2 \underline{\mathbf{y}}^{(2)}(x) + \cdots + c_k \underline{\mathbf{y}}^{(k)}(x) = \underline{\mathbf{0}} \quad \text{für alle } x \in I,$$

$c_i \in \mathbb{R}, i = 1, \dots, k$, stets folgt

$$c_1 = c_2 = \cdots = c_k = 0.$$

*Sonst heißen die Funktionen **linear abhängig auf I** .*

Bemerkung. Die Notation $\underline{\mathbf{y}}^{(1)}, \dots, \underline{\mathbf{y}}^{(k)}$ zur Bezeichnung der Funktionenfamilie **dient nur zur Durchnummerierung und hat nichts mit höheren Ableitungen zu tun**. Da in diesem Kapitel nur Ableitungen erster Ordnung auftreten, kann es zu keinen Verwechslungen kommen. Tiefgestellte Indizes bezeichnen im Folgenden wie üblich die Komponenten einer vektorwertigen Funktion.

Auf die eindeutige Lösbarkeit des Anfangswertproblems (vgl. Satz 14.1.1) wird später in einem allgemeineren Kontext eingegangen werden, für das **lineare homogene System**, d.h.

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} \quad (**)$$

sei an dieser Stelle festgehalten (vgl. Satz 14.2.1):

Satz 16.1.1

Die Funktionen $a_{ij}(x): I \rightarrow \mathbb{R}$, $1 \leq i, j \leq n$ seien stetig auf I . Dann besitzt das *homogene System (**)* auf I *genau n linear unabhängige Lösungen* $\underline{\mathbf{y}}^{(1)}, \underline{\mathbf{y}}^{(2)}, \dots, \underline{\mathbf{y}}^{(n)}$. *Jede Lösung* des homogenen Systems auf I *ist von der Form*

$$\underline{\mathbf{y}}(x) = c_1 \underline{\mathbf{y}}^{(1)}(x) + c_2 \underline{\mathbf{y}}^{(2)}(x) + \dots + c_n \underline{\mathbf{y}}^{(n)}(x)$$

für alle $x \in I$, wobei die c_i , $i = 1, \dots, n$, reelle Konstanten bezeichnen.

Bemerkung. Wieder heißen n linear unabhängige Lösungen ein **Fundamentalsystem** und die **allgemeine Lösung** besteht aus allen Linearkombinationen eines Fundamentalsystems.

Auch das Kriterium zur Prüfung auf lineare Unabhängigkeit von Lösungen ist völlig analog zu Kapitel 14 zu formulieren (vgl. Satz 14.2.2), die anschließenden Bemerkungen behalten wiederum ihre Gültigkeit.

Satz 16.1.2

Es sind n Lösungen $\underline{\mathbf{y}}^{(1)}, \dots, \underline{\mathbf{y}}^{(n)}$ des homogenen Systems (**) auf I **genau dann** linear unabhängig, wenn für (mindestens) ein $x \in I$ die **Wronski-Determinante**

$$D = \begin{vmatrix} \underline{\mathbf{y}}_1^{(1)}(x) & \underline{\mathbf{y}}_1^{(2)}(x) & \dots & \underline{\mathbf{y}}_1^{(n)}(x) \\ \underline{\mathbf{y}}_2^{(1)}(x) & \underline{\mathbf{y}}_2^{(2)}(x) & \dots & \underline{\mathbf{y}}_2^{(n)}(x) \\ \vdots & \vdots & & \vdots \\ \underline{\mathbf{y}}_n^{(1)}(x) & \underline{\mathbf{y}}_n^{(2)}(x) & \dots & \underline{\mathbf{y}}_n^{(n)}(x) \end{vmatrix}$$

von Null verschieden ist.

Bemerkung. Zu Beginn dieses Kapitels wurde eine explizite Differentialgleichung n^{ter} Ordnung auf ein System erster Ordnung zurückgeführt. Man überprüfe, dass dabei Satz 14.2.2 in Satz 16.1.2 übergeht.

Beispiel. Für $x > 0$ und $n = 2$ sei das lineare homogene System

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} \quad \text{mit} \quad A(x) = \begin{pmatrix} -\frac{1}{x(x^2+1)} & \frac{1}{x^2(x^2+1)} \\ \frac{x^2}{x^2+1} & \frac{2x^2+1}{x(x^2+1)} \end{pmatrix}$$

betrachtet, d.h. ausführlich

$$y_1' = -\frac{1}{x(x^2+1)}y_1 + \frac{1}{x^2(x^2+1)}y_2$$

$$y_2' = -\frac{x^2}{x^2+1}y_1 + \frac{2x^2+1}{x(x^2+1)}y_2.$$

Mit

$$\underline{\mathbf{y}}^{(1)}(x) := \begin{pmatrix} 1 \\ x \end{pmatrix}, \quad \underline{\mathbf{y}}^{(2)}(x) := \begin{pmatrix} -1/x \\ x^2 \end{pmatrix}, \quad x > 0,$$

berechnet man leicht

$$\begin{aligned}\underline{\mathbf{y}}^{(1)'}(x) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{x(x^2+1)} & \frac{1}{x^2(x^2+1)} \\ -\frac{x^2}{x^2+1} & \frac{2x^2+1}{x(x^2+1)} \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} \\ &= A(x)\underline{\mathbf{y}}^{(1)}(x)\end{aligned}$$

sowie

$$\begin{aligned}\underline{\mathbf{y}}^{(2)'}(x) &= \begin{pmatrix} 1/x^2 \\ 2x \end{pmatrix} = \begin{pmatrix} -\frac{1}{x(x^2+1)} & \frac{1}{x^2(x^2+1)} \\ -\frac{x^2}{x^2+1} & \frac{2x^2+1}{x(x^2+1)} \end{pmatrix} \begin{pmatrix} -1/x \\ x^2 \end{pmatrix} \\ &= A(x)\underline{\mathbf{y}}^{(2)}(x).\end{aligned}$$

Sowohl $\underline{\mathbf{y}}^{(1)}(x)$ als auch $\underline{\mathbf{y}}^{(2)}(x)$ lösen also für $x > 0$ das betrachtete System, zur Beantwortung der Frage, ob damit ein Fundamentalsystem von Lösungen gefunden ist, d.h. ob die beiden Funktionen auf $I = (0, \infty)$ linear unabhängig sind, wird die Wronski-Determinante berechnet:

$$D = \begin{vmatrix} 1 & -1/x \\ x & x^2 \end{vmatrix} = x^2 + 1 \neq 0,$$

es handelt sich also tatsächlich um ein Fundamentalsystem.

Wie findet man (zumindest prinzipiell) ein Fundamentalsystem?

Ist $A(x)$ auf I stetig (d.h. sind alle Komponentenfunktionen stetig), so fixiere man $x_0 \in I$. Weiter betrachte man die kanonische Basis $(\underline{\mathbf{e}}^{(1)}, \dots, \underline{\mathbf{e}}^{(n)})$ des \mathbb{R}^n und löse die n Anfangswertaufgaben auf I (wie bereits erwähnt, existiert jeweils eine eindeutige Lösung)

$$\begin{aligned}\underline{\mathbf{y}}^{(k)'} &= A(x)\underline{\mathbf{y}}^{(k)}, \\ \underline{\mathbf{y}}^{(k)}(x_0) &= \underline{\mathbf{e}}^{(k)}, \quad k = 1, \dots, n.\end{aligned}$$

Satz 16.1.3

Die so konstruierten Lösungen $\underline{\mathbf{y}}^{(k)}$, $k = 1, \dots, n$, bilden ein Fundamentalsystem von (**), die allgemeine Lösung dieses homogenen Systems lautet

$$\underline{\mathbf{y}}^{(hom)}(x) = \sum_{k=1}^n c_k \underline{\mathbf{y}}^{(k)}(x), \quad c_k \in \mathbb{R}, \quad k = 1, \dots, n.$$

Beweis. Übung. □

Welche **praktische** Hilfe gibt es bei der Suche nach einem Fundamentalsystem?

Zu Beginn von Kapitel 15.4 führt im Spezialfall eines Systems mit konstanten Koeffizienten ein geeigneter Ansatz von einer bekannten Lösung auf eine weitere (unabhängige) Lösung. Ein entsprechender Ansatz, genannt **Reduktion der Ordnung**, wird nun für das System (**) gemacht, in dem die Koeffizienten der Matrix A Funktionen der unabhängigen Variablen x sind.

Dazu nimmt man an, dass eine Lösung $\underline{\mathbf{y}}^{(0)}(x)$ des homogenen Systems bereits bekannt ist und dass $y_1^{(0)}(x) \neq 0$ für alle $x \in I$ (natürlich kann mit jeder anderen Komponente analog argumentiert werden). Um mit Hilfe dieser Information eine weitere linear unabhängige Lösung zu finden, macht man den

Ansatz.

$$\underline{\mathbf{y}}(x) = w(x) \underline{\mathbf{y}}^{(0)}(x) + \underline{\mathbf{z}}(x),$$

wobei die skalare Funktion $w(x)$ und die vektorwertige Funktion $\underline{\mathbf{z}}(x)$ geeignet zu wählen sind.

Bemerkung. Im Fall des Systems mit konstanten Koeffizienten wurde im ersten Schritt $w(x) = x$ und $\underline{\mathbf{z}}(x) = \underline{\mathbf{v}}^{(1)} e^{\lambda x}$ gewählt, wobei λ ein

Eigenwert der nicht diagonalisierbaren Matrix A war. Die bekannte Lösung $\underline{\mathbf{y}}^{(0)}$ war in diesem Fall $\underline{\mathbf{v}}^{(0)}e^{\lambda x}$, wobei $\underline{\mathbf{v}}^{(0)}$ ein Eigenvektor zum Eigenwert λ war.

Wird obiger Ansatz differenziert, so ergibt sich (da $\underline{\mathbf{y}}^{(0)}$ nach Voraussetzung das System löst)

$$\begin{aligned}\underline{\mathbf{y}}'(x) &= w'(x)\underline{\mathbf{y}}^{(0)}(x) + w(x)\underline{\mathbf{y}}^{(0)'}(x) + \underline{\mathbf{z}}'(x) \\ &= w'(x)\underline{\mathbf{y}}^{(0)}(x) + w(x)A(x)\underline{\mathbf{y}}^{(0)}(x) + \underline{\mathbf{z}}'(x),\end{aligned}$$

also

$$\begin{aligned}\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} &\Leftrightarrow \underline{\mathbf{z}}'(x) + w'(x)\underline{\mathbf{y}}^{(0)}(x) + w(x)A(x)\underline{\mathbf{y}}^{(0)}(x) \\ &= A(x)(w(x)\underline{\mathbf{y}}^{(0)}(x) + \underline{\mathbf{z}}(x)) \\ &\Leftrightarrow \underline{\mathbf{z}}'(x) = A(x)\underline{\mathbf{z}}(x) - w'(x)\underline{\mathbf{y}}^{(0)}(x).\end{aligned}$$

Übung. Mit der Wahl $w(x) = x$ und $\underline{\mathbf{z}}(x) = \underline{\mathbf{v}}^{(1)}e^{\lambda x}$, die im Fall von Systemen mit konstanten Koeffizienten möglich war, entspricht die letzte Bedingung genau der Definition eines Hauptvektors erster Stufe.

Obiger Ansatz liefert genau dann eine Lösung des Systems, falls für alle $x \in I$ und für alle $i = 1, \dots, n$ gilt

$$z_i'(x) = \sum_{k=1}^n a_{ik}(x)z_k - w'(x)y_i^{(0)}(x). \quad (1)$$

Die Idee ist es nun, (1) für die ausgezeichnete Komponente $i = 1$ zu betrachten. Wählt man w mit

$$w'(x) = \frac{1}{y_1^{(0)}(x)} \sum_{k=2}^n a_{1k}(x)z_k(x),$$

so wird (1) im Fall $i = 1$ zu

$$z_1' = \sum_{k=1}^n a_{1k}z_k - \left[\frac{1}{y_1^{(0)}} \sum_{k=2}^n a_{1k}z_k \right] y_1^{(0)} = a_{11}z_1,$$

die Komponente $z_1 := 0$ löst die erste Gleichung. Die verbleibenden $(n - 1)$ Gleichungen lauten für $i = 2, \dots, n$ (wegen $z_1 \equiv 0$):

$$\begin{aligned} z'_i &= \sum_{k=2}^n a_{ik}(x)z_k - \left[\frac{1}{y_1^{(0)}(x)} \sum_{k=2}^n a_{1k}(x)z_k \right] y_i^{(0)}(x) \\ &= \sum_{k=2}^n \left[a_{ik}(x) - a_{1k}(x) \frac{y_i^{(0)}(x)}{y_1^{(0)}(x)} \right] z_k \\ &=: \sum_{k=2}^n b_{ik}(x)z_k . \end{aligned}$$

Zusammenfassend ist der Satz gezeigt:

Satz 16.1.4 (*Reduktionsverfahren*)

Mit obiger Notation (insbesondere sei $\underline{y}^{(0)}$ eine bekannte Lösung von $\underline{y}' = A(x)\underline{y}$ mit $y_1^{(0)}(x) \neq 0$ für alle $x \in I$) setze man für $i, k = 2, \dots, n$:

$$b_{ik}(x) := a_{ik}(x) - a_{1k}(x) \frac{y_i^{(0)}(x)}{y_1^{(0)}(x)} .$$

Die Funktion $\underline{z}(x): I \rightarrow \mathbb{R}^{n-1}$ sei eine Lösung von

$$\underline{z}' = B(x)\underline{z}, \quad B = (b_{ik})_{2 \leq i, k \leq n},$$

wobei man schreibe

$$\underline{\mathbf{z}} = \begin{pmatrix} z_2 \\ \vdots \\ z_n \end{pmatrix}$$

und die Notation $B\underline{\mathbf{z}}$ somit offensichtlich ist.

Dann ist $\underline{\mathbf{y}}: I \rightarrow \mathbb{R}^n$,

$$\underline{\mathbf{y}}(x) := w(x)\underline{\mathbf{y}}^{(0)}(x) + \begin{pmatrix} 0 \\ z_2(x) \\ \vdots \\ z_n(x) \end{pmatrix}$$

mit $(x_0 \in I)$

$$w(x) := \int_{x_0}^x \frac{\sum_{k=2}^n a_{1k}(t)z_k(t)}{y_1^{(0)}(t)} dt$$

eine (auf I) von $\underline{\mathbf{y}}^{(0)}$ linear unabhängige Lösung des ursprünglichen Systems.

Beispiel. Es sei $n = 2$, $I = \mathbb{R}$ und das zu untersuchende Differentialgleichungssystem laute

$$\underline{\mathbf{y}}' = \begin{pmatrix} x & -1 \\ 1 & 0 \end{pmatrix} \underline{\mathbf{y}}.$$

Man rechnet leicht nach, dass eine Lösung lautet

$$\underline{\mathbf{y}}^{(0)}(x) = \begin{pmatrix} 1 \\ x \end{pmatrix}.$$

Wie in Satz 16.1.4 wird zur Reduktion der Ordnung gesetzt:

$$b_{22}(x) = a_{22}(x) - a_{12}(x) \frac{y_2^{(0)}(x)}{y_1^{(0)}(x)} = x ,$$

es ist die Differentialgleichung

$$z_2' = x z_2$$

zu untersuchen, als eine Lösung findet man

$$z_2(x) = e^{\frac{x^2}{2}} .$$

Es berechnet sich die Funktion w zu ($x_0 = 0$)

$$w(x) = \int_0^x \frac{a_{12}(t) z_2(t)}{y_1^{(0)}(t)} dt = - \int_0^x e^{\frac{x^2}{2}} dt ,$$

woraus

$$\underline{\mathbf{y}}(x) = - \int_0^x e^{-\frac{t^2}{2}} dt \begin{pmatrix} 1 \\ x \end{pmatrix} + \begin{pmatrix} 0 \\ e^{x^2/2} \end{pmatrix}$$

folgt. Eine Probe (\rightsquigarrow Übung) zeigt, dass in der Tat eine weitere (linear unabhängige) Lösung gefunden wurde.

Bisher wurde eine Frage noch nicht diskutiert:

Wie findet man eine **spezielle Lösung des inhomogenen Systems** (*)?

Der Ansatz ist analog zu Kapitel 14.4 eine **Variation der Konstanten**.

Notation. Ist $\underline{\mathbf{y}}^{(1)}, \dots, \underline{\mathbf{y}}^{(n)}$ ein Fundamentalsystem von Lösungen des homogenen Systems

$$\underline{\mathbf{y}}' = A(x) \underline{\mathbf{y}} ,$$

so nennt man die Matrix bestehend aus diesen Spaltenvektoren,

$$Y := (\underline{\mathbf{y}}^{(1)}(x) \ \underline{\mathbf{y}}^{(2)}(x) \ \dots \ \underline{\mathbf{y}}^{(n)}(x)) = \begin{pmatrix} y_1^{(1)}(x) & \dots & y_1^{(n)}(x) \\ \vdots & & \vdots \\ y_n^{(1)}(x) & \dots & y_n^{(n)}(x) \end{pmatrix}$$

eine **Fundamentalmatrix**. Weiterhin wird aus reellen Konstanten c_1, \dots, c_n der Vektor

$$\underline{\mathbf{c}} := \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$$

gebildet. Man beachte, dass mit dieser Schreibweise gilt

$$M(n, n, \mathbb{R}) \ni Y' = A(x)Y,$$

$$\mathbb{R}^n \ni Y(x)\underline{\mathbf{c}} = c_1 \underline{\mathbf{y}}^{(1)}(x) + \dots + c_n \underline{\mathbf{y}}^{(n)}(x).$$

Der Ansatz für eine spezielle Lösung von

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} + \underline{\mathbf{r}}(x)$$

lautet für eine auf I differenzierbare vektorwertige Funktion $\underline{\mathbf{k}}(x)$

$$\underline{\mathbf{y}}^{(s)}(x) = Y(x)\underline{\mathbf{k}}(x).$$

Differentiation ergibt (da Y Fundamentalmatrix ist)

$$\begin{aligned} \underline{\mathbf{y}}^{(s)'}(x) &= Y'(x)\underline{\mathbf{k}}(x) + Y(x)\underline{\mathbf{k}}'(x) \\ &= A(x)Y(x)\underline{\mathbf{k}}(x) + Y(x)\underline{\mathbf{k}}'(x) \\ &= A(x)\underline{\mathbf{y}}^{(s)}(x) + Y(x)\underline{\mathbf{k}}'(x). \end{aligned}$$

Also löst $\underline{\mathbf{y}}^{(s)}(x)$ genau dann die inhomogene Differentialgleichung, wenn gilt (die Integration ist komponentenweise zu interpretieren)

$$\begin{aligned} Y(x)\underline{\mathbf{k}}'(x) &= \underline{\mathbf{r}}(x), \quad \text{das bedeutet} \\ \underline{\mathbf{k}}(x) &= \underline{\mathbf{c}} + \int_{x_0}^x [Y(t)]^{-1} \underline{\mathbf{r}}(t) dt, \quad \underline{\mathbf{c}} \in \mathbb{R}^n, \quad x_0 \in I, \end{aligned}$$

wobei eine Fundamentalmatrix stets invertierbar ist (warum?). Es ist gezeigt:

Satz 16.1.5

Auf I sei das inhomogene System

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} + \underline{\mathbf{r}}(x)$$

betrachtet, wobei die Daten ($A(x)$ und $\underline{\mathbf{r}}(x)$) stetige Funktionen auf I seien. Es sei $Y(x)$ eine Fundamentalmatrix des zugehörigen homogenen Systems. Dann kann die **allgemeine Lösung des inhomogenen Systems** geschrieben werden als ($x_0 \in I$)

$$\underline{\mathbf{y}}^{(inhom)}(x) = Y(x) \left[\underline{\mathbf{c}} + \int_{x_0}^x [Y(t)]^{-1} \underline{\mathbf{r}}(t) dt \right], \quad \underline{\mathbf{c}} \in \mathbb{R}^n.$$

Bemerkungen.

- i) Es sei daran erinnert, dass die allgemeine Lösung des homogenen Systems geschrieben werden kann als

$$\underline{\mathbf{y}}^{(hom)} = Y(x)\underline{\mathbf{c}}, \quad \underline{\mathbf{c}} \in \mathbb{R}^n.$$

- ii) Eine spezielle Lösung des inhomogenen Systems ist wie oben bewiesen

$$\underline{\mathbf{y}}^{(s)} = Y(x) \int_{x_0}^x [Y(t)]^{-1} \underline{\mathbf{r}}(t) dt.$$

- iii) Wird das Anfangswertproblem mit

$$\underline{\mathbf{y}}(x_0) = \underline{\mathbf{y}}^{(0)}$$

betrachtet, so ist $\underline{\mathbf{c}} = [Y(x_0)]^{-1} \underline{\mathbf{y}}^{(0)}$ zu wählen.

Beispiel. Betrachtet sei für $x > 0$ das System

$$\underline{\mathbf{y}}' = A(x)\underline{\mathbf{y}} + \underline{\mathbf{r}}(x)$$

mit

$$A(x) = \begin{pmatrix} -\frac{1}{x(x^2+1)} & \frac{1}{x^2(x^2+1)} \\ -\frac{x^2}{x^2+1} & \frac{2x^2+1}{x(x^2+1)} \end{pmatrix}, \quad \underline{\mathbf{r}}(x) = \begin{pmatrix} 1/x \\ 1 \end{pmatrix}.$$

Es wurde bereits gezeigt, dass

$$\underline{\mathbf{y}}^{(1)}(x) = \begin{pmatrix} 1 \\ x \end{pmatrix} \quad \text{und} \quad \underline{\mathbf{y}}^{(2)}(x) = \begin{pmatrix} -1/x \\ x^2 \end{pmatrix}$$

linear unabhängige Lösungen des homogenen Systems sind, eine Fundamentalmatrix des homogenen Systems ist

$$Y(x) = \begin{pmatrix} 1 & -1/x \\ x & x^2 \end{pmatrix}.$$

Man findet

$$[Y(x)]^{-1} = \frac{1}{1+x^2} \begin{pmatrix} x^2 & 1/x \\ -x & 1 \end{pmatrix}$$

sowie mit der Wahl $x_0 = 1$

$$\begin{aligned} \int_1^x [Y(t)]^{-1} \underline{\mathbf{r}}(t) dt &= \int_1^x \frac{1}{1+t^2} \begin{pmatrix} t^2 & 1/t \\ -t & 1 \end{pmatrix} \begin{pmatrix} 1/t \\ 1 \end{pmatrix} dt \\ &= \int_1^x \frac{1}{1+t^2} \begin{pmatrix} t + 1/t \\ 0 \end{pmatrix} dt \\ &= \int_1^x \frac{1}{t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} dt \\ &= \begin{pmatrix} \ln(x) \\ 0 \end{pmatrix}. \end{aligned}$$

Eine spezielle Lösung des inhomogenen Systems ist nach Satz 16.1.5

$$\underline{\mathbf{y}}^{(s)} = \begin{pmatrix} 1 & -1/x \\ x & x^2 \end{pmatrix} \begin{pmatrix} \ln(x) \\ 0 \end{pmatrix} = \begin{pmatrix} \ln(x) \\ x \ln(x) \end{pmatrix}.$$

Dies kann wieder mit einer Probe (\rightsquigarrow Übung) verifiziert werden.

16.2 Systeme mit konstanten Koeffizienten

In diesem Abschnitt soll zusammengefasst werden, wie im Fall von System mit **konstanten Koeffizienten** die allgemeine Lösung explizit konstruiert werden kann. Es wird stets das System

$$\underline{\mathbf{y}}' = A\underline{\mathbf{y}}, \quad A \in M(n, n, \mathbb{R}),$$

auf $I = (-\infty, \infty)$ betrachtet. Im Wesentlichen können drei Szenarien auftreten, wobei die Fälle in der Regel gemischt auftreten.

Fall 1. Die Matrix A ist **diagonalisierbar**, wobei alle Eigenwerte und die zugehörige **Basis aus Eigenvektoren reell** sind.

Diese Situation wurde bereits zu Beginn von Kapitel 15, HMI 2, studiert, zur Erinnerung sei an dieser Stelle ein Beispiel erörtert: Es sei

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \in M(3, 3, \mathbb{R}).$$

In Kapitel 15.1, HMI 2, wurde nachgerechnet, dass A die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 1$ hat und dass die zugehörigen (reellen) Eigenräume gegeben sind durch

$$E_{\lambda_1=2} = \{\underline{\mathbf{v}} \in \mathbb{R}^3 : v_3 = 0\}$$

sowie

$$E_{\lambda_2=1} = \{t\underline{\mathbf{w}} : t \in \mathbb{R}\}, \quad \underline{\mathbf{w}} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Mit

$$\underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(3)} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

ist eine Basis aus Eigenvektoren gefunden, die allgemeine Lösung des Systems ist ($c_1, c_2, c_3 \in \mathbb{R}$)

$$\underline{\mathbf{y}}(x) = c_1 \underline{\mathbf{v}}^{(1)} e^{2x} + c_2 \underline{\mathbf{v}}^{(2)} e^{2x} + c_3 \underline{\mathbf{v}}^{(3)} e^x.$$

Fall 2. Die Matrix A ist zwar **diagonalisierbar**, aber **nicht im Reellen**.

Da man aber an reellen Lösungen des Systems interessiert ist, muss ähnlich wie im dritten Beispiel, Kapitel 14.3, HMI 2, argumentiert werden. Mit λ ist auch $\bar{\lambda}$ Eigenwert (Eigenvektoren $\underline{\mathbf{v}}$ und $\bar{\underline{\mathbf{v}}}$, vgl. Satz 15.2.2, HMI 2) und man erhält aus zwei komplexen Lösungen

$$\underline{\mathbf{v}} e^{\lambda x}, \quad \bar{\underline{\mathbf{v}}} e^{\bar{\lambda} x}$$

die reellen (nachrechnen!) Lösungen

$$\underline{\mathbf{y}}^{(1)}(x) = \operatorname{Re} \underline{\mathbf{v}} e^{\lambda x} = \frac{1}{2} (\underline{\mathbf{v}} e^{\lambda x} + \bar{\underline{\mathbf{v}}} e^{\bar{\lambda} x}),$$

$$\underline{\mathbf{y}}^{(2)}(x) = \operatorname{Im} \underline{\mathbf{v}} e^{\lambda x} = \frac{1}{2i} (\underline{\mathbf{v}} e^{\lambda x} - \bar{\underline{\mathbf{v}}} e^{\bar{\lambda} x}).$$

Lautet das System etwa

$$\underline{\mathbf{y}}' = A \underline{\mathbf{y}} = \begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix} \underline{\mathbf{y}},$$

so bestimmt man zunächst als Eigenwerte und Eigenvektoren der Matrix A

$$\lambda_1 = 1 + 2i, \quad \underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 1 \\ -2i \end{pmatrix},$$

$$\lambda_2 = 1 - 2i, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 1 \\ 2i \end{pmatrix}.$$

Rechnet man im Komplexen, so ergibt sich als Fundamentalsystem unmittelbar

$$\underline{\mathbf{z}}^{(1)}(x) = e^{(1+2i)x} \begin{pmatrix} 1 \\ -2i \end{pmatrix},$$

$$\underline{\mathbf{z}}^{(2)}(x) = e^{(1-2i)x} \begin{pmatrix} 1 \\ 2i \end{pmatrix}.$$

Im Reellen betrachtet man

$$\begin{aligned} \underline{\mathbf{y}}^{(1)}(x) &= \frac{1}{2} \left[e^{(1+2i)x} \begin{pmatrix} 1 \\ -2i \end{pmatrix} + e^{(1-2i)x} \begin{pmatrix} 1 \\ 2i \end{pmatrix} \right] \\ &= \frac{1}{2} e^x \left[(\cos(2x) + i \sin(2x)) \begin{pmatrix} 1 \\ -2i \end{pmatrix} \right. \\ &\quad \left. + (\cos(2x) - i \sin(2x)) \begin{pmatrix} 1 \\ 2i \end{pmatrix} \right] \\ &= \frac{1}{2} e^x \begin{pmatrix} 2 \cos(2x) \\ 4 \sin(2x) \end{pmatrix} \\ &= e^x \begin{pmatrix} \cos(2x) \\ 2 \sin(2x) \end{pmatrix}. \end{aligned}$$

Als linear unabhängige reelle Lösung findet man analog

$$\underline{\mathbf{y}}^{(2)}(x) = e^x \begin{pmatrix} \sin(2x) \\ -2 \cos(2x) \end{pmatrix},$$

mit anderen Worten ist die allgemeine (reelle) Lösung

$$\underline{\mathbf{y}}(x) = e^x \begin{pmatrix} c_1 \cos(2x) + c_2 \sin(2x) \\ 2c_1 \sin(2x) - 2c_2 \cos(2x) \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

Fall 3. Die Matrix A ist **nicht diagonalisierbar**.

Dieser Fall sei anhand des einführenden Beispiels aus Abschnitt 15.4 diskutiert, es sei also

$$\underline{\mathbf{y}}' = A\underline{\mathbf{y}} = \begin{pmatrix} 2 & -4 & 2 \\ 0 & -2 & -2 \\ 0 & 8 & 6 \end{pmatrix} \underline{\mathbf{y}}.$$

Als Eigenvektor und Hauptvektoren (zum einzigen Eigenwert $\lambda = 2$) erster und zweiter Stufe (die einer Kettenbedingung genügen) wurden gefunden

$$\underline{\mathbf{v}}^{(0)} = \begin{pmatrix} 8 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(1)} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}, \quad \underline{\mathbf{v}}^{(2)} = \begin{pmatrix} 0 \\ 1/8 \\ 1/4 \end{pmatrix}.$$

Ebenso wurde bereits verifiziert, dass

$$\underline{\mathbf{y}}^{(0)}(x) = \underline{\mathbf{v}}^{(0)} e^{2x}, \quad \underline{\mathbf{y}}^{(1)}(x) = \underline{\mathbf{v}}^{(0)} x e^{2x} + \underline{\mathbf{v}}^{(1)} e^{2x}$$

linear unabhängige Lösungen sind. Mit Hilfe des Hauptvektors zweiter Stufe konstruiert man analog eine dritte linear unabhängige Lösung (Probe!):

$$\underline{\mathbf{y}}^{(2)}(x) = \underline{\mathbf{v}}^{(0)} \frac{1}{2!} x^2 e^{2x} + \underline{\mathbf{v}}^{(1)} \frac{x}{1!} e^{2x} + \underline{\mathbf{v}}^{(2)} e^{2x}.$$

Teil VII

Analysis von Funktionen mehrerer Veränderlicher

Kapitel 17

Differentialrechnung von Funktionen mehrerer Veränderlichen

Vorweg sei betont, dass es sich hier aus verschiedenen Gründen um einen zentralen Teil der Vorlesung handelt:

- i)* Die Beschreibung von realen Vorgängen geschieht in der Regel über die Analysis von Funktionen mehrerer Veränderlicher, eindimensionale Betrachtungsweisen greifen oft zu kurz;
- ii)* die Geometrie von Abbildungen $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ kann lediglich mit sehr beschränkten Hilfsmitteln der Anschauung zugänglich gemacht werden;
- iii)* wichtige Aspekte dieses Teils entbehren eines eindimensionalen Analogons;
- iv)* schließlich gehen die folgenden Betrachtungen weit über den üblicherweise behandelten Schulstoff hinaus.

17.1 Kurven im \mathbb{R}^n (differenzierbare Kurve; Tangentenvektor; reguläre Kurve; singulärer Punkt; Spur; Doppelpunkt; Rektifizierbarkeit; Bogenlänge; Parametertransformation)

Bevor die Analysis von Funktionen $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ im allgemeinen Fall diskutiert wird, soll zunächst der Spezialfall $m = 1$ angesprochen werden, d.h. das Urbild (der Definitionsbereich) der betrachteten Funktionen

ist eine Teilmenge der reellen Achse. Dies ist einerseits eine wesentliche Vereinfachung, andererseits wird hier ein wichtiger geometrischer Begriff vorgestellt, nämlich:

Definition 17.1.1

i) Unter einer **Kurve im \mathbb{R}^n** , $n \geq 2$, versteht man eine **stetige** Abbildung $\gamma: I \rightarrow \mathbb{R}^n$,

$$\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{pmatrix}, \quad t \in I,$$

wobei $I \subset \mathbb{R}$ ein (verallgemeinertes) Intervall bezeichnet.

ii) Eine Kurve γ heißt **differenzierbar** (bzw. stetig differenzierbar), wenn alle Funktionen (Komponenten) γ_k , $1 \leq k \leq n$, differenzierbar (bzw. stetig differenzierbar) sind.

iii) Ist γ differenzierbar, so heißt

$$\gamma'(t) = \begin{pmatrix} \gamma'_1(t) \\ \vdots \\ \gamma'_m(t) \end{pmatrix}, \quad t \in I,$$

Tangentenvektor (oder **Tangentialvektor**) der Kurve γ zum Parameterwert t . Im Fall $\gamma'(t) \neq 0$ heißt $\gamma'(t)/\|\gamma'(t)\|$ **Tangenteneinheitsvektor**.

*iv) Ist $\gamma: I \rightarrow \mathbb{R}^n$ stetig differenzierbar, so heißt die Kurve **regulär**, falls $\gamma'(t) \neq \mathbf{0}$ für alle $t \in I$. Andernfalls heißt ein Parameterwert $t \in I$ mit $\gamma'(t) = \mathbf{0}$ **singulär**.*

Bemerkungen.

- i) Die unabhängige Variable wird hier mit t bezeichnet, da sie oft als Zeit interpretiert wird.*
- ii) Die Lösungen der im letzten Kapitel diskutierten Differentialgleichungssysteme sind Kurven im \mathbb{R}^n .*
- iii) Per definitionem ist eine Kurve eine **Abbildung**, die Menge $\gamma(I) \subset \mathbb{R}^m$ heißt **Spur** der Kurve und ist von der Abbildung selbst zu unterscheiden. **Achtung**: Unterschiedliche Kurven können dieselbe Spur haben.*
- iv) Eine Kurve muss nicht notwendig injektiv sein, es können **Doppelpunkte** existieren (d.h. $\gamma(t_1) = \gamma(t_2)$ für $t_1 \neq t_2 \in I$).*
- v) An “abgeschlossenen Intervallenden” sind die Begriffe Stetigkeit und Differenzierbarkeit “einseitig” zu interpretieren.*

Beispiele.

- i) Die Kurve*

$$\gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi),$$

hat als Spur die Einheitskreislinie im \mathbb{R}^2 . Die gleiche Spur (zweimal durchlaufen) hat die Kurve

$$\tilde{\gamma}(t) = \begin{pmatrix} \cos(2t) \\ \sin(2t) \end{pmatrix}, \quad t \in [0, 2\pi).$$

ii) Es seien $a > 0$, $b \neq 0$ fixiert. Die Kurve

$$\gamma(t) = \begin{pmatrix} a \cos(t) \\ a \sin(t) \\ bt \end{pmatrix}, \quad t \in \mathbb{R},$$

hat als Spur eine **Helix** (Schraubenlinie) mit **Ganghöhe** $2\pi b$ auf dem Zylinder $x_1^2 + x_2^2 = a^2$ (vgl. Abbildung 17.1).

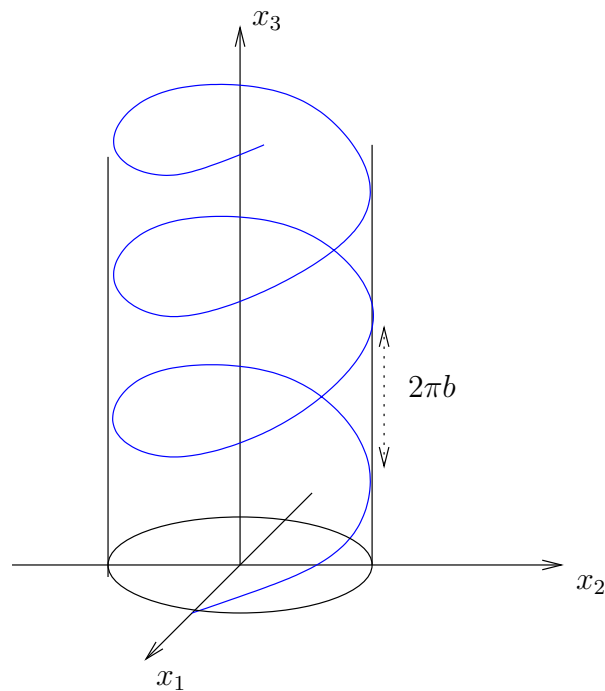


Abbildung 17.1: Eine Helix mit Ganghöhe $2\pi b$.

iii) Der Punkt $0 = \gamma(1) = \gamma(-1)$ ist ein Doppelpunkt der Kurve (vgl. Abbildung 17.2)

$$\gamma(t) = \begin{pmatrix} t^2 - 1 \\ t^3 - t \end{pmatrix}, \quad t \in \mathbb{R}.$$

Trotzdem macht der Begriff ‘Tangentialvektoren zu den Parameterwerten $t = 1$ und $t = -1$ ’ Sinn.

iv) Die **Neilsche Parabel** (vgl. Abbildung 17.3),

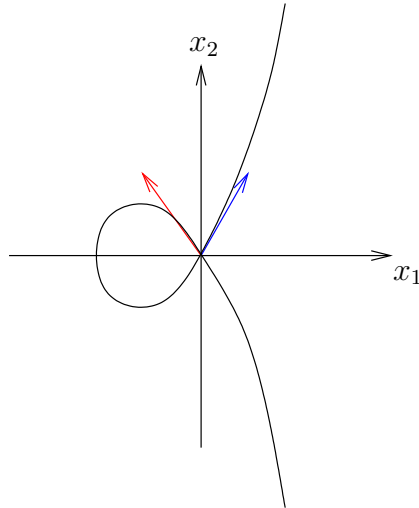


Abbildung 17.2: Eine reguläre Kurve mit Doppelpunkt.

$$\gamma(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}, \quad t \in \mathbb{R},$$

ist im Punkt $t = 0$ singulär.

v) Übung: Man betrachte die Kurven $\gamma, \tilde{\gamma}: \mathbb{R} \rightarrow \mathbb{R}^2$,

$$\gamma(t) = \begin{pmatrix} t \\ t \end{pmatrix}, \quad \tilde{\gamma}(t) = \begin{pmatrix} t^3 \\ t^3 \end{pmatrix},$$

und skizziere deren Spur. Sind die Kurven regulär? (Fazit: Ist eine Kurve “ungeschickt” parametrisiert, so sieht man ihrer Spur singuläre Punkte nicht an.)

Interpretationen. Genau wie zu Beginn von Kapitel 11 bereits besprochen, gibt es zwei Interpretationen eines Tangentialvektors an eine Kurve.

Die Vorstellung bei der [geometrischen Interpretation](#) ist, dass der Vektor $\gamma'(t)$ in die Richtung der [Tangente](#) an die Kurve (im Punkt $\gamma(t)$) zeigt.

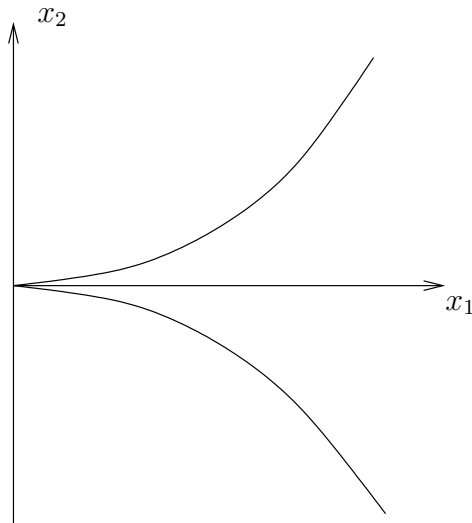


Abbildung 17.3: Die Neilsche Parabel.

Bei der **kinematischen Interpretation einer Kurve** wird der Parameter $t \in I$ als die Zeit aufgefasst, $\gamma(t) \in \mathbb{R}^n$ als Ort. Dann beschreibt die Kurve die **zeitliche Bewegung** eines Punktes im \mathbb{R}^n . Beispielsweise kann der Ort von N Teilchen im \mathbb{R}^3 als Punkt im $\mathbb{R}^{n=3N}$ interpretiert werden. Die Ableitung $\dot{\gamma}(t)$ ist der **Geschwindigkeitsvektor** zum Zeitpunkt t . Hier wird wieder der Unterschied zwischen einer Kurve und ihrer Spur deutlich: Der Kurvenbegriff beinhaltet neben der Spur als Bild der Kurve auch die Information, in welcher Richtung und mit welcher Geschwindigkeit die Spur durchlaufen wird.

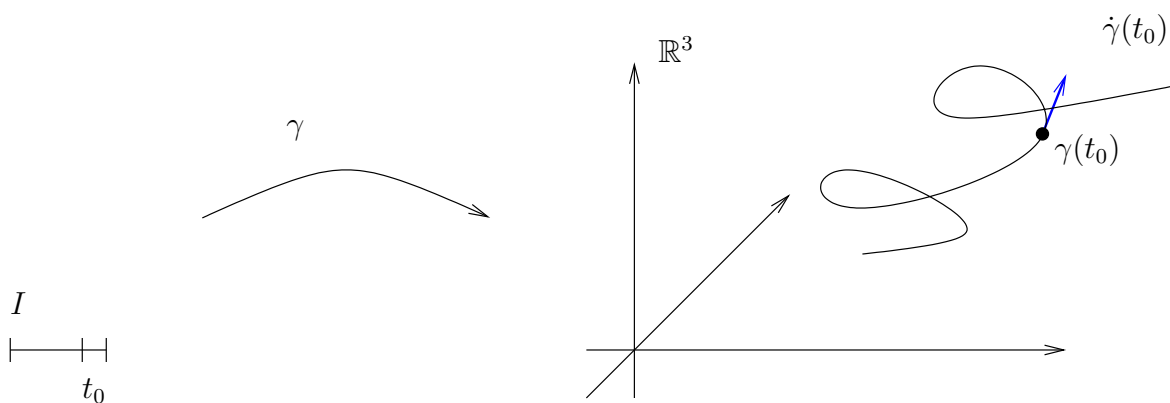


Abbildung 17.4: Zur kinematischen Interpretation eines Tangentialvektors.

Eine wesentliche **geometrische** Frage ist:

Was ist die (Bogen-) Länge einer Kurve?

Bemerkung. Auch wenn man eine intuitive Vorstellung von der Länge einer Kurve hat, so ist dieser Begriff keinesfalls für Kurven definiert, denen nicht elementargeometrisch eine Länge zugeordnet werden kann. Ein analoges Problem gab es zu Beginn von Kapitel 12, in dem der Flächeninhalt über das bestimmte Integral definiert wurde und nicht umgekehrt.

Idee zur Definition der Länge. Man approximiere eine Kurve mit **Polygonzügen**, d.h. Punkte auf der Spur der Kurve werden (affin) linear (mit Strecken) verbunden (vgl. Abbildung 17.5).

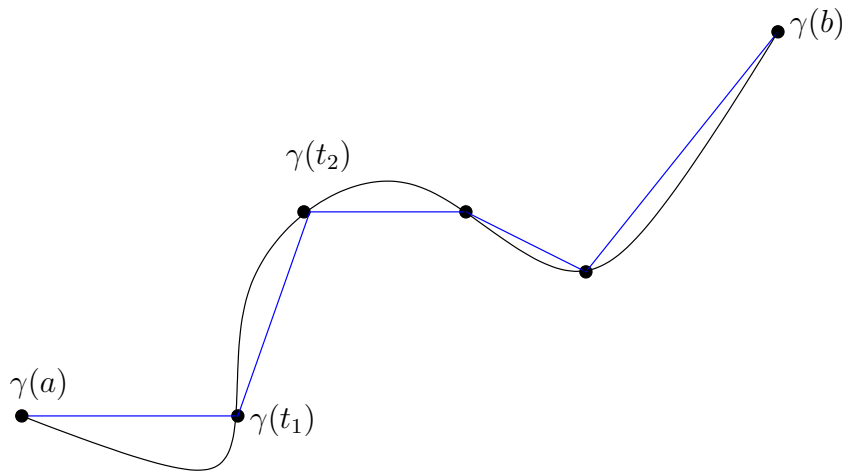


Abbildung 17.5: Zur Rektifizierbarkeit einer Kurve.

Zur präzisen Definition sei in Erinnerung gerufen, dass nach Definition 12.1.1 Eine **Zerlegung** $\mathcal{Z} = \{t_0, t_1, \dots, t_N\}$ eines Intervalls $[a, b]$ eine Menge von Punkten

$$a = t_0 < t_1 < \dots < t_N = b$$

ist. Damit wird definiert

Definition 17.1.2

Es sei $\gamma: [a, b] \rightarrow \mathbb{R}^n$ eine Kurve und für eine beliebige Zerlegung \mathcal{Z} des Intervalls $[a, b]$ sei

$$L(\mathcal{Z}) := \sum_{i=1}^N \|\gamma(t_{i+1}) - \gamma(t_i)\| .$$

Gibt es eine Konstante M , so dass für jede solche Zerlegung $L(\mathcal{Z}) \leq M$ gilt, so heißt γ **rektifizierbar** und

$$L := \sup_{\mathcal{Z}} L(\mathcal{Z})$$

heißt die (**Bogen-**) **Länge** der Kurve.

Bemerkung. Polygonzüge sind rektifizierbar, die Bogenlänge ist genau die elementargeometrisch bekannte (\rightsquigarrow Übung). Ansonsten wäre die Definition auch nicht sinnvoll.

Wie kann man die Bogenlänge einer Kurve **praktisch** berechnen?

Satz 17.1.1

Jede **stetig differenzierbare** Kurve $\gamma: [a, b] \rightarrow \mathbb{R}^n$ ist rektifizierbar, und ihre Länge berechnet sich zu

$$L = \int_a^b \|\gamma'(t)\| dt .$$

Beweisidee. Man approximiere obiges Integral (definitionsgemäß) über Riemannsche Summen. \square

Bemerkungen und Beispiele.

- i)* Wenn die Bogenlänge sinnvoll definiert ist, muss sie etwa auch bei der Kreislinie vom Radius r der elementargeometrisch bekannten entsprechen. Es sei also $r > 0$ fixiert und

$$\begin{aligned}\gamma(t) &= \begin{pmatrix} r \cos(t) \\ r \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi], \quad \text{d.h.} \\ \gamma'(t) &= \begin{pmatrix} -r \sin(t) \\ r \cos(t) \end{pmatrix}.\end{aligned}$$

Wie erwartet folgt

$$L = \int_0^{2\pi} \sqrt{r^2 \sin^2(t) + r^2 \cos^2(t)} \, dt = \int_0^{2\pi} r \, dt = 2\pi r.$$

- ii)* Im \mathbb{R}^2 sei ein **Graph** betrachtet, d.h. eine Kurve der Form

$$\gamma(t) = \begin{pmatrix} t \\ g(t) \end{pmatrix}, \quad t \in [a, b].$$

Ist γ stetig differenzierbar, so gilt

$$\begin{aligned}\gamma'(t) &= \begin{pmatrix} 1 \\ g'(t) \end{pmatrix}, \quad \text{also} \\ L &= \int_a^b \sqrt{1 + (g')^2(t)} \, dt.\end{aligned}$$

- iii)* Die Kreislinie vom Radius r kann auch mit doppelter Geschwindigkeit (oder ebenso mit variabler Geschwindigkeit) durchlaufen werden. D.h. es sei $r > 0$ fixiert und

$$\begin{aligned}\gamma(t) &= \begin{pmatrix} r \cos(2t) \\ r \sin(2t) \end{pmatrix}, \quad t \in [0, \pi], \\ \gamma'(t) &= \begin{pmatrix} -r2 \sin(2t) \\ r2 \cos(2t) \end{pmatrix}.\end{aligned}$$

Nun folgt

$$L = \int_0^\pi \sqrt{r^2 4 \sin^2(t) + r^2 4 \cos^2(t)} dt = \int_0^\pi 2r dt = 2\pi r .$$

Wie erwartet bleibt die Länge invariant.

In Verallgemeinerung des letzten Beispiels stellt sich die Frage nach

Parametertransformationen für Kurven?

Die geometrische Idee ist in Abbildung 17.6 veranschaulicht, präzise wird definiert:

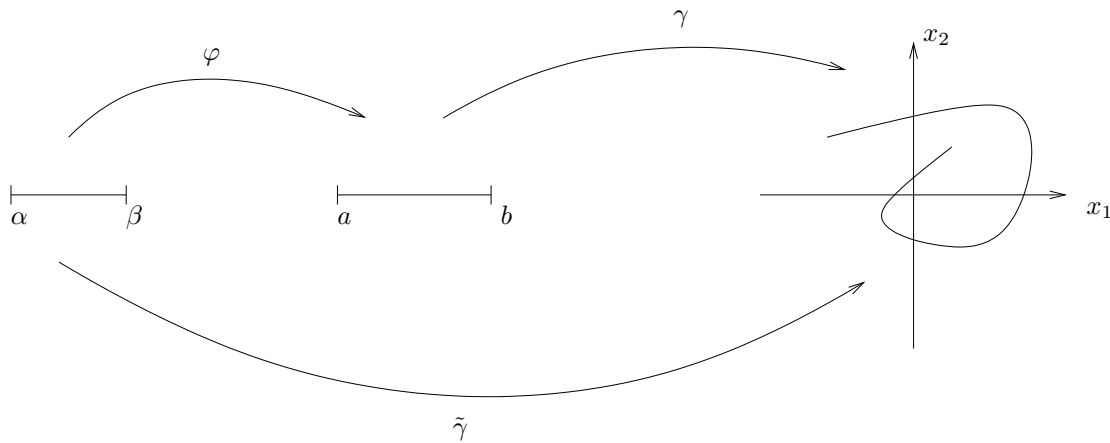


Abbildung 17.6: Eine Parametertransformation.

Definition 17.1.3

Es sei $\gamma: [a, b] \rightarrow \mathbb{R}^n$ eine Kurve, $[\alpha, \beta] \subset \mathbb{R}$ ein weiteres Intervall und $\varphi: [\alpha, \beta] \rightarrow [a, b]$ eine bijektive, stetige Abbildung. Dann ist die zusammengesetzte Abbildung

$$\tilde{\gamma} := \gamma \circ \varphi: [\alpha, \beta] \rightarrow \mathbb{R}^n$$

wieder eine Kurve im \mathbb{R}^n (mit gleicher Spur), $\tilde{\gamma}$ geht aus γ durch die *Parametertransformation* φ hervor. φ heißt *orientierungstreu* (bzw. *orientierungsumkehrend*), wenn die Abbildung *monoton wächst* (bzw. *monoton fällt*). Sind sowohl φ als auch die Umkehrabbildung $\varphi^{-1}: [a, b] \rightarrow [\alpha, \beta]$ stetig differenzierbar, so heißt φ eine *C^1 -Parametertransformation*.

Bemerkung. Die geometrischen Eigenschaften (der Spur) einer Kurve sollten unter Parametertransformation *invariant* bleiben. Beispielsweise sollte die Länge einer Kurve nicht von der speziellen Parametrisierung abhängen (\rightsquigarrow Übungen).

17.2 Ableitungen (partielle Ableitung; (stetig) partiell differenzierbar; die Klasse $C^1(U)$; Richtungsableitung; Nabla-Operator; Gradient; Tangentialebene; Differential; Jacobi-Matrix; Funktionalmatrix; totale Differenzierbarkeit; Polarkoordinaten; krummlinige Koordinaten; Differentialoperatoren (in krummlinigen Koordinaten); Vektorfeld; Divergenz; Rotation; Laplace-Operator; höhere Ableitungen; Vertauschbarkeit von Ableitungen)

Im letzten Paragraphen wurden Kurven im \mathbb{R}^n diskutiert, d.h. Abbildungen $f: \mathbb{R} \rightarrow \mathbb{R}^n$. Im Gegensatz dazu sei jetzt zunächst der *Bildbe-*

reich reell, also

$$f : U \subset \mathbb{R}^m \rightarrow \mathbb{R}, \quad U \ni \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \mapsto f(\underline{\mathbf{x}}) \in \mathbb{R}.$$

In diesem Abschnitt ist stets $U \subset \mathbb{R}^m$, $m \in \mathbb{N}$, eine offene Menge.

Auf den ersten Blick ist es nicht evident, wie der Ableitungsbegriff auf den Fall mehrerer Veränderlicher verallgemeinert werden kann – hier kann kein Differenzenquotient betrachtet werden, da die Division durch eine vektorwertige Größe nicht sinnvoll erklärt werden kann.

Konzept.

- Man variiert die unabhängige Variable nur in eine Koordinatenrichtung $\underline{\mathbf{e}}^{(i)}$ und hält die anderen Koordinaten fest (vgl. Abbildung 17.7).
- Das so entstandene “Teilproblem” kann analog zum Fall einer Funktion einer Variablen behandelt werden.
- Wiederholt man dies sukzessive für jede Koordinatenrichtung, so ist möglicherweise das Verhalten von f als Funktion mehrerer Veränderlicher verstanden.

Bemerkung. Es sei daran erinnert, dass dieses Konzept der **partiellen Differenzierbarkeit** bereits kurz in Definition 12.5.1 angesprochen wurde.

Definition 17.2.1

Eine Abbildung $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ heißt im Punkt $\underline{\mathbf{x}}^{(0)} \in U$ partiell differenzierbar bzgl. der i^{ten} Koordinatenrichtung, falls der Grenzwert

$$\begin{aligned} \frac{\partial f}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) &:= D_i f(\underline{\mathbf{x}}^{(0)}) \\ &:= \lim_{\substack{h \rightarrow 0 \\ h \neq 0, \underline{\mathbf{x}}^{(0)} + h\underline{\mathbf{e}}^{(i)} \in U}} \frac{f(\underline{\mathbf{x}}^{(0)} + h\underline{\mathbf{e}}^{(i)}) - f(\underline{\mathbf{x}}^{(0)})}{h} \end{aligned}$$

existiert, wobei $\underline{\mathbf{e}}^{(i)}$, $1 \leq i \leq m$, den i^{ten} Einheitsvektor der Standardbasis des \mathbb{R}^m bezeichne. $D_i f(\underline{\mathbf{x}}^{(0)})$ heißt die i^{te} partielle Ableitung von f im Punkt $\underline{\mathbf{x}}^{(0)}$.

Bemerkung. Für ein fixiertes $\underline{\mathbf{x}}^{(0)} \in U$ und für ein fixiertes $1 \leq i \leq m$ wird also bei $x_i^{(0)}$ die Funktion einer Veränderlichen betrachtet ($t \in (x_i^{(0)} - \varepsilon, x_i^{(0)} + \varepsilon)$):

$$t \mapsto f^{(i)}(t) := f \begin{pmatrix} x_1^{(0)} \\ \vdots \\ x_{i-1}^{(0)} \\ t \\ x_{i+1}^{(0)} \\ \vdots \\ x_m^{(0)} \end{pmatrix} .$$

Mit dieser Notation ist

$$D_i f(\underline{\mathbf{x}}^{(0)}) = \lim_{h \rightarrow 0, h \neq 0} \frac{f^{(i)}(x_i^{(0)} + h) - f^{(i)}(x_i^{(0)})}{h} = f^{(i)'}(x_i^{(0)}) .$$

Mit anderen Worten: Man leite nach der i^{ten} Variablen ab, die restlichen $n - 1$ Variablen sind fixiert (eingefroren) (vgl. Abbildung 17.7). Deshalb übertragen sich die Rechenregeln für “d/dx” auf die Operatoren D_i .

Von besonderem Interesse sind Funktionen, die für alle $\underline{\mathbf{x}} \in U$ in alle Richtungen partiell differenzierbar sind:

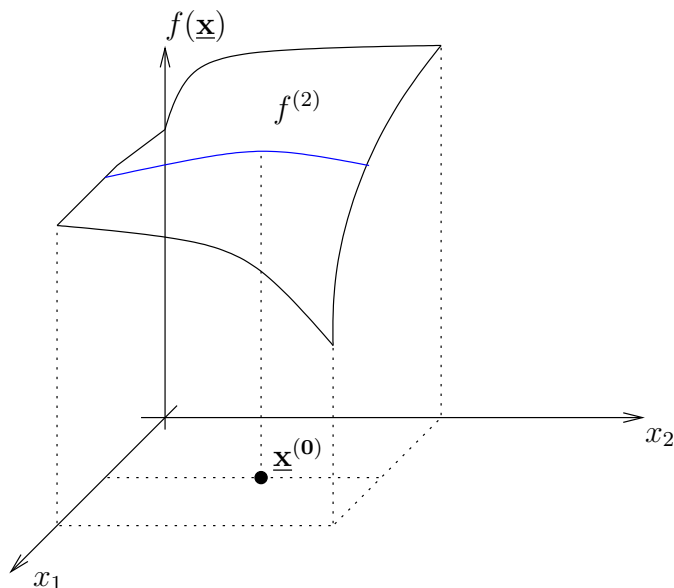


Abbildung 17.7: Zur partiellen Differentiation.

Definition 17.2.2

Eine Funktion $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ heißt *partiell differenzierbar in U* , falls $D_i f(\underline{x})$ für alle Punkte $\underline{x} \in U$ und für alle $i = 1, \dots, m$ existiert. Sind alle partiellen Ableitungen (als Funktionen von U nach \mathbb{R})

$$D_i f : U \rightarrow \mathbb{R}, \quad U \ni \underline{x} \mapsto D_i f(\underline{x})$$

stetig, so heißt f *stetig partiell differenzierbar*. Notation: $f \in C^1(U)$.

Bemerkungen und Beispiele.

i) Man betrachte die Funktion

$$f(\underline{x}) = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} = \|\underline{x}\|, \quad U = \mathbb{R}^n - \{\underline{0}\}.$$

Ist $\underline{x} \neq \underline{0}$ fixiert, so ist

$$t \mapsto \sqrt{x_1^2 + \cdots + t^2 + \cdots + x_n^2}$$

$\uparrow i$

als Funktion einer Variablen differenzierbar, also ist f partiell differenzierbar. Es gilt

$$D_i f(\underline{\mathbf{x}}) = \frac{t}{\sqrt{x_1^2 + \cdots + t^2 + \cdots + x_n^2}} \Big|_{t=x_i} = \frac{x_i}{f(\underline{\mathbf{x}})} = \frac{x_i}{\|\underline{\mathbf{x}}\|}.$$

Die Funktion ist auf U stetig partiell differenzierbar, wie im Fall der Betragsfunktion einer Veränderlichen ist dabei der Nullpunkt aus der Menge U herauszunehmen, in dem die Funktion nicht partiell differenzierbar ist.

ii) Es wurde bereits bei Funktionen einer Veränderlichen aufgezeigt (Definition 11.1.2, *iii*), und anschließende Bemerkungen), dass die Ableitung einer differenzierbaren Funktion nicht stetig sein muss. Entsprechend gilt: **Aus partieller Differenzierbarkeit allein folgt nicht, dass die partielle Ableitungen stetige Funktionen sind.** Deshalb ist der Begriff “stetig partiell differenzierbar” in Definition 17.2.2 geprägt.

iii) Aus partieller Differenzierbarkeit folgt **nicht einmal die Stetigkeit der Funktion selbst** (im Gegensatz zum Fall $n = 1$)! Dies belegt:

Beispiel. (\rightsquigarrow Übungen) Es sei $f: \mathbb{R}^m \rightarrow \mathbb{R}$, $m \geq 2$, definiert durch

$$f(\underline{\mathbf{x}}) = \begin{cases} \frac{x_1 x_2 \cdots x_m}{(x_1^2 + x_2^2 + \cdots + x_m^2)^m} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Dann ist f zwar partiell differenzierbar, f ist im Nullpunkt aber nicht stetig.

Partielle Differenzierbarkeit allein kann also nicht die geeignete Verallgemeinerung des eindimensionalen Ableitungsbegriffes sein.

Im Folgenden wird jedoch **als wesentliche Vereinfachung stets angenommen, dass die Funktion zumindest von der Klasse C^1 ist.** Dann ist auch die Funktion selbst stetig, wie man es von einem sinnvollen Ableitungsbegriff erwartet.

Um nun das Verhalten einer Funktion nahe eines Punktes zu charakterisieren (und nicht nur in fixierte Richtungen), sei vorangestellt:

Erinnerung. Im Fall $m = 1$ ist der Graph von f die Spur der Kurve im \mathbb{R}^2

$$\gamma(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix},$$

der Richtungsvektor der Tangente im Punkt t ist gegeben durch

$$\gamma'(t) = \begin{pmatrix} 1 \\ f'(t) \end{pmatrix}.$$

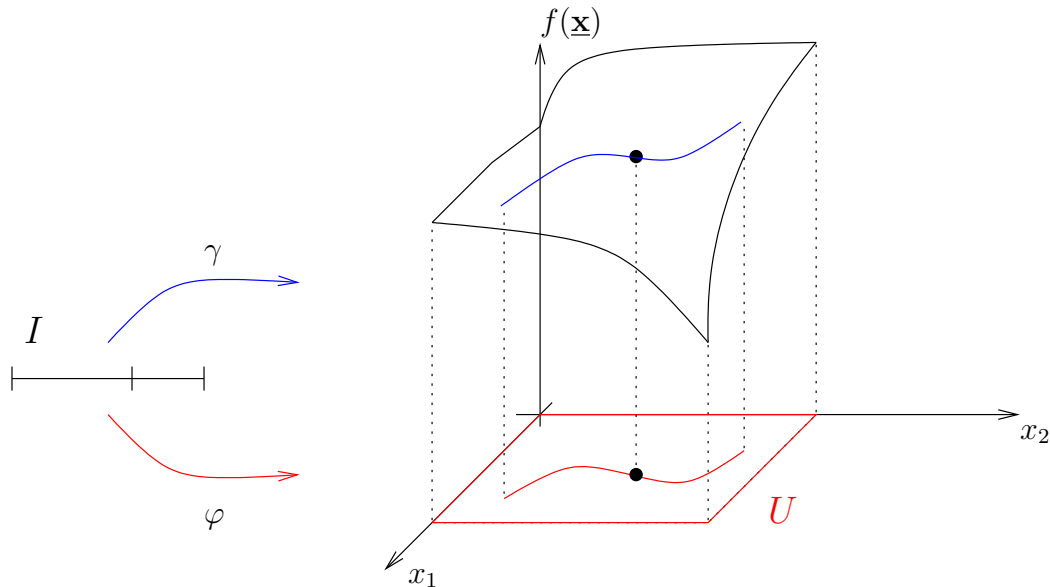
Im Fall $m > 1$ ist der Graph aber nicht mehr die Spur nur einer Kurve.

Idee. Die Geometrie des Graphen $\{(\underline{\mathbf{x}}, f(\underline{\mathbf{x}})) : \underline{\mathbf{x}} \in U\} \subset \mathbb{R}^{m+1}$ soll durch das Studium der Gesamtheit aller **Kurven auf der Fläche** analysiert werden: Dazu sei $f: U \rightarrow \mathbb{R}$ von der Klasse C^1 und φ sei eine C^1 -Abbildung eines (verallgemeinerten) Intervalls I nach U (**eine Kurve im \mathbb{R}^m**). Somit ist

$$\gamma: I \rightarrow \mathbb{R}^{m+1}, \quad \gamma(t) := \begin{pmatrix} \varphi_1(t) \\ \vdots \\ \varphi_m(t) \\ f \circ \varphi(t) \end{pmatrix}$$

eine **Kurve auf dem Graphen von f** (vgl. Abbildung 17.8).

Um die Differenzierbarkeit der Kurve einzusehen, wird eine **Kettenregel** für die letzte Komponente benötigt:


 Abbildung 17.8: Eine Kurve auf dem Graphen von f .

Satz 17.2.1

Es sei $f \in C^1(U)$ und $\varphi: [a, b] \rightarrow \mathbb{R}^m$ eine C^1 -Kurve im \mathbb{R}^m mit $\varphi([a, b]) \subset U$. Dann ist die Komposition $f \circ \varphi: [a, b] \rightarrow \mathbb{R}$ ebenfalls von der Klasse C^1 und es gilt für alle $t_0 \in (a, b)$

$$\frac{d}{dt} f(\varphi(t_0)) = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\varphi(t_0)) \varphi'_i(t_0).$$

Beweisidee.

- Setze $\underline{x}^{(0)} := \varphi(t_0)$, $\underline{\xi} := \varphi(t_0 + h)$.
- Verbinde $\underline{x}^{(0)}$ und $\underline{\xi}$ "längs der Koordinatenachsen".
- Der Mittelwertsatz (Satz 11.2.2) liefert Satz 17.2.1. □

Beispiele.

- i) Es sei f wie oben und $\underline{x}^{(0)} \in U$ fixiert. Man betrachte für $1 \leq i \leq m$ die Funktion $\varphi: (-\varepsilon, \varepsilon) \rightarrow U$, $\varepsilon > 0$ hinreichend klein,

$$\varphi(t) = \underline{x}^{(0)} + t \underline{e}^{(i)}.$$

Dann ist nach Satz 17.2.1

$$\frac{d}{dt} f(\underline{\mathbf{x}}^{(0)} + t\underline{\mathbf{e}}^{(i)})|_{t=0} = D_i f(\underline{\mathbf{x}}^{(0)}) .$$

ii) Es sei wieder f wie oben und $\underline{\mathbf{x}}^{(0)} \in U$ fixiert. Anstelle eines kanonischen Basisvektors kann ebenso ein Vektor $\underline{\mathbf{v}}$ betrachtet werden, der hier auf die Länge eins normiert sei, $\|\underline{\mathbf{v}}\| = 1$. Nun sei $\varphi: (-\varepsilon, \varepsilon) \rightarrow U$, $\varepsilon > 0$ hinreichend klein,

$$\varphi(t) = \underline{\mathbf{x}}^{(0)} + t\underline{\mathbf{v}} .$$

Die **Richtungsableitung** von f im Punkt $\underline{\mathbf{x}}^{(0)}$ in Richtung $\underline{\mathbf{v}}$ ist nach Satz 17.2.1 (vgl. Abbildung 17.9)

$$D_{\underline{\mathbf{v}}} f(\underline{\mathbf{x}}^{(0)}) := \frac{d}{dt} f(\underline{\mathbf{x}}^{(0)} + t\underline{\mathbf{v}})|_{t=0} = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) v_i .$$

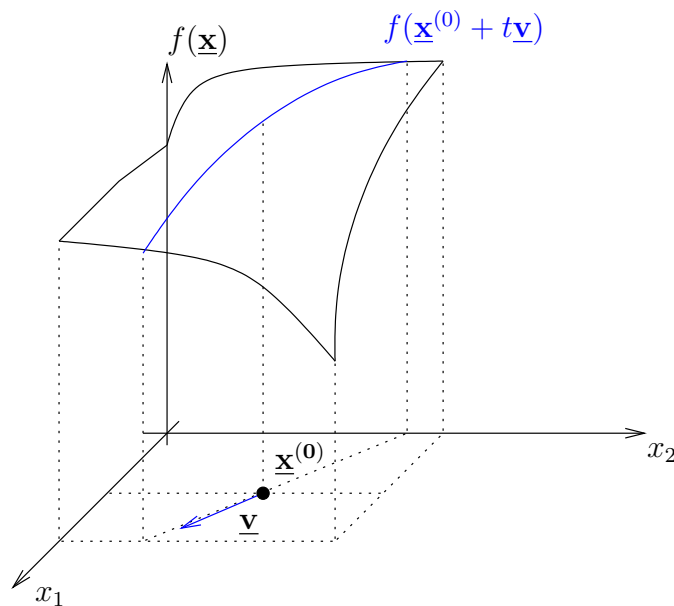


Abbildung 17.9: Zur Richtungsableitung.

Mit Hilfe der Richtungsableitung wird das Verhalten von f bei $\underline{\mathbf{x}}^{(0)}$ in Richtung von $\underline{\mathbf{v}}$ analysiert. Ist $\underline{\mathbf{v}} = \underline{\mathbf{e}}^{(i)}$, so erhält man wieder die i^{te} partielle Ableitung, $i = 1, \dots, m$.

Formal setzt man im nächsten Schritt:

Definition 17.2.3

Es sei $f: \mathbb{R}^m \subset U \rightarrow \mathbb{R}$ von der Klasse C^1 . Formal wird der **Nabla-Operator** definiert als der **Zeilenvektor**

$$\nabla := \left(\frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \cdots \quad \frac{\partial}{\partial x_m} \right).$$

Dann heißt

$$\text{grad} f(\underline{\mathbf{x}}) := \nabla f(\underline{\mathbf{x}}) := \left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_m} \right)(\underline{\mathbf{x}})$$

der **Gradient** von f im Punkt $\underline{\mathbf{x}} \in U$.

Mit dieser Schreibweise ergibt sich für die Richtungsableitung

$$D_{\underline{\mathbf{v}}} f(\underline{\mathbf{x}}^{(0)}) = \nabla f(\underline{\mathbf{x}}^{(0)}) \underline{\mathbf{v}}.$$

Interpretationen des Gradienten.

- i) Betrachtet man in einem fixierten Punkt $\underline{\mathbf{x}}^{(0)}$ die Richtungsableitungen in alle möglichen Richtungen $\underline{\mathbf{v}}$, also den Ausdruck

$$\nabla f(\underline{\mathbf{x}}^{(0)}) \underline{\mathbf{v}}, \quad \|\underline{\mathbf{v}}\| = 1,$$

so erkennt man, dass dieser Ausdruck maximal wird, falls $\underline{\mathbf{v}}$ in Richtung von ∇f (eigentlich des transponierten Vektors) zeigt: Als Spaltenvektor im \mathbb{R}^m zeigt $\nabla f(\underline{\mathbf{x}}^{(0)})$ in die Richtung des stärksten Anstiegs von f im Punkt $\underline{\mathbf{x}}^{(0)}$.

- ii) Zu einer ähnlichen Interpretation gelangt man, wenn die Funktion f mit Hilfe von **Höhenlinien** diskutiert wird (\rightsquigarrow Übungen).

- iii) Es sei der Anschaulichkeit wegen $m = 2$, d.h. $f: \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$ eine Abbildung der Klasse $C^1(U)$. Es sei weiter ein beliebiger Punkt $\underline{\mathbf{x}}^{(0)} \in U$ fixiert, in dessen Nähe die Abbildung f durch das Verhalten von Kurven auf dem Graphen analysiert werden soll. Dazu sei φ eine beliebige C^1 -Kurve in U , $\varphi: (-\varepsilon, \varepsilon) \rightarrow U$, $\varepsilon > 0$

hinreichend klein, mit $\varphi(0) = \underline{\mathbf{x}}^{(0)}$. Die Kurve γ ,

$$\gamma: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^3, \quad \gamma(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ f \circ \varphi(t) \end{pmatrix},$$

auf dem Graphen von f hat im Punkt $t = 0$ nach Satz 17.2.1 als Tangentenvektor

$$\begin{aligned} \gamma'(0) &= \begin{pmatrix} \varphi_1'(0) \\ \varphi_2'(0) \\ \nabla f(\underline{\mathbf{x}}^{(0)}) \begin{pmatrix} \varphi_1'(0) \\ \varphi_2'(0) \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} \varphi_1'(0) \\ \varphi_2'(0) \\ \frac{\partial f}{\partial x_1}(\underline{\mathbf{x}}^{(0)})\varphi_1'(0) + \frac{\partial f}{\partial x_2}(\underline{\mathbf{x}}^{(0)})\varphi_2'(0) \end{pmatrix} \end{aligned}$$

also

$$\gamma'(0) = \varphi_1'(0) \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x_1}(\underline{\mathbf{x}}^{(0)}) \end{pmatrix} + \varphi_2'(0) \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial x_2}(\underline{\mathbf{x}}^{(0)}) \end{pmatrix}.$$

Setzt man schließlich

$$\underline{\mathbf{u}} := \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x_1}(\underline{\mathbf{x}}^{(0)}) \end{pmatrix}, \quad \underline{\mathbf{v}} := \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial x_2}(\underline{\mathbf{x}}^{(0)}) \end{pmatrix},$$

wobei diese beiden Vektoren **nur von der Abbildung f und dem Punkt $\underline{\mathbf{x}}^{(0)}$ abhängen**, so erkennt man, dass **Tangentenvektoren an beliebige Kurven auf dem Graphen von f** in der sogenannten **Tangentialebene**

$$T_{\underline{\mathbf{x}}^{(0)}}f := \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = \lambda \underline{\mathbf{u}} + \mu \underline{\mathbf{v}}, \lambda, \mu \in \mathbb{R}\}$$

liegen, die von $\underline{\mathbf{u}}$ und $\underline{\mathbf{v}}$ aufgespannt wird. Anschaulich stellt man sich diese Ebene im Punkt $\begin{pmatrix} \underline{\mathbf{x}}^{(0)} \\ f(\underline{\mathbf{x}}^{(0)}) \end{pmatrix} \in \mathbb{R}^3$ “angeheftet” vor (in vollständiger analogie zur Tangente an die Kurve im Fall $m = 1$). Die Situation ist in Abbildung 17.10 skizziert.

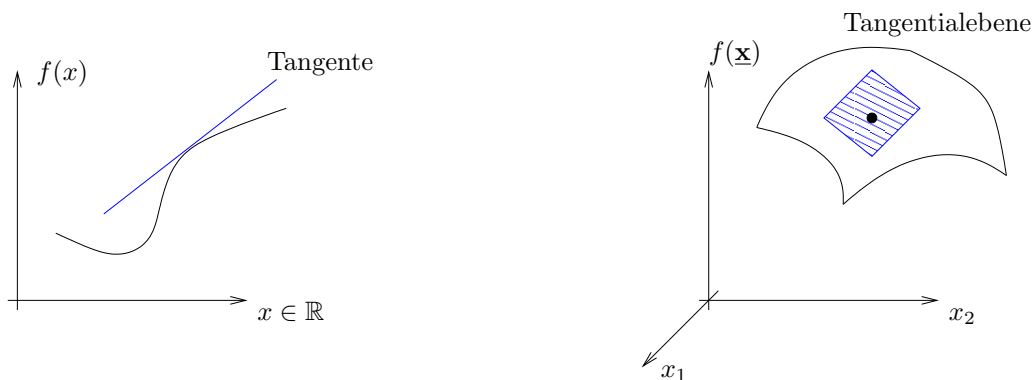


Abbildung 17.10: Zur Tangentialebene an den Graphen von f .

Beispiel. Man betrachte die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(\underline{\mathbf{x}}) = c + x_1^2 + x_2^2,$$

wobei $c > 0$ eine positive Konstante bezeichne. Es handelt sich hier um ein **Paraboloid**, das in Abbildung 17.11 skizziert ist.

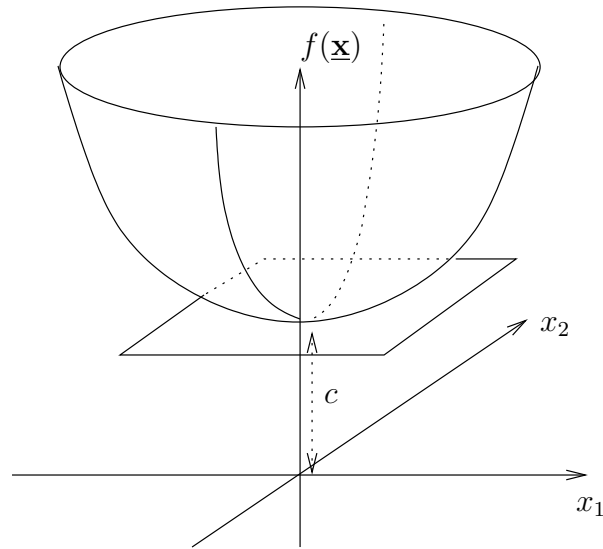


Abbildung 17.11: Ein Paraboloid.

Mit obiger Notation ist

$$\underline{\mathbf{u}} = \begin{pmatrix} 1 \\ 0 \\ 2x_1 \end{pmatrix}, \quad \underline{\mathbf{v}} = \begin{pmatrix} 0 \\ 1 \\ 2x_2 \end{pmatrix},$$

im Punkt $\underline{\mathbf{0}}$ wird die Tangentialebene $T_{\underline{\mathbf{0}}}f$ aufgespannt von den Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

$T_{\underline{\mathbf{0}}}f$ ist die $\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}$ -Ebene im \mathbb{R}^3 .

iv) In Kapitel 11 (siehe auch Kapitel 13), wurde die Ableitung einer Funktion einer Veränderlichen als Approximation erster Ordnung interpretiert, d.h.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \dots,$$

wobei “...” für Terme höherer Ordnung steht, die in der Nähe eines fixierten Punktes x_0 im besten Fall “vernachlässigt” werden

können. Es handelt sich hier um eine (affin) lineare Approximation.

Soll analog eine Funktion $f: \mathbb{R}^m \rightarrow \mathbb{R}$ (affin) linear approximiert werden, so sei daran erinnert (Kapitel 9), dass eine lineare Abbildung $L: \mathbb{R}^m \rightarrow \mathbb{R}$ durch einen Zeilenvektor

$$\underline{\mathbf{a}}^T = (a_1 \ a_2 \ \dots \ a_m)$$

dargestellt wird, wobei die Darstellung hier immer bezüglich der kanonischen Basis des \mathbb{R}^m erfolge. Für “hinreichend glatte” Funktionen (und das sind beileibe nicht alle) $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ ist demnach als Analogon zum eindimensionalen Fall zu einem fixierten Punkt $\underline{\mathbf{x}}^{(0)} \in U$ ein Zeilenvektor $\underline{\mathbf{a}}^T$ (abhängig von f und von $\underline{\mathbf{x}}^{(0)}$) gesucht, sodass “in der Nähe von $\underline{\mathbf{x}}^{(0)}$ ”

$$f(\underline{\mathbf{x}}) = f(\underline{\mathbf{x}}^{(0)}) + \underline{\mathbf{a}}^T (\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}) + \dots, \quad (*)$$

wobei “...” wie oben zu interpretieren ist. Ist (*) richtig, so gilt für h hinreichend klein und $1 \leq i \leq m$:

$$f(\underline{\mathbf{x}}^{(0)} + h\underline{\mathbf{e}}^{(i)}) = f(\underline{\mathbf{x}}^{(0)}) + \underline{\mathbf{a}}^T h\underline{\mathbf{e}}^{(i)} + \dots,$$

wobei $\underline{\mathbf{e}}^{(i)}$ wie üblich den i^{ten} Einheitsvektor der kanonischen Basis des \mathbb{R}^m bezeichne. Umgeformt bedeutet dies

$$\frac{f(\underline{\mathbf{x}}^{(0)} + h\underline{\mathbf{e}}^{(i)}) - f(\underline{\mathbf{x}}^{(0)})}{h} = \underline{\mathbf{a}}^T \underline{\mathbf{e}}^{(i)} + \dots = a_i + \dots, \quad (**)$$

im Grenzwert $h \rightarrow 0$ sieht man (nach Definition 17.2.1 konvergiert die linke Seite von (**)) gegen $D_i f(\underline{\mathbf{x}}^{(0)})$, dass für “hinreichend glatte” Funktionen $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ “in der Nähe” eines jeden Punktes $\underline{\mathbf{x}}^{(0)} \in U$ gilt

$$f(\underline{\mathbf{x}}) = f(\underline{\mathbf{x}}^{(0)}) + \nabla f(\underline{\mathbf{x}}^{(0)}) (\underline{\mathbf{x}} - \underline{\mathbf{x}}^{(0)}) + \dots$$

Der Gradient wird in diesem Sinne als (affin) lineare Approximation einer Funktion f bei $\underline{\mathbf{x}}^{(0)}$ interpretiert.

Um die letzte Bemerkung für eine vektorwertige Funktion mehrerer Veränderlicher $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ zusammenzufassen, wird zunächst definiert:

Definition 17.2.4¹

Es sei $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$ von der Klasse $C^1(U; \mathbb{R}^n)$, d.h. *alle* partiellen Ableitungen *aller* Komponentenfunktionen existieren in *jedem* Punkt aus U und seien *stetige* Funktionen auf U . Dann heißt die Matrix $A = A(\underline{\mathbf{x}}^{(0)})$, bestehend aus allen partiellen Ableitungen in einem Punkt $\underline{\mathbf{x}}^{(0)} \in U$,

$$A := Df(\underline{\mathbf{x}}^{(0)}) := J_f(\underline{\mathbf{x}}^{(0)}) := \left(\frac{\partial f_i}{\partial x_j} \right)_{\substack{j=1, \dots, m \\ i=1, \dots, n}} \in M(n, m, \mathbb{R})$$

das *Differential* oder die *Jacobi-Matrix* oder die *Funktionalmatrix* von f in $\underline{\mathbf{x}}^{(0)}$.

Beispiele.

i) Im Fall $n = 1$ ist $A(\underline{\mathbf{x}}^{(0)})$ der Gradient von f im Punkt $\underline{\mathbf{x}}^{(0)}$.

ii) Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$,

$$f(\underline{\mathbf{x}}) = \begin{pmatrix} x_1^2 + x_2 \\ x_1 x_2 \\ e^{x_1} \end{pmatrix}.$$

Dann ist f von der Klasse $C^1(\mathbb{R}^2; \mathbb{R}^3)$ (Übung!) und für jedes $\underline{\mathbf{x}} \in \mathbb{R}^2$ ist

$$A = A(\underline{\mathbf{x}}) = \begin{pmatrix} 2x_1 & 1 \\ x_2 & x_1 \\ e^{x_1} & 0 \end{pmatrix}.$$

Mit obiger Notation gilt

¹C.G. Jacobi, 1804–1851; Königsberg, Berlin.

Satz 17.2.2

Es sei $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$, von der Klasse $C^1(U; \mathbb{R}^n)$. Dann ist f im jedem Punkt $\underline{\mathbf{x}}^{(0)} \in U$ (total) differenzierbar (m.a.W. in U (total) differenzierbar), d.h. es gibt eine lineare Abbildung $A: \mathbb{R}^m \rightarrow \mathbb{R}^n$, sodass in einer Umgebung von $\underline{\mathbf{x}}^{(0)}$ gilt:

$$f(\underline{\mathbf{x}}^{(0)} + \underline{\xi}) = f(\underline{\mathbf{x}}^{(0)}) + A\underline{\xi} + \varphi(\underline{\xi}).$$

Hierbei ist φ in einer Umgebung der Null definiert mit Werten im \mathbb{R}^n , es gilt:

$$\lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}} \frac{\varphi(\underline{\xi})}{\|\underline{\xi}\|} = \underline{\mathbf{0}}.$$

Die lineare Abbildung A (ihre Matrixdarstellung bzgl. der kanonischen Basen) ist *eindeutig bestimmt*. Sie ist *genau die Jacobi-Matrix* im Punkt $\underline{\mathbf{x}}^{(0)}$.

Beweisidee. Die der letzten Behauptung zugrunde liegende Idee wurde bereits als Interpretation des Gradienten im Fall $n = 1$ skizziert. \square

Bemerkung. In Satz 17.2.2 wird der Begriff der (totalen) Differenzierbarkeit definiert. Es gelten die Implikationen

$$\begin{aligned} \text{stetig partiell differenzierbar} &\Rightarrow \text{(total) differenzierbar} \\ &\Rightarrow \text{partiell differenzierbar.} \end{aligned}$$

Die **Umkehrungen** sind im Allgemeinen **falsch**.

Beispiele.

i) Es sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\underline{\mathbf{x}}) = \|\underline{\mathbf{x}}\|^2.$$

Es ist f von der Klasse $C^1(U)$, also nach Satz 17.2.2 (total) differenzierbar. Dies kann ebenso über die Definition verifiziert werden: Es gilt in jedem Punkt $\underline{\mathbf{x}}^{(0)} \in \mathbb{R}^m$:

$$\begin{aligned} f(\underline{\mathbf{x}}^{(0)} + \underline{\xi}) - f(\underline{\mathbf{x}}^{(0)}) &= \sum_{i=1}^m (x_i^{(0)} + \xi_i)^2 - \sum_{i=1}^m x_i^{(0)2} \\ &= \sum_{i=1}^m (2x_i^{(0)}\xi_i + \xi_i^2) \\ &= 2(x_1^{(0)} \dots x_m^{(0)})\underline{\xi} + \|\underline{\xi}\|^2. \end{aligned}$$

Ein Vergleich mit der Definition der (totalen) Differenzierbarkeit zeigt, dass f in jedem $\underline{\mathbf{x}}^{(0)} \in \mathbb{R}^m$ differenzierbar ist mit

$$A = A(\underline{\mathbf{x}}^{(0)}) = 2(x_1^{(0)}, \dots, x_m^{(0)}).$$

Wie in Satz 17.2.2 behauptet, ist A genau der Gradient von f im Punkt $\underline{\mathbf{x}}^{(0)}$.

ii) Es sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$,

$$f(\underline{\mathbf{x}}) = \|\underline{\mathbf{x}}\|.$$

Auf $\mathbb{R}^n - \{\underline{\mathbf{0}}\}$ ist f von der Klasse C^1 , also (total) differenzierbar. Im Punkt $\{\underline{\mathbf{0}}\}$ beobachtet man zunächst:

$$f(\underline{\mathbf{0}} + \underline{\xi}) - f(\underline{\mathbf{0}}) = \|\underline{\xi}\| = \left[\sum_{i=1}^n \xi_i^2 \right]^{\frac{1}{2}}$$

Wäre f in $\underline{\mathbf{x}}^{(0)}$ (total) differenzierbar, so hätte man nach der Definition

$$\lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}} \frac{f(\underline{\mathbf{0}} + \underline{\xi}) - f(\underline{\mathbf{0}})}{\|\underline{\xi}\|} = 1 = \lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}} \frac{A\underline{\xi}}{\|\underline{\xi}\|}.$$

Der Widerspruch folgt aus

$$\lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}} \frac{A\underline{\xi}}{\|\underline{\xi}\|} = - \lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}} \frac{A(-\underline{\xi})}{\|-\underline{\xi}\|},$$

also ist f in $\underline{\mathbf{0}}$ nicht (total) differenzierbar, wie es von der Diskussion der Betragsfunktion in einer Veränderlichen zu erwarten war.

Eine Produktregel für die Ableitungen von Funktionen mehrerer Veränderlicher wird in den Übungen diskutiert, es bleibt die Frage:

Gilt im allgemeinen Fall eine **Kettenregel**?

Satz 17.2.3

Es seien $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^k$ offen, $g: U \rightarrow \mathbb{R}^k$, $f: V \rightarrow \mathbb{R}^n$, es gelte $g(U) \subset V$. Die Abbildungen f, g seien von der Klasse C^1 . Dann ist auch die Abbildung $f \circ g: \mathbb{R}^m \supset U \rightarrow \mathbb{R}^n$ von der Klasse C^1 und in jedem Punkt $\underline{\mathbf{x}}^{(0)} \in U$ gilt

$$D(f \circ g)(\underline{\mathbf{x}}^{(0)}) = Df(g(\underline{\mathbf{x}}^{(0)}))Dg(\underline{\mathbf{x}}^{(0)}) .$$

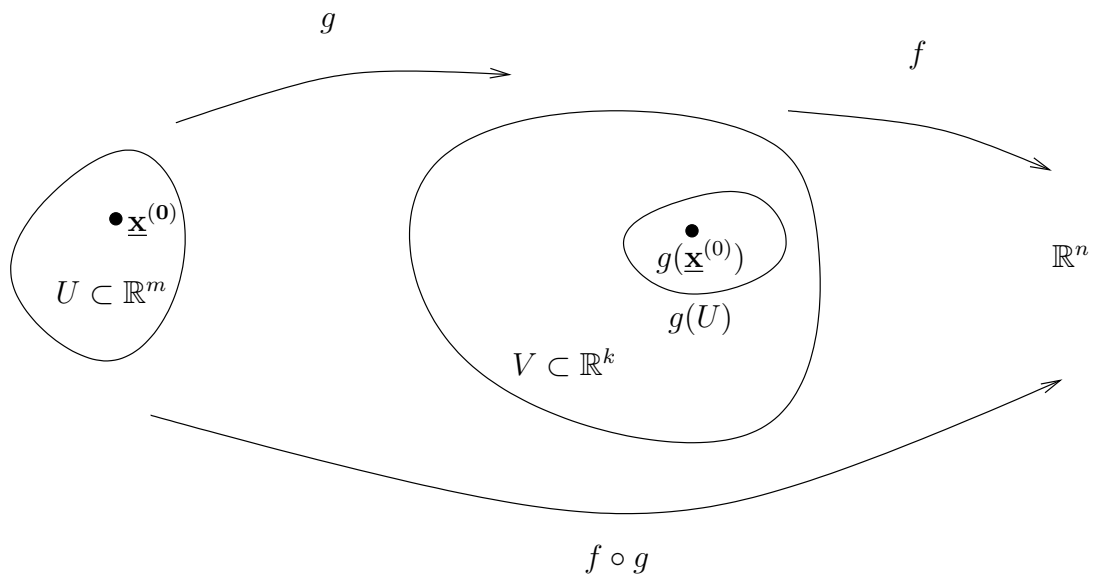


Abbildung 17.12: Zur Kettenregel.

Beispiele.

i) Im Spezialfall f reellwertig ($n = 1$) gilt ($h := f \circ g$)

$$\begin{aligned} Dh(\underline{\mathbf{x}}^{(0)}) &= Df(g(\underline{\mathbf{x}}^{(0)}))Dg(\underline{\mathbf{x}}^{(0)}) \\ &= \left(\frac{\partial f}{\partial y_1}(g(\underline{\mathbf{x}}^{(0)})) \dots \frac{\partial f}{\partial y_k}(g(\underline{\mathbf{x}}^{(0)})) \right) \\ &\quad \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\underline{\mathbf{x}}^{(0)}) & \dots & \frac{\partial g_1}{\partial x_m}(\underline{\mathbf{x}}^{(0)}) \\ \vdots & & \vdots \\ \frac{\partial g_k}{\partial x_1}(\underline{\mathbf{x}}^{(0)}) & \dots & \frac{\partial g_k}{\partial x_m}(\underline{\mathbf{x}}^{(0)}) \end{pmatrix}. \end{aligned}$$

Insbesondere ist ($i = 1, \dots, m$)

$$\frac{\partial h}{\partial x_i}(\underline{\mathbf{x}}^{(0)}) = \sum_{j=1}^k \frac{\partial f}{\partial y_j}(g(\underline{\mathbf{x}}^{(0)})) \frac{\partial g_j}{\partial x_i}(\underline{\mathbf{x}}^{(0)}).$$

ii) Es sei (Polarkoordinaten)

$$\begin{aligned} U &= \left\{ \begin{pmatrix} r \\ \varphi \end{pmatrix} : 0 < r, 0 < \varphi < 2\pi \right\}, \\ g : \begin{pmatrix} r \\ \varphi \end{pmatrix} &\mapsto \underline{\mathbf{x}} = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}, \\ f : \mathbb{R}^2 &\rightarrow \mathbb{R}. \end{aligned}$$

Dann gilt für $h = f \circ g$:

$$\frac{\partial h}{\partial r} = \frac{\partial f}{\partial x_1} \cos(\varphi) + \frac{\partial f}{\partial x_2} \sin(\varphi), \quad (1)$$

$$\frac{\partial h}{\partial \varphi} = \frac{\partial f}{\partial x_1} (-r \sin(\varphi)) + \frac{\partial f}{\partial x_2} r \cos(\varphi). \quad (2)$$

Multipliziert man die Gleichung (1) mit $r \sin(\varphi)$, Gleichung (2) mit $\cos(\varphi)$ und addiert sie anschließend, so folgt

$$\frac{\partial}{\partial x_2} f = \sin(\varphi) \frac{\partial}{\partial r} h + \frac{\cos(\varphi)}{r} \frac{\partial}{\partial \varphi} h. \quad (3)$$

Wird die Gleichung (3) wiederum in (1) eingesetzt, so folgt

$$\frac{\partial}{\partial x_1} f = \cos(\varphi) \frac{\partial}{\partial r} h - \frac{\sin(\varphi)}{r} \frac{\partial}{\partial \varphi} h. \quad (4)$$

Die Beziehungen (3) und (4) werden **symbolisch** geschrieben als

$$\frac{\partial}{\partial x_1} = \cos(\varphi) \frac{\partial}{\partial r} - \frac{\sin(\varphi)}{r} \frac{\partial}{\partial \varphi}, \quad (5)$$

$$\frac{\partial}{\partial x_2} = \sin(\varphi) \frac{\partial}{\partial r} + \frac{\cos(\varphi)}{r} \frac{\partial}{\partial \varphi}, \quad (6)$$

es handelt sich um den Nabla-Operator in Polarkoordinaten. Analoge Ausdrücke können auch für andere **krummlinige Koordinaten** hergeleitet werden.

In den physikalischen Anwendungen spielen neben dem Gradienten weitere spezielle **Differentialoperatoren** eine herausragende Rolle:

Definition 17.2.5

Es seien alle im Folgenden auftretenden Ableitungen stetige Funktionen. Ein **Vektorfeld** auf U ist eine Abbildung

$$F : \mathbb{R}^m \supset U \rightarrow \mathbb{R}^m ,$$

die also jedem $\underline{\mathbf{x}} \in U$ einen Vektor $\underline{\mathbf{F}}(\underline{\mathbf{x}}) \in \mathbb{R}^m$ zuordnet (Beispiel: $\underline{\mathbf{x}} \mapsto \nabla f(\underline{\mathbf{x}})$ als Spaltenvektor interpretiert).

i) Es heißt

$$\operatorname{div} F(\underline{\mathbf{x}}) := \langle \nabla^T, F \rangle(\underline{\mathbf{x}}) := \sum_{i=1}^m \frac{\partial F_i}{\partial x_i}(\underline{\mathbf{x}}) \in \mathbb{R}$$

die **Divergenz** des Vektorfeldes F im Punkt $\underline{\mathbf{x}} \in U$.

ii) Ist speziell $m = 3$ und, so heißt

$$\operatorname{rot} F(\underline{\mathbf{x}}) := (\nabla^T \times F)(\underline{\mathbf{x}}) := \begin{pmatrix} \frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \\ \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \end{pmatrix}(\underline{\mathbf{x}}) \in \mathbb{R}^3$$

die **Rotation** des Vektorfeldes v im Punkt $\underline{\mathbf{x}} \in U$.

iii) Man setzt

$$\Delta f(\underline{\mathbf{x}}) := \operatorname{div} \operatorname{grad} f(\underline{\mathbf{x}}) = \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_i} \right) (\underline{\mathbf{x}}) ,$$

wobei Δ **Laplace-Operator** heißt.

Bemerkungen. (werden z.T. später deutlich)

- i) Die Divergenz ist ein Maß für “**Quellen und Senken**” eines Vektorfeldes (Fluid-Mechanik: “div Geschwindigkeitsfeld = 0” bedeutet **Massenerhaltung**).
- ii) Die Rotation gibt Informationen über mögliche “**Wirbel**” eines Vektorfeldes.
- iii) Die **partielle Differentialgleichung** $\Delta f = 0$ heißt die **Potentialgleichung**, die Lösungen sind **harmonische Funktionen**.
- iv) Für Beziehungen wie **rot grad = 0** etc. sei auf die Übungen verwiesen.

Beispiele.

- i) Es sei $F(\underline{\mathbf{x}}) = \underline{\mathbf{x}}$. Dann gilt

$$\operatorname{div} F(\underline{\mathbf{x}}) = \sum_{i=1}^m \frac{\partial F_i}{\partial x_i} = \sum_{i=1}^m \frac{\partial x_i}{\partial x_i} = m.$$

- ii) Es sei $f(\underline{\mathbf{x}}) = \left(\sum_{i=1}^m x_i^2 \right)^{1/2}$, $U = \mathbb{R}^m - \underline{\mathbf{0}}$. Es gilt (∇f als Spaltenvektor interpretiert)

$$\nabla f(\underline{\mathbf{x}}) = \frac{\underline{\mathbf{x}}}{f(\underline{\mathbf{x}})} \in \mathbb{R}^m;$$

$$\begin{aligned} \Delta f(\underline{\mathbf{x}}) &= \operatorname{div} \operatorname{grad} f(\underline{\mathbf{x}}) \\ &= \sum_{i=1}^m \frac{\partial}{\partial x_i} \frac{x_i}{f(\underline{\mathbf{x}})} \\ &= \sum_{i=1}^m \left[\frac{1}{f(\underline{\mathbf{x}})} - \frac{x_i}{f^2(\underline{\mathbf{x}})} \frac{x_i}{f(\underline{\mathbf{x}})} \right] \\ &= \left[\sum_{i=1}^m \frac{1}{f(\underline{\mathbf{x}})} \right] - \frac{1}{f(\underline{\mathbf{x}})} \\ &= \frac{m-1}{f(\underline{\mathbf{x}})}. \end{aligned}$$

Oben sind bereits des Öfteren die partiellen Ableitungen $D_i f$ selbst als Abbildungen $U \rightarrow \mathbb{R}^n$ aufgefasst worden. So können auch **höhere Ableitungen** eingeführt werden, wobei hier der Einfachheit wegen der skalare Fall $n = 1$ betrachtet wird:

Definition 17.2.6

*Es sei $f: U \rightarrow \mathbb{R}$ partiell differenzierbar. Sind alle partiellen Ableitungen $D_i f: U \rightarrow \mathbb{R}$, $i = 1, \dots, m$ selbst wieder partiell differenzierbar, so heißt f **zweimal partiell differenzierbar** mit **partiellen Ableitungen zweiter Ordnung** $D_j D_i f$, $i = 1, \dots, m$, $j = 1, \dots, m$. **Zweimal stetig partiell differenzierbar** (Notation $C^2(U)$) bedeutet, dass zusätzlich alle partiellen Ableitungen erster und zweiter Ordnung stetig sind.*

Bemerkung.

- i) Induktiv werden Ableitungen k^{ter} Ordnung, $k > 2$, definiert.
- ii) Um Probleme zu vermeiden, werden im Folgenden wieder alle auftretenden Ableitungen als stetige Funktionen angenommen.

Wie hängt $D_j D_i f$ mit $D_i D_j f$ zusammen?

Idee. Klarerweise sind die **Ableitungen vertauschbar**, falls

$$f(\underline{\mathbf{x}}) = x_1^a x_2^b, \quad a, b \in \mathbb{R}, \quad U = \mathbb{R}^2.$$

Satz 17.2.4

*Ist $f: U \rightarrow \mathbb{R}$ **zweimal stetig partiell differenzierbar**, so gilt für alle $\underline{\mathbf{x}} \in U$, $i, j = 1, \dots, n$*

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := D_j D_i f(\underline{\mathbf{x}}) = D_i D_j f(\underline{\mathbf{x}}).$$

Vorsicht. Ohne obige Annahme der Stetigkeit aller Ableitungen, d.h. ohne die Voraussetzung “zweimal stetig partiell differenzierbar”, ist Satz 17.2.4 **falsch**. Beispiel (Übung!):

$$f(\underline{\mathbf{x}}) = \begin{cases} x_1 x_2 \frac{x_1^2 - x_2^2}{x_1^2 + x_2^2} & \text{für } \underline{\mathbf{x}} \neq \underline{\mathbf{0}}, \\ 0 & \text{für } \underline{\mathbf{x}} = \underline{\mathbf{0}}. \end{cases}$$

Beispiel. ([Laplace-Operator in Polarkoordinaten](#)) Die zweimalige Anwendung von (3) und (4) liefert in Polarkoordinaten (wieder in der symbolischen Schreibweise) (\rightsquigarrow Übungen)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{r} \frac{\partial}{\partial r}.$$

17.3 Der Satz von Taylor für Funktionen mehrerer Veränderlicher, lokale Extrema (Multiindex; kritische Punkte; notwendige und hinreichende Bedingungen für die Existenz lokaler Extrema; Definitheit einer Matrix; Hessesche Matrix)

In Kapitel 13 wurden der Satz von Taylor und Fehlerabschätzungen in einem Kapitel zusammengefasst. Zuvor waren bereits (Kapitel 11.2) lokale Extrema unabhängig davon diskutiert worden. Ebenso können aber notwendige und hinreichende Bedingungen für lokale Extrema aus dem Satz von Taylor hergeleitet werden. Dieser Zugang wird nun im Fall von Funktionen mehrerer Veränderlicher gewählt.

Zunächst sei daran erinnert, dass im letzten Paragraphen das Differential als “affin lineare Approximation” einer Abbildung $f: \mathbb{R}^m \rightarrow \mathbb{R}$ interpretiert wurde. Ebenso wurde der Begriff “höhere Ableitungen” zur Verfügung gestellt.

In diesem Abschnitt wird die Taylorsche Formel als “**Approximation höherer Ordnung**” eingeführt und anschließend werden, wie bereits erwähnt, **lokale Extrema** diskutiert.

Wieder ist im Folgenden $U \subset \mathbb{R}^m$ offen, $f: U \rightarrow \mathbb{R}$ ist stets eine **skalare** Funktion. Es sei ausdrücklich darauf hingewiesen, dass Extrema von Funktionen nur im skalaren Fall diskutiert werden können (warum?).

Bezeichnungen. Multiindizes: Es sei $\alpha = (\alpha_1, \dots, \alpha_m)$ ein m -Tupel natürlicher Zahlen mit

$$|\alpha| := \alpha_1 + \dots + \alpha_m, \quad \alpha! := \alpha_1! \alpha_2! \dots \alpha_m!$$

Ist f eine $|\alpha|$ -mal stetig differenzierbare Funktion, so setzt man

$$D^\alpha f := D_1^{\alpha_1} D_2^{\alpha_2} \dots D_m^{\alpha_m} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}} \quad \text{mit}$$

$$D_i^{\alpha_i} := \underbrace{D_i \dots D_i}_{\alpha_i\text{-mal}}, \quad i = 1, \dots, m.$$

Weiterhin sei $\underline{x}^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_m^{\alpha_m}$. Dann gilt die **Taylorische Formel**² (es wird exemplarisch die **Lagrangesche Restgliedformel** angegeben, vgl. Korollar 13.1.1, *ii*))

Satz 17.3.1

Es seien $\underline{x}^{(0)} \in U$, $\underline{\xi} \in \mathbb{R}^m$ derart, dass für alle $0 \leq t \leq 1$ gilt $\underline{x}^{(0)} + t\underline{\xi} \in U$. Die Funktion $f: U \rightarrow \mathbb{R}$ sei von der Klasse C^{k+1} . Dann existiert ein $\theta \in [0, 1]$ mit

$$f(\underline{x}^{(0)} + \underline{\xi}) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(\underline{x}^{(0)})}{\alpha!} \underline{\xi}^\alpha + \sum_{|\alpha|=k+1} \frac{D^\alpha f(\underline{x}^{(0)} + \theta \underline{\xi})}{\alpha!} \underline{\xi}^\alpha.$$

Beweisidee. Man betrachte die Funktion einer Variablen

$$g: [0, 1] \rightarrow \mathbb{R}, \quad g(t) := f(\underline{x}^{(0)} + t\underline{\xi}).$$

□

²Die Bezeichnungen aus Kapitel 13 übertragen sich in natürlicher Weise und werden nicht erneut eingeführt.

Bemerkungen.

i) Demnach ist

$$f(\underline{\mathbf{x}}^{(0)} + \underline{\xi}) = \sum_{|\alpha| \leq k} \frac{D^\alpha f(\underline{\mathbf{x}}^{(0)})}{\alpha!} \underline{\xi}^\alpha + o(\|\underline{\xi}\|^k),$$

wobei $o(\|\underline{\xi}\|^k)$ für eine Funktion $\varphi(\underline{\xi})$ steht mit

$$\varphi(\underline{\mathbf{0}}) = 0 \quad \text{und} \quad \lim_{\underline{\xi} \rightarrow \underline{\mathbf{0}}, \underline{\xi} \neq \underline{\mathbf{0}}} \frac{\varphi(\underline{\xi})}{\|\underline{\xi}\|^k} = 0.$$

Im Hinblick auf diese Notation sei an die Definition der Landauschen Symbole, Definition 13.2.1, erinnert.

ii) Damit die richtigen Vorfaktoren gewählt sind, ist unbedingt darauf zu achten, dass bei der Definition von $D^\alpha f$ die partiellen Ableitungen **in geordneter Reihenfolge** beginnend mit D_1 bis hin zu $D_m f$ auftauchen.

Beispiel. Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = e^{x_1 x_2}$. Zu berechnen sind zunächst die partiellen Ableitungen:

$$D_1 f(\underline{\mathbf{x}}) = x_2 e^{x_1 x_2}, \quad D_2 f(\underline{\mathbf{x}}) = x_1 e^{x_1 x_2},$$

$$D_1 D_1 f(\underline{\mathbf{x}}) = x_2^2 e^{x_1 x_2}, \quad D_1 D_2 f(\underline{\mathbf{x}}) = (1 + x_1 x_2) e^{x_1 x_2},$$

$$D_2 D_2 f(\underline{\mathbf{x}}) = x_1^2 e^{x_1 x_2},$$

$$D_1 D_1 D_1 f(\underline{\mathbf{x}}) = x_2^3 e^{x_1 x_2}, \quad D_1 D_1 D_2 f(\underline{\mathbf{x}}) = (2x_2 + x_1 x_2^2) e^{x_1 x_2},$$

$$D_1 D_2 D_2 f(\underline{\mathbf{x}}) = (2x_1 + x_2 x_1^2) e^{x_1 x_2}, \quad D_2 D_2 D_2 f(\underline{\mathbf{x}}) = x_1^3 e^{x_1 x_2}.$$

Es folgt für jedes fixierte $\underline{\mathbf{x}} \in \mathbb{R}^2$ und für alle $\underline{\xi} \in \mathbb{R}^2$

$$\begin{aligned} e^{(x_1+\xi_1)(x_2+\xi_2)} &= e^{x_1 x_2} + x_2 e^{x_1 x_2} \xi_1 + x_1 e^{x_1 x_2} \xi_2 \\ &\quad + \frac{1}{2} x_2^2 e^{x_1 x_2} \xi_1^2 + (1 + x_1 x_2) e^{x_1 x_2} \xi_1 \xi_2 + \frac{1}{2} x_1^2 e^{x_1 x_2} \xi_2^2 \\ &\quad + R_2(\underline{\xi}), \end{aligned}$$

wobei mit der Notation $\bar{x}_1 := x_1 + \theta\xi_1$, $\bar{x}_2 := x_2 + \theta\xi_2$, $0 < \theta < 1$ wie in Satz 17.3.1, gilt

$$R_2(\underline{\xi}) = \frac{1}{3!}\bar{x}_2^3 e^{\bar{x}_1\bar{x}_2}\xi_1^3 + \frac{1}{2!}(2\bar{x}_2 + \bar{x}_1\bar{x}_2^3)e^{\bar{x}_1\bar{x}_2}\xi_1^2\xi_2 + \frac{1}{2!}(2\bar{x}_1 + \bar{x}_2\bar{x}_1^2)e^{\bar{x}_1\bar{x}_2}\xi_1\xi_2^2 + \frac{1}{3!}\bar{x}_1^3 e^{\bar{x}_1\bar{x}_2}\xi_2^3.$$

Insbesondere geht also $R_2(\underline{\xi})$ wie $\|\underline{\xi}\|^3$ gegen 0 bei $\|\underline{\xi}\| \rightarrow 0$.

Im nächsten Schritt sollen **Extremwertaufgaben in mehreren Veränderlichen** diskutiert werden. Wie im Fall einer Variablen (vgl. Satz 11.2.1) wird zunächst eine **notwendige Bedingung** (Satz 17.3.2) studiert. Als Vorbereitung ist das Analogon zu Definition 11.2.1 festzuhalten:

Definition 17.3.1

*Es sei $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$. Dann hat f im Punkt $\underline{\mathbf{x}}^{(0)} \in U$ ein **lokales Maximum** (bzw. **lokales Minimum**) ($\underline{\mathbf{x}}^{(0)}$ ist ein **lokaler Maximierer** (bzw. **lokaler Minimierer**)), falls eine Umgebung $V \subset U$ von $\underline{\mathbf{x}}^{(0)}$ existiert mit*

$$f(\underline{\mathbf{x}}^{(0)}) \geq f(\underline{\mathbf{x}}) \quad (\text{bzw. } f(\underline{\mathbf{x}}^{(0)}) \leq f(\underline{\mathbf{x}}))$$

*für alle $\underline{\mathbf{x}} \in V$. Gilt zusätzlich für $\underline{\mathbf{x}} \neq \underline{\mathbf{x}}^{(0)}$ sogar die strikte Ungleichung, so handelt es sich um einen **isolierten Maximierer** (bzw. **isolierten Minimierer**).*

Bemerkungen.

- i)* In der Sprechweise von Definition 11.2.1 hat f in einem isolierten Maximierer (Minimierer) ein strenges lokales Maximum (Minimum).
- ii)* Wie im Fall $m = 1$ ist ein lokales Extremum von einem globalen (oder absoluten) Extremum (vgl. Kapitel 17.4) zu unterscheiden.

Satz 17.3.2

Im Punkt $\underline{\mathbf{x}}^{(0)} \in U$ habe f ein *lokales Extremum* (d.h. ein lokales Maximum oder Minimum), wobei f von der Klasse C^1 sei. Dann gilt

$$\nabla f(\underline{\mathbf{x}}^{(0)}) = \underline{\mathbf{0}},$$

d.h. $\underline{\mathbf{x}}^{(0)}$ ist ein *kritischer Punkt* von f .

Beweisidee. Vgl. Satz 11.2.1. □

Bemerkungen.

- i)* Mit einer kleinen Nachlässigkeit in der Notation wird $\nabla f(\underline{\mathbf{x}}^{(0)})$ im Folgenden meist als Spaltenvektor im \mathbb{R}^m interpretiert.
- ii)* Die Bedingung, dass $\nabla f(\underline{\mathbf{x}}^{(0)})$ verschwindet, ist lediglich **notwendig, aber nicht hinreichend** dafür, dass in $\underline{\mathbf{x}}^{(0)}$ ein lokales Extremum vorliegt (völlig analog zum Fall $m = 1$).
- iii)* An dieser Stelle ist es wichtig, dass U eine **offene Menge** ist (warum?).

Beispiele.

- i)* Lokale Maximierer, Minimierer und **Sattelpunkte** sind in Abbildung 17.13 veranschaulicht.
- ii)* Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = \sin(x_1) \sin(x_2)$. Es gilt

$$\nabla f(\underline{\mathbf{x}}) = \begin{pmatrix} \cos(x_1) \sin(x_2) \\ \sin(x_1) \cos(x_2) \end{pmatrix},$$

also

$$\nabla f(\underline{\mathbf{0}}) = \underline{\mathbf{0}}.$$

Aber: Für $x_1 = x_2$ ist $f(\underline{\mathbf{x}}) = \sin^2(x_1) > 0$, für $x_1 = -x_2$ ist $f(\underline{\mathbf{x}}) = -\sin^2(x_1) < 0$. Es ist $f(\underline{\mathbf{0}}) = 0$, beliebig nahe bei $\underline{\mathbf{0}}$ existieren aber Punkte, in denen f positive bzw. negative Werte annimmt. f hat in $\underline{\mathbf{0}}$ kein lokales Extremum, $\underline{\mathbf{0}}$ ist ein **Sattelpunkt**.

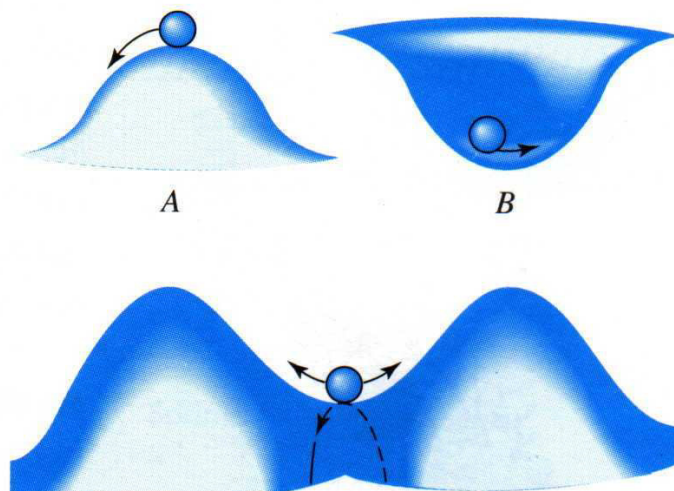


Abbildung 17.13: A: Maximierer; B: Minimierer; C: Sattelpunkt (aus “Hildebrandt/Tromba”).

Hinreichende Bedingungen für lokale Extrema?

Satz 11.2.5 hält hinreichende Bedingungen für die Existenz lokaler Extrema im Fall $m = 1$ fest. Dort wird mit Hilfe des Vorzeichens der zweiten Ableitung argumentiert. Im Fall mehrerer Veränderlicher ist zunächst gar nicht klar, wie eine Vorzeichenbedingung zu formulieren ist, um eine Verallgemeinerung des Satzes zu erhalten. Man setzt:

Definition 17.3.2

Es sei $A \in M(m, m, \mathbb{R})$ eine *symmetrische Matrix*.

i) A heißt *positiv definit*, falls

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle > 0 \quad \text{für alle } \underline{\mathbf{v}} \in \mathbb{R}^m - \{\underline{\mathbf{0}}\} .$$

ii) A heißt *positiv semidefinit*, falls

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle \geq 0 \quad \text{für alle } \underline{\mathbf{v}} \in \mathbb{R}^m .$$

iii) A heißt *negativ definit* (bzw. *negativ semidefinit*), falls die Matrix $-A$ positiv definit (bzw. positiv semidefinit) ist.

iv) A heißt *indefinit*, falls $\underline{\mathbf{v}}, \underline{\mathbf{w}} \in \mathbb{R}^m$ existieren mit

$$\langle \underline{\mathbf{v}}, A\underline{\mathbf{v}} \rangle > 0 , \quad \langle \underline{\mathbf{w}}, A\underline{\mathbf{w}} \rangle < 0 .$$

Beispiele.

i) Die Matrix $A = I_m \in M(m, m, \mathbb{R})$ ist positiv definit, da für alle $\underline{\mathbf{v}} \in \mathbb{R}^m - \{\underline{\mathbf{0}}\}$ gilt

$$\langle \underline{\mathbf{v}}, I_m \underline{\mathbf{v}} \rangle = \|\underline{\mathbf{v}}\|^2 > 0 .$$

ii) Es sei

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in M(2, 2, \mathbb{R}) .$$

Sind $\underline{\mathbf{e}}^{(1)}, \underline{\mathbf{e}}^{(2)}$ die kanonischen Basisvektoren des \mathbb{R}^2 , so gilt

$$\langle \underline{\mathbf{e}}^{(1)}, A\underline{\mathbf{e}}^{(1)} \rangle = 1 , \quad \langle \underline{\mathbf{e}}^{(2)}, A\underline{\mathbf{e}}^{(2)} \rangle = -1 ,$$

die Matrix A ist indefinit.

Bemerkung. Anhand obiger Beispiele erkennt man, dass die Vorzeichen der Eigenwerte Auskunft über die Definitheit (Indefinitheit) einer Matrix geben. Da hier nur symmetrische Matrizen betrachtet werden, sei an Kapitel 15.2 (insbesondere Satz 15.2.3) erinnert, aus dem die Existenz einer Orthonormalbasis aus Eigenvektoren folgt. Man überlege sich, welche Kombinationen von Vorzeichen der Eigenwerte einer Matrix deren Definitheit bzw. Indefinitheit implizieren (Übung!).

Zur weiteren Behandlung von Extremwertaufgaben ist die folgende Matrix der zweiten Ableitungen zu bilden:

Definition 17.3.3 ³

Es sei $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Die **Hessesche Matrix** von f im Punkt $\underline{\mathbf{x}}^{(0)} \in U$ ist die **symmetrische** $m \times m$ -Matrix

$$(\text{Hess } f)(\underline{\mathbf{x}}^{(0)}) := (D_j D_i f(\underline{\mathbf{x}}^{(0)}))_{\substack{1 \leq j \leq m \\ 1 \leq i \leq m}} \in M(m, m, \mathbb{R}).$$

Mit den Definitionen 17.3.2 und 17.3.3 kann schließlich formuliert werden:

Satz 17.3.3

Es sei $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Existiert ein Punkt $\underline{\mathbf{x}}^{(0)} \in U$ mit $\nabla f(\underline{\mathbf{x}}^{(0)}) = \mathbf{0}$, so gilt:

- i) Ist $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$ **positiv definit**, so ist $\underline{\mathbf{x}}^{(0)}$ ein **isolierter Minimierer** von f .
- ii) Ist $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$ **negativ definit**, so ist $\underline{\mathbf{x}}^{(0)}$ ein **isolierter Maximierer** von f .
- iii) Ist $(\text{Hess } f)(\underline{\mathbf{x}}^{(0)})$ **indefinit**, so besitzt f in $\underline{\mathbf{x}}^{(0)}$ **kein lokales Extremum**.

³L.O. Hesse, 1811–1874; Königsberg, Halle, Heidelberg, München.

Beispiel. Im \mathbb{R}^3 sei das **Paraboloid** $c + x_1^2 + x_2^2 = x_3$, $c = \text{konst.}$ betrachtet, m.a.W.: Man betrachte die Funktion

$$f(\underline{\mathbf{x}}) := c + x_1^2 + x_2^2.$$

Es gilt

$$\nabla f(\underline{\mathbf{x}}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}, \quad \nabla f(\underline{\mathbf{0}}) = \underline{\mathbf{0}},$$

die Hessesche Matrix

$$(\text{Hess } f)(\underline{\mathbf{x}}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

ist stets positiv definit, $\underline{\mathbf{x}}^{(0)} = \underline{\mathbf{0}}$ ist also ein isolierter Minimierer, wie es schon in Abbildung 17.11 angedeutet wurde.

Bemerkungen.

- i)* Im Fall $m = 2$ kann die Situation wie folgt zusammengefasst werden: Ist in einem kritischen Punkt $\underline{\mathbf{x}}^{(0)} \in U$

$$(\text{Hess } f)(\underline{\mathbf{x}}^{(0)}) = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

so ist

- $a > 0$, $ac - b^2 > 0$ **hinreichend für ein Minimum**;
- $a < 0$, $ac - b^2 > 0$ **hinreichend für ein Maximum**;
- $a \geq 0$, $ac - b^2 \geq 0$ **notwendig für ein Minimum**;
- $a \leq 0$, $ac - b^2 \geq 0$ **notwendig für ein Maximum**.

Ist hingegen $ac - b^2 < 0$, so kann $\underline{\mathbf{x}}^{(0)}$ **keine Extremalstelle** sein.

- ii)* Zur systematischen Untersuchung des Falls $m > 2$ sei auf die Literatur verwiesen.

Beispiel. Ist $m = 2$ und

$$f(\underline{\mathbf{x}}) = 6x_1x_2 - 3x_2^2 - 2x_1^3,$$

so sind die einzigen kritischen Punkte

$$\underline{\mathbf{x}}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \underline{\mathbf{x}}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Die Hessesche Matrix ist

$$(\text{Hess } f)(\underline{\mathbf{x}}) = \begin{pmatrix} -12x_1 & 6 \\ 6 & -6 \end{pmatrix}.$$

Demnach kann $\underline{\mathbf{x}}^{(0)}$ keine Extremalstelle sein, in $\underline{\mathbf{x}}^{(1)}$ hat f ein lokales Maximum.

17.4 Implizite Funktionen, Umkehrsatz, Extrema mit Nebenbedingungen (Lagrangescher Multiplikator; globale Extrema)

In diesem Paragraphen wird zunächst der Satz über implizite Funktionen skizziert, als wesentliche Konsequenzen werden der Umkehrsatz (von zentraler Bedeutung etwa bei Koordinatentransformationen) und die Diskussion von Extrema unter Nebenbedingungen folgen. Mit letzterer eröffnet sich die Möglichkeit, auch im Fall von Funktionen mehrerer Veränderlicher, die Frage nach **globalen** Extrema zu erörtern.

Worum es im Satz über implizite Funktionen geht.

- i)* Zu einer Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ (im einfachsten Fall) wird die **implizite Gleichung**

$$g(x, y) = 0$$

untersucht. Es stellt sich die Frage: Wann kann man diese Gleichung **nach der zweiten Variablen y „auflösen“**, d.h. wann existiert eine Funktion $y = f(x)$ mit $g(x, f(x)) = 0$ und ist diese Funktion evtl. eindeutig bestimmt?

- ii) Beispielsweise können auf diese Art die **Höhenlinien einer Funktion $h(x, y)$ analysiert werden**, indem man setzt: $g(x, y) := h(x, y) - c$.
- iii) Auf weitere Folgerungen (wie oben angedeutet) im Anschließen eingegangen werden.

Beispiele.

- i) Man betrachte die (affin) lineare Funktion $g(x, y) = 3x + by - 4$, $b \in \mathbb{R}$ fixiert, also die lineare Gleichung

$$3x + by - 4 = 0 .$$

Ist $b \neq 0$, so kann diese Gleichung nach y aufgelöst werden mit

$$y = f(x) = \frac{4}{b} - \frac{3}{b}x .$$

Die Einschränkung $b \neq 0$ korrespondiert mit der ersten wesentlichen Voraussetzung des Satzes über implizite Funktionen. Ohne diese Einschränkung kann die Auflösbarkeit nach y nicht erwartet werden.

- ii) Um die zweite wesentliche Voraussetzung des Satzes über implizite Funktionen zu verstehen, betrachte man die Einheitskreislinie im \mathbb{R}^2 , die gegeben ist durch die implizite Gleichung (mit der Schreibweise $\mathbb{R}^2 \ni \underline{x} = (x, y)$)

$$g(x, y) = x^2 + y^2 - 1 = 0 .$$

Hier kommen sowohl

$$f(x) = \sqrt{1 - x^2} \quad \text{als auch} \quad \tilde{f}(x) = -\sqrt{1 - x^2}$$

als "Auflösungen" nach y in Frage, **es existiert im Allgemeinen also höchstens eine lokale Lösung des Problems in einer Umgebung U eines fixierten Punktes $(x^{(0)}, y^{(0)})$** . M.a.W.: Man kann nicht erwarten, dass die komplette Nullstellenmenge von g mit Hilfe einer Funktion $y = f(x)$ charakterisiert werden kann, es werden in verschiedenen Bereichen unterschiedliche Auflösungen heranzuziehen sein (vgl. Abbildung 17.14).

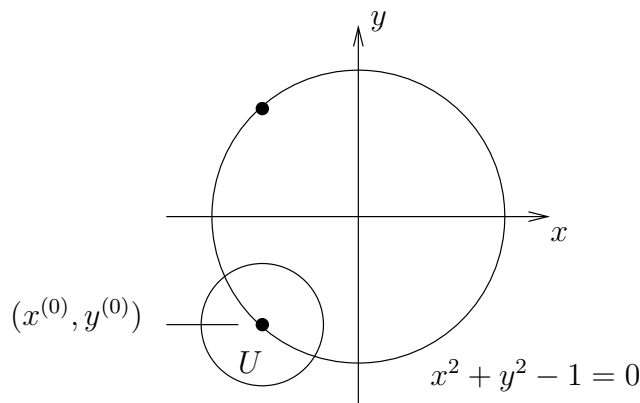


Abbildung 17.14: Die Gleichung $x^2 + y^2 - 1 = 0$ ist höchstens lokal auflösbar.

iii) In Verallgemeinerung des ersten Beispiels sei nun $g: \mathbb{R}^m \times \mathbb{R}^n \rightarrow$

$$\mathbb{R}^n, \underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m, \underline{\mathbf{y}} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n. \text{ Mit der offensichtli-}$$

chen Bedeutung der Notation wird das **lineare Gleichungssystem**

$$g(\underline{\mathbf{x}}, \underline{\mathbf{y}}) = \underline{\mathbf{0}}$$

untersucht. Es handelt sich hier um die n linearen Gleichungen in den $m + n$ Unbekannten $x_1, \dots, x_m, y_1, \dots, y_n$:

$$g_j(\underline{\mathbf{x}}, \underline{\mathbf{y}}) = \sum_{k=1}^m a_{jk} x_k + \sum_{l=1}^n b_{jl} y_l - c_j = 0, \quad j = 1, \dots, n,$$

wobei alle a_{jk} , b_{jl} und $c_j \in \mathbb{R}$ fixiert seien. Mit $A = (a_{jk}) \in$

$$M(n, m, \mathbb{R}), B = (b_{jl}) \in M(n, n, \mathbb{R}) \text{ und } \underline{\mathbf{c}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n, \text{ kann}$$

das System geschrieben werden als

$$B\underline{\mathbf{y}} = \underline{\mathbf{c}} - A\underline{\mathbf{x}}.$$

Es kann also genau dann nach $\underline{\mathbf{y}}$ aufgelöst werden, wenn gilt

$$\det B \neq 0.$$

Idee. Die Matrix B ist die Jacobi Matrix der Funktion g in Abhängigkeit von der Variablen \underline{y} . Ist **lokal** die erste Ableitung einer Funktion von dominierender Ordnung, so sollte ein Analogon im allgemeinen Fall richtig sein. In der Tat gilt:

Satz 17.4.1 (*Satz über implizite Funktionen*)

Es seien $U_1 \subset \mathbb{R}^m$ und $U_2 \subset \mathbb{R}^n$ offene Mengen und

$$\begin{aligned} g : U_1 \times U_2 &\rightarrow \mathbb{R}^n, \\ (\underline{x}, \underline{y}) &\mapsto g(\underline{x}, \underline{y}), \end{aligned}$$

eine Abbildung von der Klasse C^1 . Es sei weiter $(\underline{x}^{(0)}, \underline{y}^{(0)}) \in U_1 \times U_2$ ein fixierter Punkt mit

$$g(\underline{x}^{(0)}, \underline{y}^{(0)}) = \underline{\mathbf{0}},$$

und die Matrix

$$D_{\underline{y}}g := \left(\frac{\partial g_i}{\partial y_j} \right)_{\substack{1 \leq j \leq n \\ 1 \leq i \leq n}} \in M(n, n, \mathbb{R})$$

sei in diesem Punkt **invertierbar**. Dann gibt es offene **Umgebungen** $V_1 \subset U_1$ von $\underline{x}^{(0)}$ und $V_2 \subset U_2$ von $\underline{y}^{(0)}$ und eine stetig differenzierbare Abbildung $f: V_1 \rightarrow V_2$ mit

$$g(\underline{x}, f(\underline{x})) = \underline{\mathbf{0}} \quad \text{für alle } \underline{x} \in V_1.$$

Ist $(\underline{x}, \underline{y}) \in V_1 \times V_2$ ein Punkt mit $g(\underline{x}, \underline{y}) = \underline{\mathbf{0}}$, so folgt $\underline{y} = f(\underline{x})$.

Bemerkungen.

- i) Wie Abbildung 17.15 zeigt, ist die Verkleinerung von U_1 und U_2 tatsächlich notwendig.

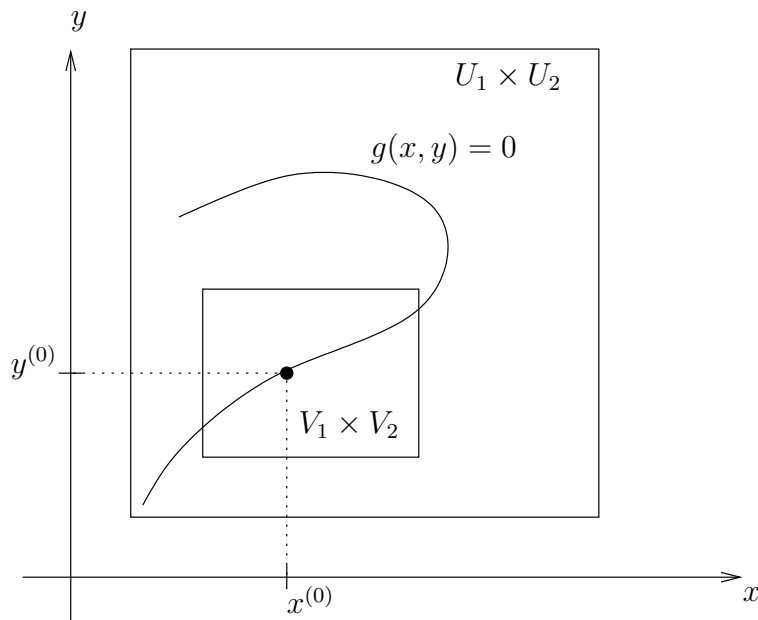


Abbildung 17.15: Zum Satz über implizite Funktionen.

- ii) Es gilt nach der Kettenregel (lokal bei $(\underline{\mathbf{x}}^{(0)}, f(\underline{\mathbf{x}}^{(0)}))$)

$$Df(\underline{\mathbf{x}}) = - \left(D_{\underline{\mathbf{y}}}g(\underline{\mathbf{x}}, f(\underline{\mathbf{x}})) \right)^{-1} D_{\underline{\mathbf{x}}}g(\underline{\mathbf{x}}, f(\underline{\mathbf{x}})) .$$

- iii) Die Invertierbarkeit von $D_{\underline{\mathbf{y}}}g$ ist nur möglich, wenn $U_2 \subset \mathbb{R}^n$, d.h. “die Dimensionen müssen zueinander passen” (warum?).

Beispiel. Im einfachsten Fall $m = n = 1$ betrachte man das wohlbekannte Paraboloid, das als Graph der Funktion $h: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$h(\underline{\mathbf{x}}) = x_1^2 + x_2^2 + 1$$

gegeben ist. Es soll nun die Höhenlinie $\{\underline{\mathbf{x}} \in \mathbb{R}^2 : h(\underline{\mathbf{x}}) = 6\}$ mit Hilfe des Satzes über implizite Funktionen diskutiert werden (in dem Fall eine Kreislinie). Dazu sei in der Notation von Satz 17.4.1 die Gleichung

$$g(x, y) := x^2 + y^2 - 5 = 0$$

betrachtet und der Punkt $(x^{(0)}, y^{(0)}) = (1, 2)$ fixiert.

Der Punkt genügt offensichtlich der Gleichung und es ist

$$D_y g(1, 2) = 2y|_{(x=1, y=2)} = 4 \neq 0,$$

der Satz über implizite Funktionen ist somit anwendbar und liefert eine Auflösung $y = f(x)$ in der Nähe des betrachteten Punktes. Für die Ableitung folgt (wiederum in der Nähe dieses Punktes)

$$Df = -(2y)^{-1}(2x), \quad Df(1, 2) = -\frac{1}{4}2 = -\frac{1}{2}.$$

Probe. Klarerweise handelt es sich im obigen Beispiel um eine Kreislinie vom Radius $\sqrt{5}$ um den Ursprung. Diese ist **in der Nähe von $(1, 2)$** gegeben durch

$$y = f(x) = \sqrt{5 - x^2}.$$

Es folgt (wie oben bereits hergeleitet)

$$f'(1) = \frac{1}{2} \frac{1}{\sqrt{5 - x^2}} (-2x)|_{x=1} = -\frac{1}{2}.$$

Im Gegensatz zu diesem Beispiel kann im Allgemeinen die Funktion f nicht explizit angegeben werden, man kann lediglich ihre Ableitung in dem fixierten Punkt, also ihr ‘Verhalten erster Ordnung’ bestimmen.

Ohne die Diskussion des Satzes über implizite Funktionen zu vertiefen, sei hier eine wichtige Folgerung betont (vgl. Satz 11.1.6):

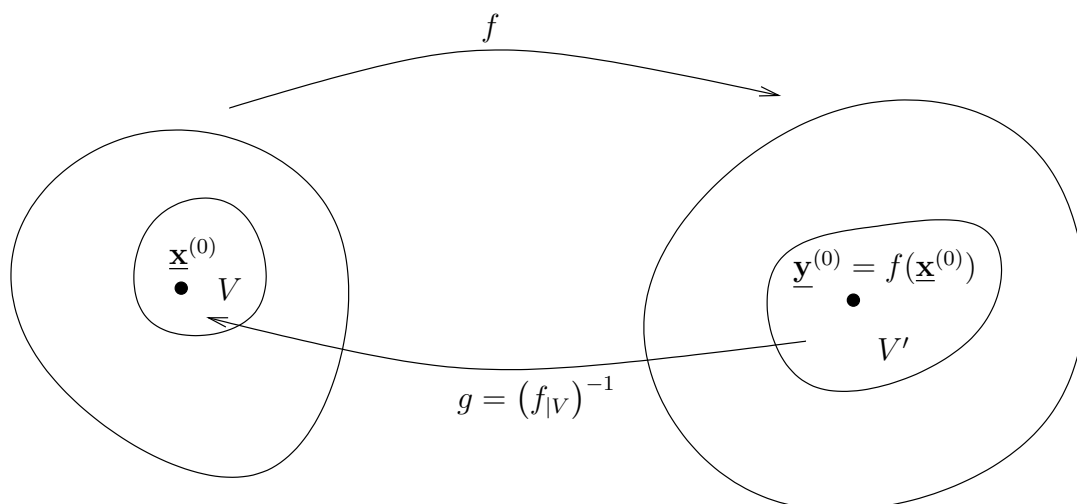


Abbildung 17.16: Zum Umkehrsatz.

Satz 17.4.2 (*Umkehrsatz*, siehe Abbildung 17.16)

Es sei $U \subset \mathbb{R}^m$ offen und $f: U \rightarrow \mathbb{R}^m$ stetig differenzierbar. Es sei weiter $\underline{\mathbf{x}}^{(0)} \in U$, $\underline{\mathbf{y}}^{(0)} = f(\underline{\mathbf{x}}^{(0)})$ und die **Jacobi-Matrix $Df(\underline{\mathbf{x}}^{(0)})$ sei invertierbar**. Dann gibt es eine offene Umgebung $V \subset U$ von $\underline{\mathbf{x}}^{(0)}$ und eine offene Umgebung V' von $\underline{\mathbf{y}}^{(0)}$, sodass f die Menge V **bijektiv** auf V' abbildet und die **Umkehrabbildung**

$$g := (f|_V)^{-1} : V' \rightarrow V$$

stetig differenzierbar ist mit

$$Dg(\underline{\mathbf{y}}^{(0)}) = (Df(\underline{\mathbf{x}}^{(0)}))^{-1} .$$

Bemerkungen.

- i) Wie im Satz über implizite Funktionen kann nur die **lokale Umkehrbarkeit** geschlossen werden.
- ii) Die (lokale) Umkehrbarkeit ist etwa bei Koordinatentransformationen wichtig, die ein Problem auf beispielsweise den Symmetrien angepasste Koordinaten transformieren. Hierbei müssen die neuen Koordinaten (zumindest lokal) 1-1 auf die alten umgerechnet werden können.
- iii) Invertierbarkeit kann nur gelten, wenn die Dimension des Urbildraumes gleich der des Bildraumes ist, d.h. $U \subset \mathbb{R}^m$ und das Bild liegt ebenfalls im \mathbb{R}^m .

Beispiel. Polarkoordinaten: Es sei

$$f : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \underline{\mathbf{x}} := \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix},$$

$$Df \left(\begin{pmatrix} r \\ \varphi \end{pmatrix} \right) = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix}.$$

Dann gilt

$$\det \left(Df \left(\begin{pmatrix} r \\ \varphi \end{pmatrix} \right) \right) = r > 0 ,$$

die Abbildung f ist nach Satz 17.4.2 in jedem Punkt $\begin{pmatrix} r \\ \varphi \end{pmatrix} \in (0, \infty) \times$

\mathbb{R} **lokal invertierbar** mit $(g := (f|_V)^{-1})$

$$\begin{aligned} Dg \left(f \left(\begin{pmatrix} r \\ \varphi \end{pmatrix} \right) \right) &= \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix}^{-1} \\ &= \frac{1}{r} \begin{pmatrix} r \cos(\varphi) & r \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} . \end{aligned}$$

Aus

$$r = \sqrt{x_1^2 + x_2^2}, \quad \frac{x_1}{r} = \cos(\varphi), \quad \frac{x_2}{r} = \sin(\varphi),$$

folgt schließlich

$$Dg(\underline{\mathbf{x}}) = \frac{1}{\sqrt{x_1^2 + x_2^2}} \begin{pmatrix} x_1 & x_2 \\ -\frac{x_2}{\sqrt{x_1^2 + x_2^2}} & \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \end{pmatrix} .$$

Bemerkungen.

i) Natürlich gilt in diesem Beispiel

$$f \left(\begin{pmatrix} r \\ \varphi \end{pmatrix} \right) = f \left(\begin{pmatrix} r \\ \varphi + 2\pi k \end{pmatrix} \right), \quad k \in \mathbb{Z},$$

es gibt also **keine globale Umkehrabbildung**.

ii) Weitere Beispiele werden in den Übungen diskutiert.

Eine zweite wesentliche Konsequenz aus dem Satz über implizite Funktionen ist der folgende Satz 17.4.3 über **Extrema mit Nebenbedingungen**. Es sei ausdrücklich darauf hingewiesen, dass eine Vielzahl von physikalischen Problemen Nebenbedingungen involvieren, beispielsweise ist bei einem Pendel der Abstand der schwingenden Masse zum Drehpunkt konstant etc. Auch die Diskussion von **globalen** (oder **absoluten**) Extrema einer Funktion führt in der Regel zu einer Nebenbedingung, nämlich dass die Funktion gesondert auf dem Rand ihres Definitionsbereiches zu untersuchen ist. Hier soll jedoch als Motivation ein geometrisches Beispiel erörtert werden.

Beispiel. Es soll der Flächeninhalt A eines Dreieckes bei **gegebenem Umfang** maximiert werden. Zur Erinnerung: Es gilt die **Heronische Flächenformel**

$$A^2 = s(s-a)(s-b)(s-c) .$$

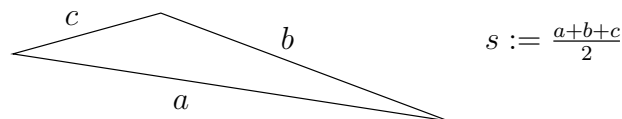


Abbildung 17.17: Zur Heronischen Flächenformel.

Intuitive Vermutung. Die Lösung ist ein **gleichseitiges Dreieck**, d.h.

$$a = b = c = \frac{2}{3}s .$$

Satz 17.4.3

Es sei $U \subset \mathbb{R}^m$ offen, $g: U \rightarrow \mathbb{R}$ stetig differenzierbar und

$$M := \{ \underline{\mathbf{x}} \in U : g(\underline{\mathbf{x}}) = 0 \} .$$

Es sei weiter $\underline{\mathbf{x}}^{(0)} \in M$ ein Punkt mit

$$\nabla g(\underline{\mathbf{x}}^{(0)}) \neq \underline{\mathbf{0}} ,$$

und $f: U \rightarrow \mathbb{R}$ sei eine stetig differenzierbare Funktion, die im Punkt $\underline{\mathbf{x}}^{(0)}$ ein lokales Maximum (Minimum) unter der Nebenbedingung $g = 0$ besitze, d.h.:

Für eine Umgebung $V \subset U$ von $\underline{\mathbf{x}}^{(0)}$ gelte

$$f(\underline{\mathbf{x}}^{(0)}) \geq f(\underline{\mathbf{x}}) \quad (\text{bzw. } f(\underline{\mathbf{x}}^{(0)}) \leq f(\underline{\mathbf{x}}))$$

für alle $\underline{\mathbf{x}} \in M \cap V$. Dann gibt es einen *Lagrangeschen Multiplikator* $\lambda \in \mathbb{R}$ mit

$$\nabla f(\underline{\mathbf{x}}^{(0)}) = \lambda \nabla g(\underline{\mathbf{x}}^{(0)}) .$$

Bemerkung. Analog zu Satz 17.3.2 handelt es sich hier um eine **notwendige Bedingung**, d.h. **falls** f ein lokales Extremum unter der Nebenbedingung $g = 0$ besitzt und **falls** ∇g in dem entsprechenden Punkt nicht verschwindet, dann existiert ein Lagrangescher Multiplikator mit obigen Eigenschaften. M.a.W.: Mit dem Satz werden mögliche Kandidaten für Extrema unter Nebenbedingungen ermittelt.

Zurück zu obigem Beispiel. Hier ist der Umfang oder äquivalent $s > 0$ fixiert. Mit der Notation

$$\underline{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

lautet die Aufgabe

$$f(\underline{\mathbf{x}}) := s(s - x_1)(s - x_2)(s - x_3) \rightarrow \max$$

unter der Nebenbedingung

$$g(\underline{\mathbf{x}}) := x_1 + x_2 + x_3 - 2s = 0 .$$

In dem Beispiel ist

$$M := \{ \underline{\mathbf{x}} \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ = U : g(\underline{\mathbf{x}}) = 0 \},$$

$$\nabla g(\underline{\mathbf{x}}) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \neq \underline{\mathbf{0}} \quad \text{für alle } \underline{\mathbf{x}} \in M .$$

Falls ein Maximum in einem Punkt $\underline{\mathbf{x}}^{(0)}$ existiert, so folgt also aus Satz 17.4.3

$$\begin{aligned} \nabla f(\underline{\mathbf{x}}^{(0)}) &= - \begin{pmatrix} s(s-x_2)(s-x_3) \\ s(s-x_1)(s-x_3) \\ s(s-x_1)(s-x_2) \end{pmatrix} \\ &= \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\ g(\underline{\mathbf{x}}^{(0)}) &= x_1 + x_2 + x_3 - 2s = 0. \end{aligned}$$

Man hat insgesamt (d.h. inklusive der Nebenbedingung) vier Gleichungen in vier Unbekannten, in diesem speziellen Fall gibt es auch vier mögliche Lösungen:

1^{te} Möglichkeit:

$$\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \lambda \end{pmatrix} = \begin{pmatrix} \frac{2}{3}s \\ \frac{2}{3}s \\ \frac{2}{3}s \\ -\frac{s^3}{9} \end{pmatrix} \Rightarrow f(\underline{\mathbf{x}}^{(0)}) = \frac{s^4}{27} > 0.$$

2^{te} bis 4^{te} Möglichkeit:

$$\begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \lambda \end{pmatrix} = \begin{pmatrix} s \\ s \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \lambda \end{pmatrix} = \begin{pmatrix} s \\ 0 \\ s \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ x_3^{(0)} \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ s \\ s \\ 0 \end{pmatrix}.$$

In all diesen Fällen gilt

$$f(\underline{\mathbf{x}}^{(0)}) = 0.$$

Falls also eine Lösung der Aufgabe existiert, so kommt nur der erste Kandidat in Frage, es handelt sich dann tatsächlich um ein gleichseitiges Dreieck.

Bemerkung. Ist $f: \mathbb{R}^m \supset U \rightarrow \mathbb{R}$ von der Klasse C^1 und ist man an globalen (absoluten) Maximierern (Minimierern) von f interessiert, d.h. an Punkten $\underline{\mathbf{x}}^{(0)} \in U$ mit

$$f(\underline{\mathbf{x}}^{(0)}) \geq f(\underline{\mathbf{x}}) \quad (\text{bzw. } f(\underline{\mathbf{x}}^{(0)}) \leq f(\underline{\mathbf{x}}))$$

für alle $\underline{\mathbf{x}} \in U$, so kann man nach Satz 10.1.5 deren Existenz schließen, falls *Ukompakt* ist. Die notwendige Bedingung aus Satz 17.3.2 greift aber nur im Innern von U , Extrema am Rand sind nicht durch eine horizontale Tangentialebene charakterisiert (selbst wenn man am Rand differenzieren dürfte). Dies sieht man auch leicht anhand eindimensionaler Beispiele ein (siehe Abbildung 17.18). Wie in den Übungen diskutiert wird, liefert hier Satz 17.4.3 ein wichtiges Hilfsmittel zum Auffinden der globalen Extrema.

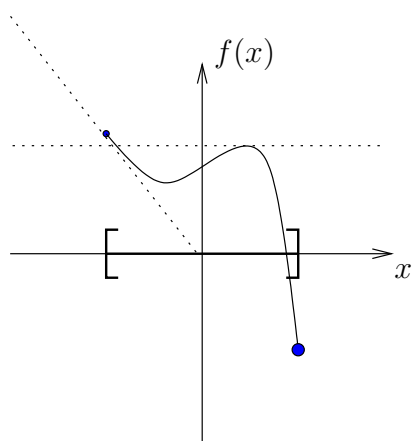


Abbildung 17.18: Extrema am Rand.

Kapitel 18

Kurvenintegrale

18.1 Definition und Eigenschaften (Weg; Wegunabhängigkeit des Kurvenintegrals; geschlossener Weg; konservative Vektorfelder; Potential)

Man betrachte eine stetig differenzierbare Kurve

$$\gamma : I = [a, b] \rightarrow U \subset \mathbb{R}^m .$$

In einem Kraftfeld $F: U \rightarrow \mathbb{R}^m$ bewege sich ein Massenpunkt längs dieser Kurve.

Wie berechnet sich die verrichtete Arbeit?

Betrachtet sei dazu eine Zerlegung $\mathcal{Z} = \{t_0, t_1, \dots, t_N\}$ (siehe Kapitel 17.1) des Intervalls $[a, b]$. Im Punkt $\gamma(t_{i-1})$ wirkt die Kraft $F(\gamma(t_{i-1}))$. Die zu verrichtende Arbeit (“Kraft \times Weg”) ist jedoch nur vom Anteil der **Kraft in Kurvenrichtung** abhängig.

Bei der Bewegung des Massenpunktes vom Ort $\gamma(t_{i-1})$ zum Ort $\gamma(t_i)$ ist also (bei glattem Kraftfeld) näherungsweise die Arbeit

$$\langle F(\gamma(t_{i-1})), \underbrace{\gamma(t_i) - \gamma(t_{i-1})}_{\approx \gamma'(t_{i-1})(t_i - t_{i-1})} \rangle$$

verrichtet (vgl. Abbildung 18.1). Summation und Grenzübergang $N \rightarrow \infty$ ergeben für die verrichtete Arbeit W den Ausdruck

$$W = \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt .$$

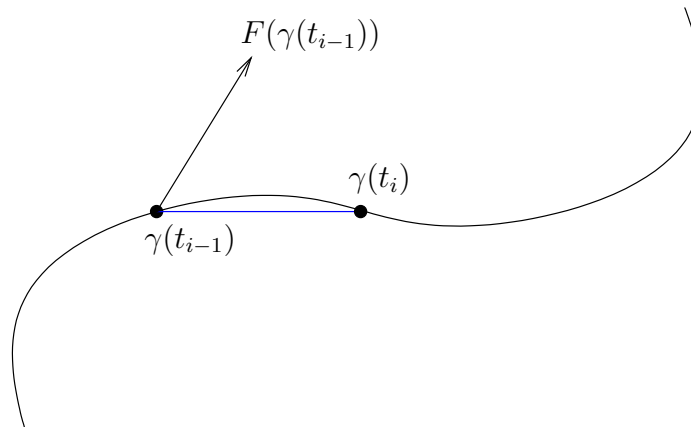


Abbildung 18.1: Zur Berechnung der verrichteten Arbeit.

Dies motiviert

Definition 18.1.1

Es sei $U \subset \mathbb{R}^m$ offen, $F: U \rightarrow \mathbb{R}^m$ sei ein stetiges Vektorfeld und $\gamma: I = [a, b] \rightarrow U$ sei eine glatte Kurve. Dann heißt

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle := \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt =: \int_{\gamma} \left[\sum_{i=1}^m F_i dx_i \right]$$

das **Kurvenintegral** des Vektorfeldes F längs der Kurve γ .

Bemerkungen.

- i) Für **stückweise glatte Kurven** γ (vgl. Abbildung 18.2) (zu beachten: Kurven sind per Definitionem aber zumindest stetig), die sich aus glatten “Teilkurven” $\gamma^{(1)}, \dots, \gamma^{(k)}$ zusammensetzen, setzt man

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle := \sum_{i=1}^k \int_{\gamma^{(i)}} \langle F, d\mathbf{x} \rangle .$$

- ii) Die Einführung des Kurvenintegrals ist oben motiviert durch den Begriff der Arbeit. Bewegt sich nun ein Massenpunkt längs der gleichen Spur einer Kurve mit unterschiedlicher Geschwindigkeit, so

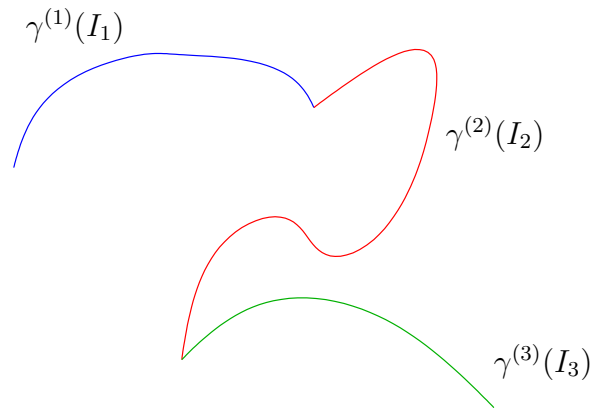


Abbildung 18.2: Eine stückweise glatte Kurve (deren Spur).

bleibt die verrichtete Arbeit die gleiche. Dies soll nun anhand der Definition des Kurvenintegrals verifiziert werden: Man betrachte dazu eine Kurve $\gamma: [a, b] \rightarrow \mathbb{R}^m$ sowie eine glatte Parametertransformation (siehe Definition 17.1.3) $\varphi: [\alpha, \beta] \rightarrow [a, b]$. Ist die Transformation **orientierungstreu**, so gilt für die Kurve $\tilde{\gamma}: [\alpha, \beta] \rightarrow \mathbb{R}^m$, $\tilde{\gamma} := \gamma \circ \varphi$:

$$\begin{aligned} \int_{\tilde{\gamma}} \langle F, d\underline{\mathbf{x}} \rangle &= \int_{\alpha}^{\beta} \langle F(\tilde{\gamma}(\tau)), \tilde{\gamma}'(\tau) \rangle d\tau \\ &= \int_{\alpha}^{\beta} \langle F \circ \gamma(\varphi(\tau)), \gamma'(\varphi(\tau)) \rangle \varphi'(\tau) d\tau \\ &\stackrel{\text{Subst.}}{=} \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt = \int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle. \end{aligned}$$

iii) Identifiziert man (als sogenannte **Äquivalenzrelation**) Kurven, die durch orientierungserhaltende Parametertransformationen auseinander hervorgehen, so spricht man von einem **Weg**. Nach ii) ist es dann auch berechtigt, von einem **Wegintegral** zu sprechen. Auf die genaue Unterscheidung zwischen Kurve und Weg wird hier meist nicht eingegangen.

iv) Ist $\tilde{\gamma}$ eine Kurve, die mittels einer **orientierungsumkehrenden** Parametertransformation aus γ hervorgeht, so sieht man wie oben

$$\int_{\tilde{\gamma}} \langle F, d\underline{\mathbf{x}} \rangle = - \int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle.$$

Aus diesem Grunde wird manchmal $-\gamma$ für einen in umgekehrter Orientierung durchlaufenen Weg geschrieben.

Beispiele.

- i) Es fließe ein konstanter Strom I durch einen unendlich langen Leiter im \mathbb{R}^3 . Es wird ein Magnetfeld

$$B : U = \{\underline{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0\} \rightarrow \mathbb{R}^3$$

aufgebaut mit

$$B(\underline{x}) = I \begin{pmatrix} -\frac{x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \\ 0 \end{pmatrix}.$$

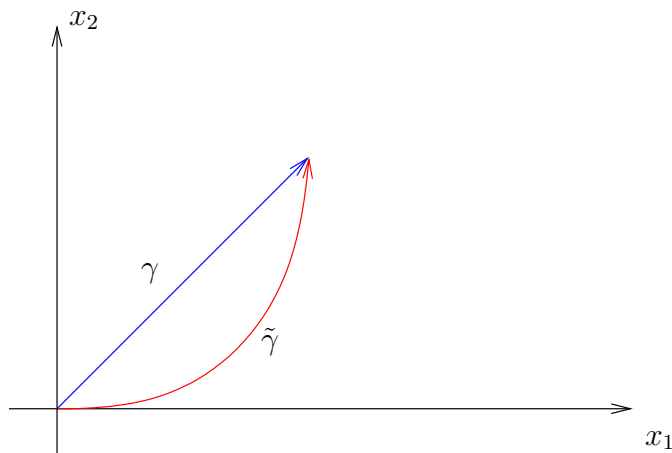


Abbildung 18.3: Magnetfeld um einen Leiter.

Berechnet werden soll das Kurvenintegral längs der Kurve γ ,

$$t \in [0, 2\pi) \mapsto \begin{pmatrix} r \cos(t) \\ r \sin(t) \\ konst. \end{pmatrix}, \quad r > 0.$$

Es ist

$$\begin{aligned} \int_{\gamma} \langle B, d\mathbf{x} \rangle &= I \int_0^{2\pi} \frac{1}{r^2} \left\langle \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 0 \end{pmatrix}, \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 0 \end{pmatrix} \right\rangle dt \\ &= I \int_0^{2\pi} 1 dt = 2\pi I . \end{aligned}$$

Auf dieses Beispiel wird in Kürze noch zurückgegriffen werden.

ii) Das Vektorfeld $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sei gegeben durch

$$F(\mathbf{x}) = \begin{pmatrix} x_1^2 + x_2^2 \\ x_1 x_2 \end{pmatrix} .$$

Betrachtet seien weiter die Kurven

$$\begin{aligned} \gamma : t \in [0, 1] &\rightarrow \begin{pmatrix} t \\ t \end{pmatrix} \in \mathbb{R}^2 , \\ \tilde{\gamma} : t \in [0, 1] &\rightarrow \begin{pmatrix} t \\ t^2 \end{pmatrix} \in \mathbb{R}^2 . \end{aligned}$$

Zu beachten ist dabei: **Anfangs- und Endpunkt beider Kurven stimmen überein.**

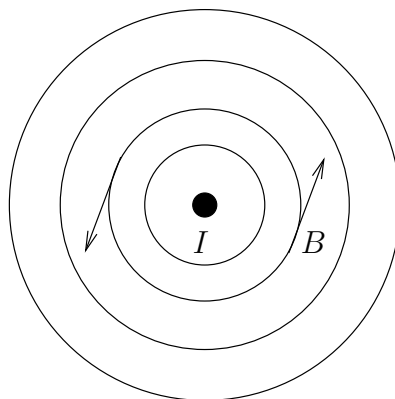


Abbildung 18.4: Zur Wegabhängigkeit des Kurvenintegrals.

Es gilt aber:

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle = \int_0^1 \left\langle \begin{pmatrix} t^2 + t^2 \\ t^2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle dt = \int_0^1 3t^2 dt = 1 ,$$

$$\begin{aligned} \int_{\tilde{\gamma}} \langle F, d\mathbf{x} \rangle &= \int_0^1 \left\langle \begin{pmatrix} t^2 + t^4 \\ t^3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2t \end{pmatrix} \right\rangle dt = \int_0^1 (3t^4 + t^2) dt \\ &= \frac{14}{15} . \end{aligned}$$

Im Allgemeinen hängt das Kurvenintegral also nicht nur vom Anfangs- und Endpunkt der Kurve ab, es kommt auch auf die spezielle Wahl des Weges an.

In welchen Fällen ist das Kurvenintegral **wegunabhängig**, d.h. nur abhängig vom Anfangs- und Endpunkt?

Dabei ist die **Wegunabhängigkeit äquivalent zu**

$$\int_{\gamma} \langle F, d\mathbf{x} \rangle = 0 \quad \text{für jeden geschlossenen Weg ,}$$

d.h. für jede geschlossen durchlaufene Kurve γ , d.h. $\gamma(a) = \gamma(b)$. Dies sieht man anhand von Abbildung 18.5 ein, wenn der Weg $\gamma = \gamma^{(1)} - \gamma^{(2)}$ betrachtet wird (vgl. obige Bemerkung zu orientierungsumkehrenden Parametertransformationen).

Beispiel. Wird eine Masse im Gravitationsfeld angehoben, so ist die verrichtete Arbeit nur abhängig von der überwundenen Höhe, der Weg, auf dem das geschehen ist, spielt keine Rolle. Das Gravitationsfeld ist **konservativ**.

Werden hingegen Reibungsverluste berücksichtigt, so wird die spezielle Wahl des Weges eine wichtige Rolle spielen.

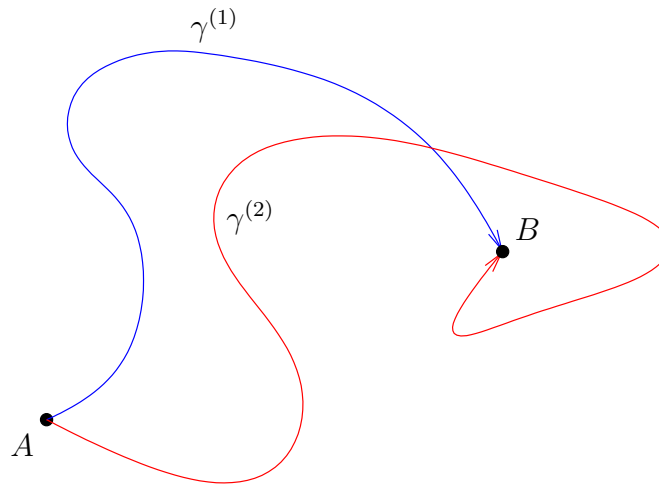


Abbildung 18.5: Wegabhängigkeit und geschlossene Wege.

Definition 18.1.2

Es sei $U \subset \mathbb{R}^m$ offen und $F: U \rightarrow \mathbb{R}^m$ sei ein stetiges Vektorfeld. Falls eine Funktion $\varphi: U \rightarrow \mathbb{R}$ existiert mit

$$F(\underline{\mathbf{x}}) = \nabla\varphi(\underline{\mathbf{x}}) \quad \text{für alle } \underline{\mathbf{x}} \in U ,$$

so heißt φ ein **Potential** von F . Das Vektorfeld F bezeichnet man in diesem Fall als **konservativ**.

Satz 18.1.1

Es sei $U \subset \mathbb{R}^m$ offen und $F: U \rightarrow \mathbb{R}^m$ sei ein stetiges, **konservatives** Vektorfeld. Ist $\gamma: [a, b] \rightarrow U$ eine (stückweise) glatte Kurve und ist φ ein **Potential** von F , so gilt

$$\int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle = \varphi(\gamma(b)) - \varphi(\gamma(a)) .$$

Beweisidee. Man berechne $\int_a^b \frac{d}{dt} \varphi(\gamma(t)) dt$ (\rightsquigarrow Übungen). \square

Beispiele.

i) Es sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$F(\underline{\mathbf{x}}) = \begin{pmatrix} x_1 + x_2^3 \\ 3x_1x_2^2 \\ 1 \end{pmatrix}.$$

Dann gilt $F = \nabla\varphi$ mit

$$\varphi(\underline{\mathbf{x}}) = \frac{1}{2}x_1^2 + x_1x_2^3 + x_3.$$

Zu beachten ist hier (nachrechnen!)

$$\operatorname{rot} F = \underline{\mathbf{0}}.$$

Allgemein wurde für beliebiges $\varphi: \mathbb{R}^3 \subset U \rightarrow \mathbb{R}$ von der Klasse C^2 in den Übungen nachgerechnet

$$\operatorname{rot}(\operatorname{grad} \varphi) = \underline{\mathbf{0}}.$$

Hat demnach ein (glattes) Vektorfeld $F: \mathbb{R}^3 \subset U \rightarrow \mathbb{R}^3$ ein Potential, so ist das **Verschwinden von $\operatorname{rot} F$ eine notwendige Bedingung dafür.**

ii) Zurück zum Beispiel des unendlich langen Leiters: Es sei wieder

$$B(\underline{\mathbf{x}}) = \begin{pmatrix} -\frac{x_2}{x_1^2 + x_2^2} \\ \frac{x_1}{x_1^2 + x_2^2} \\ 0 \end{pmatrix}.$$

Betrachtet man auf $U = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1 > 0\}$ die Funktion

$$\varphi(\underline{\mathbf{x}}) = \arctan\left(\frac{x_2}{x_1}\right),$$

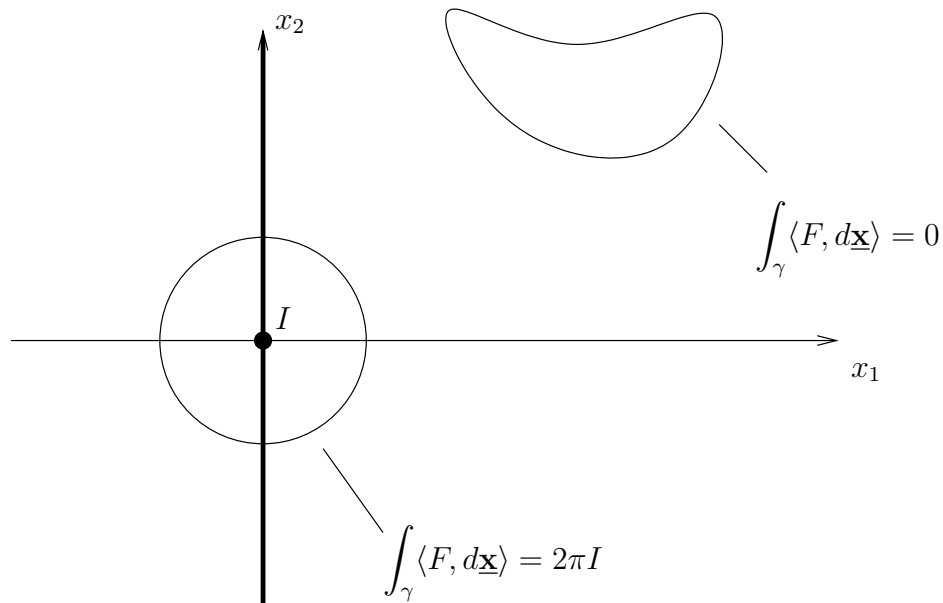


Abbildung 18.6: B kann kein “globales” Potential haben.

so erkennt man

$$\frac{\partial \varphi}{\partial x_1} = \frac{1}{1 + \frac{x_2^2}{x_1^2}} \left(-\frac{1}{x_1^2} x_2 \right) = -\frac{x_2}{x_1^2 + x_2^2},$$

$$\frac{\partial \varphi}{\partial x_2} = \frac{1}{1 + \frac{x_2^2}{x_1^2}} \frac{1}{x_1} = \frac{x_1}{x_1^2 + x_2^2},$$

$$\frac{\partial \varphi}{\partial x_3} = 0,$$

φ ist also auf U ein Potential von B .

Nach Satz 18.1.1 kann B aber kein Potential auf $\{\underline{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0\}$ haben (vgl. auch Abbildung 18.6).

Übung. Man berechne $\text{rot } B$ auf $\{\underline{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 \neq 0\}$ und interpretiere das Ergebnis.

Kapitel 19

Integralrechnung im \mathbb{R}^n

19.1 Das Riemannsche Integral auf quadrierbaren Mengen (Zellen; Jordansches Maß; Zerlegung von Zellen; Untersumme; Obersumme; Unterintegral; Oberintegral; Riemann-Integrierbarkeit; Volumenelement; Mehrfachintegral; charakteristische Funktion; kanonische Fortsetzung; Normalbereich; Cavalierisches Prinzip; Satz von Fubini)

Idee. Ist $f: M \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine beschränkte Funktion, so sollte das Integral $\int_M f \, dV$ ein Maß für das **Volumen des vom Graphen und von M berandeten Körpers sein**, so wie es in Abbildung 19.1 angedeutet ist.

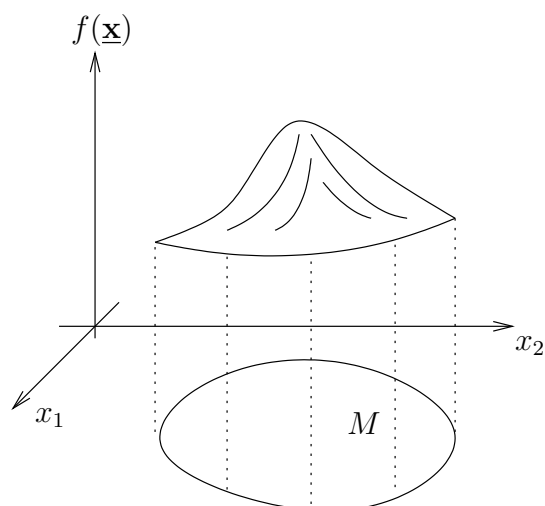


Abbildung 19.1: Zur Idee des Riemannschen Integrals.

Problem. Völlig analog zu den Ausführungen zu Beginn des Kapitels 12.1, HMI 2, zur eindimensionalen Integration gilt jedoch: Man hat zwar eine anschauliche Vorstellung des Begriffes “Volumen”, es ist jedoch nicht klar, wie das Volumen des vom Graphen berandeten Körpers überhaupt definiert ist, falls dieser Körper nicht gerade eine Vereinigung von Quadern (o.ä.) ist.

Vorgehensweise.

- i)* Zunächst wird der Integralbegriff über sogenannte **Zellen im \mathbb{R}^n** eingeführt, wobei Zellen den eindimensionalen Intervallbegriff verallgemeinern.
- ii)* Damit kann definiert werden, was das (**Jordansche Maß** einer Menge (der Urbildmenge $M \subset \mathbb{R}^n$ von f) ist. Dies war im eindimensionalen Fall nicht nötig, da dort im Wesentlichen nur über Intervalle zu integrieren ist, denen in natürlicher Weise eine Länge zugeordnet werden kann. Im mehrdimensionalen Fall hat man es mit einer wesentlich größeren Vielfalt von zu untersuchenden Mengen zu tun.
- iii)* Das Maß soll selbstverständlich bei einfachen geometrischen Objekten (beispielsweise Zellen im \mathbb{R}^n den elementargeometrisch bekannten Inhalt liefern.
- iv)* Schließlich wird das Integral über **quadrierbare Mengen im \mathbb{R}^n** definiert und man erhält eine präzise Formulierung der oben genannten Anschauung.
- v)* Die Idee zur Konstruktion in *i)* bis *iv)* ist: Man approximiere mit Quadern, anschließend wird zur Grenze übergegangen.

Der Zugang zum **mehrdimensionalen Riemannschen Integralbegriff auf Zellen** im \mathbb{R}^n ist völlig analog zur Theorie im Falle einer Variablen.

Im Folgenden sei stets $f: \mathbb{R}^n \supset M \rightarrow \mathbb{R}$ eine **skalare Funktion** veranschaulicht mit Hilfe des **Graphen von f** . Wie in Kapitel 12.1 sei f zunächst stets als **beschränkt** angenommen.

Wie bereits erwähnt tritt hier an die Stelle von Intervallen (und deren Zerlegungen):

Definition 19.1.1

i) Unter einer **Zelle** C im \mathbb{R}^n versteht man einen **abgeschlossenen Quader mit achsenparallelen Kanten**, d.h. eine Punktmenge der Form

$$C = I_1 \times I_2 \times \dots \times I_n =: [\underline{\mathbf{a}}, \underline{\mathbf{b}}] \subset \mathbb{R}^n \quad \text{mit:}$$

$$\underline{\mathbf{a}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \underline{\mathbf{b}} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

$$I_j = [a_j, b_j] \subset \mathbb{R}, \quad a_j < b_j, \quad j = 1, \dots, n.$$

ii) Es heißt $\mu(C) := |C| := \prod_{j=1}^n (b_j - a_j)$ der **Inhalt** (das **Maß**) von $[\underline{\mathbf{a}}, \underline{\mathbf{b}}]$.

iii) Eine **Menge** $\mathcal{C} = \{C^{(1)}, \dots, C^{(m)}\}$ von **“nicht überlappenden”** Zellen des \mathbb{R}^n mit

$$C = [\underline{\mathbf{a}}, \underline{\mathbf{b}}] = \bigcup_{k=1}^m C^{(k)}$$

heißt eine **Zerlegung** der Zelle C .

iv) Ist $f: \mathbb{R}^n \rightarrow \mathbb{R}$ *beschränkt* auf einer Zelle C und \mathcal{C} eine Zerlegung von C , so ist

$$\underline{m}_k := \inf\{f(\underline{\mathbf{x}}) : \underline{\mathbf{x}} \in C^{(k)}\},$$

$$\overline{m}_k := \sup\{f(\underline{\mathbf{x}}) : \underline{\mathbf{x}} \in C^{(k)}\},$$

$$\underline{S}_{\mathcal{C}}(f) := \sum_{k=1}^m \underline{m}_k \mu(C^{(k)}) \quad \text{die } \textit{Untersumme},$$

$$\overline{S}_{\mathcal{C}}(f) := \sum_{k=1}^m \overline{m}_k \mu(C^{(k)}) \quad \text{die } \textit{Obersumme}$$

von f zur Zerlegung \mathcal{C} .

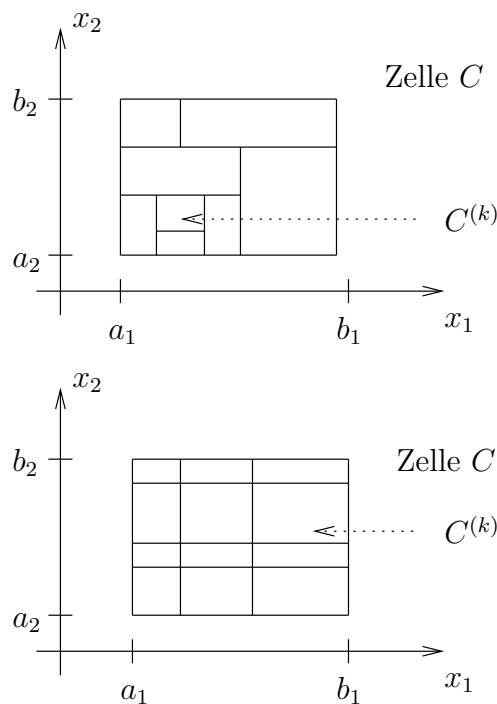


Abbildung 19.2: Zerlegungen einer Zelle.

Beispiele.

- i) Mögliche Zerlegungen von $[\underline{\mathbf{a}}, \underline{\mathbf{b}}]$ sind in Abbildung 19.2 dargestellt.
- ii) Unter- und Obersummen werden ganz analog zum Fall $n = 1$ konstruiert (siehe Abbildung 19.3)

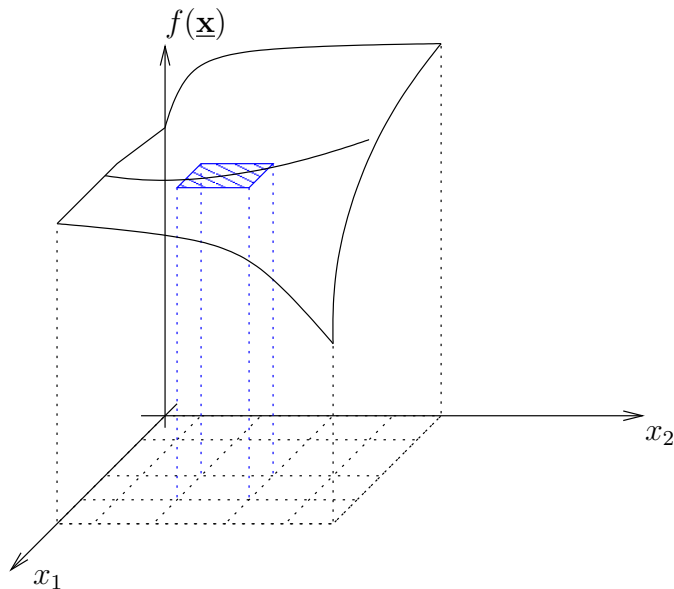


Abbildung 19.3: Zur Definition von Unter- und Obersummen.

Bemerkung. Alle bekannten Abschätzungen für Unter- und Obersummen aus dem Fall $n = 1$ übertragen sich analog.

Definition 19.1.2

i) Ist $f: C \rightarrow \mathbb{R}$ beschränkt, so sind das *Unterintegral* $\underline{\mathcal{I}}(f)$ und das *Oberintegral* $\overline{\mathcal{I}}(f)$ definiert als

$$\underline{\mathcal{I}}(f) := \sup \{ \underline{S}_C(f) : C \text{ ist eine Zerlegung von } C \},$$

$$\overline{\mathcal{I}}(f) := \inf \{ \overline{S}_C(f) : C \text{ ist eine Zerlegung von } C \}.$$

ii) f heißt (Riemann-) *integrierbar* auf C , falls gilt

$$\underline{\mathcal{I}}(f) = \overline{\mathcal{I}}(f) =: \mathcal{I}(f) .$$

iii) Die übliche Notation im Folgenden ist (falls f integrierbar ist):

$$\begin{aligned} \mathcal{I}(f) &= \int_C f \, dV = \int_C f(\underline{\mathbf{x}}) \, dV \\ &= \int \int \dots \int_C f(\underline{\mathbf{x}}) \, dx_1 \dots dx_n , \end{aligned}$$

$\mathcal{I}(f)$ heißt das *Riemannsches Integral* von f über die Zelle C , dV heißt das *Volumenelement*.

Bemerkung. Die Bemerkungen und Integrabilitätskriterien aus dem Fall $n = 1$ übertragen sich analog, beispielsweise gilt

Satz 19.1.1

Ist $f: \mathbb{R}^n \rightarrow \mathbb{R}$ *stetig* auf der Zelle C , so ist f integrierbar auf C .

Wie berechnet man Integrale im \mathbb{R}^n ?

Satz 19.1.2

Es sei o.E. $n = 2$ (für $n \geq 3$ gelten analoge Aussagen).
Es sei weiter $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ integrierbar auf der Zelle

$$C = \{ \underline{\mathbf{x}} \in \mathbb{R}^2 : a \leq x_1 \leq b, c \leq x_2 \leq d \} .$$

i) Falls für jedes *fixierte* $x_1 \in [a, b]$ das Integral $\int_c^d f(\underline{\mathbf{x}}) dx_2$ existiert, so existiert das *“iterierte”* Integral

$$\int_a^b \left[\int_c^d f(\underline{\mathbf{x}}) dx_2 \right] dx_1$$

und es gilt

$$\int_C f(\underline{\mathbf{x}}) dV = \int_a^b \left[\int_c^d f(\underline{\mathbf{x}}) dx_2 \right] dx_1 .$$

ii) Falls für jedes *fixierte* $x_2 \in [c, d]$ das Integral $\int_a^b f(\underline{\mathbf{x}}) dx_1$ existiert, so gilt die analoge Aussage.

Bemerkungen.

- i) Mit diesem Satz ist überhaupt erst die Notation als **Mehrfachintegral** $\int \int_C f(\underline{\mathbf{x}}) dx_1 dx_2$ gerechtfertigt.
- ii) Die Berechnung eines Integrals im \mathbb{R}^n wird also zurückgeführt auf zwei nacheinander auszuführende “eindimensionale” Integrationen. Damit **übertragen sich (unter den Voraussetzungen von Satz 19.1.2) die bekannten Eigenschaften aus der Integralrechnung einer Variablen** (Linearität ...)
- iii) Allein aus der Existenz der iterierten Integrale in Satz 19.1.2 **folgt nicht** die Existenz von $\int_C f(\underline{\mathbf{x}}) dV$ (Beispiel: \rightsquigarrow Übungen).
- iv) Es sei jetzt f beispielsweise stetig auf C , so dass alle auftretenden Integrale existieren. Dann gilt nach Satz 19.1.2 für das Zweifachintegral (mit der Notation $f(\underline{\mathbf{x}}) = f(x_1, x_2)$)

(a)

$$\int_C f(\underline{\mathbf{x}}) dx = \int_a^b \left[\int_c^d f(x_1, x_2) dx_2 \right] dx_1 ,$$

also wird für festes x_1 erst das Einfachintegral

$$\int_c^d f(x_1, x_2) \, dx_2 =: g(x_1)$$

bestimmt, dann die Funktion $g(x_1)$ bzgl. x_1 integriert.

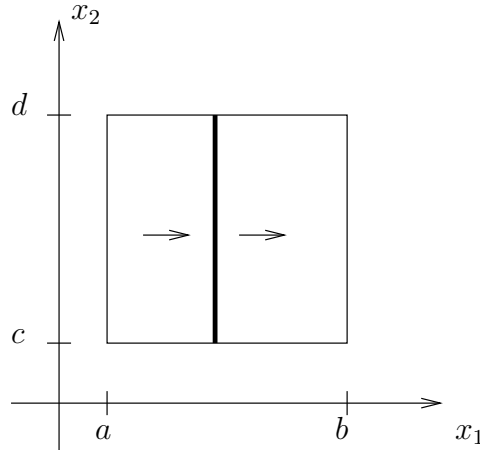


Abbildung 19.4: Eine Möglichkeit zur Berechnung eines Zweifachintegrals.

(b)

$$\int_C f(\underline{x}) \, dx = \int_c^d \left[\int_a^b f(x_1, x_2) \, dx_1 \right] dx_2 .$$

Hier ist die Situation umgekehrt (siehe Abbildung 19.5).

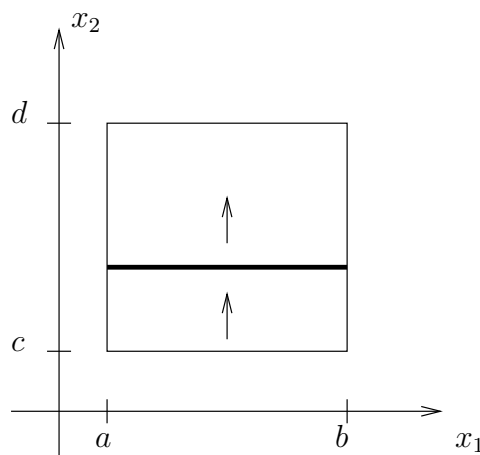


Abbildung 19.5: Eine zweite Möglichkeit zur Berechnung eines Zweifachintegrals.

Beispiele.

i) Es sei $C = \{\underline{\mathbf{x}} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 1 \leq x_2 \leq 2\}$,

$$f(x_1, x_2) = x_1^{x_2}.$$

Dann ist f stetig auf C , insbesondere sind die Voraussetzungen aus Satz 19.1.2 erfüllt. Es gilt

$$\int_C x_1^{x_2} dV = \int_1^2 \left[\int_0^1 x_1^{x_2} dx_1 \right] dx_2,$$

$$\int_0^1 x_1^{x_2} dx_1 = \frac{x_1^{x_2+1}}{x_2+1} \Big|_0^1 = \frac{1}{x_2+1}$$

und damit

$$\int_C x_1^{x_2} dV = \int_1^2 \frac{1}{x_2+1} dx_2 = \ln(1+x_2) \Big|_1^2 = \ln(3/2).$$

ii) Es sei jetzt $n = 3$,

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4\}.$$

Als stetiger Integrand f sei betrachtet

$$f(\underline{\mathbf{x}}) = x_1 + x_2 + x_3.$$

Mit der Notation

$$C_{x_2, x_3} = \left\{ \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} : 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4 \right\}$$

folgt

$$\int_C f(\underline{\mathbf{x}}) dV = \int_{C_{x_2, x_3}} \left[\underbrace{\int_0^2 (x_1 + x_2 + x_3) dx_1}_{=: g(x_2, x_3)} \right] dV(x_2, x_3),$$

wobei $dV(x_2, x_3)$ das zweidimensionale Volumenelement (bzgl. der Variablen x_2, x_3) bezeichne. Eine nochmalige Anwendung von Satz 19.1.2 liefert

$$\int_{C_{x_2, x_3}} g(x_2, x_3) dV(x_2, x_3) = \int_2^4 \left[\int_0^1 g(x_2, x_3) dx_2 \right] dx_3$$

und insgesamt:

$$\begin{aligned}
 \int_C f(\underline{\mathbf{x}}) \, dV &= \int_2^4 \left[\int_0^1 \left[\int_0^2 (x_1 + x_2 + x_3) \, dx_1 \right] dx_2 \right] dx_3 \\
 &= \int_2^4 \left[\int_0^1 \left[\frac{x_1^2}{2} + x_1(x_2 + x_3) \right]_0^2 dx_2 \right] dx_3 \\
 &= \int_2^4 \left[\int_0^1 [2 + 2x_2 + 2x_3] \, dx_2 \right] dx_3 \\
 &= \int_2^4 [2x_2(1 + x_3) + x_2^2]_0^1 dx_3 \\
 &= \int_2^4 [3 + 2x_3] \, dx_3 = [3x_3 + x_3^2]_2^4 = 18.
 \end{aligned}$$

iii) (\rightsquigarrow Übungen) Ist $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ von der Form

$$f(\underline{\mathbf{x}}) = f_1(x_1)f_2(x_2)$$

mit stetigen Funktionen $f_1, f_2: \mathbb{R} \rightarrow \mathbb{R}$ auf Intervallen $[a, b]$ bzw. $[c, d] \subset \mathbb{R}$, so ist $(C = [a, b] \times [c, d])$

$$\int_C f(\underline{\mathbf{x}}) \, dV = \left[\int_a^b f_1(x_1) \, dx_1 \right] \left[\int_c^d f_2(x_2) \, dx_2 \right].$$

Bemerkung. Analog zur letzten Bemerkung aus Kapitel 12.1 sind über die Definition des Riemannsches Integrals bereits Volumina “einfacher” Mengen im \mathbb{R}^{n+1} bestimmt und mit Hilfe von Satz 19.1.1 berechenbar. Beispielsweise findet man für das Volumen V der Menge (des Körpers)

$$K := \{\underline{\mathbf{x}} \in \mathbb{R}^3 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 0 \leq x_3 \leq x_1 + x_2\}$$

im \mathbb{R}^3 mit $C := \{\underline{\mathbf{x}} \in \mathbb{R}^2 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1\}$ und $f: C \rightarrow \mathbb{R}$, $f(\underline{\mathbf{x}}) = x_1 + x_2$:

$$V = \int_C f(\underline{\mathbf{x}}) \, dV = \int_0^2 \left[\int_0^1 (x_1 + x_2) \, dx_2 \right] dx_1 = 3.$$

Im \mathbb{R}^n will man aber nicht nur über Zellen integrieren!

Dazu muss zumindest einer großen Klasse von Mengen im \mathbb{R}^n (man beachte den Unterschied zu der letzten Bemerkung) ein Maß oder ein Inhalt zugewiesen werden. Um den Inhalt von recht allgemeinen Mengen mit Hilfe der Definition des Riemannsches Integrals über Zellen zu definieren, bedient man sich eines “Tricks” (vgl. auch Abbildung 19.6):

Definition 19.1.3¹

Eine beschränkte Menge M des \mathbb{R}^n heißt *quadrierbar* (oder *Jordan-messbar*), wenn ihre durch

$$\chi_M(\underline{\mathbf{x}}) := \begin{cases} 1 & \text{für } \underline{\mathbf{x}} \in M, \\ 0 & \text{für } \underline{\mathbf{x}} \in \mathbb{R}^n - M \end{cases}$$

definierte *charakteristische Funktion* $\chi_M: \mathbb{R}^n \rightarrow \mathbb{R}$ auf einer Zelle C mit $\overline{M} \subset \overset{\circ}{C}$ integrierbar ist. Es heißt

$$\mu(M) := \int_C \chi_M(\underline{\mathbf{x}}) \, dV$$

(*n*-dimensionaler) *Inhalt von M* oder *Volumen von M* oder (*n*-dimensionales) *Jordansches Maß von M*

Bemerkungen.

i) Die Definition des Inhalts einer quadrierbaren Menge ist **unabhängig von der speziellen Wahl der Zelle C** .

ii) Die Definition ist **konsistent** mit der des Maßes einer Zelle

$$\mu(C) = |C| = \prod_{j=1}^n (b_j - a_j) .$$

iii) **“Glatte”, beschränkte Mengen sind quadrierbar.** (Genauer Kriterien sind in der Literatur beschrieben, Quadrierbarkeit folgt etwa, falls ∂M eine Nullmenge ist oder falls ∂M lokal Graph einer stetigen Funktion ist.)

¹Zur Erinnerung: Nach Definition 6.4.4 bezeichnet \overline{M} den Abschluss von M , $\overset{\circ}{C}$ ist der offene Kern von C .

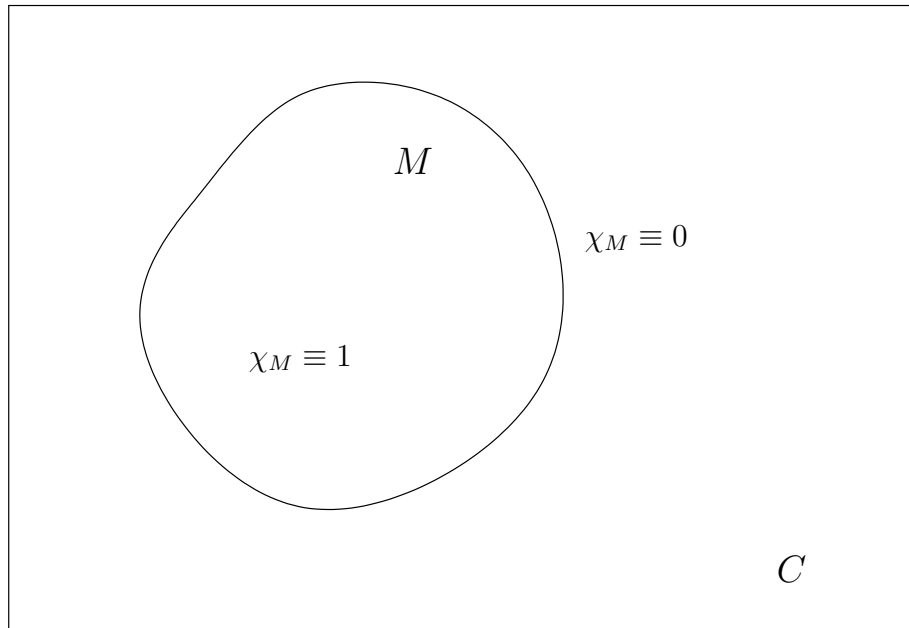


Abbildung 19.6: Zum Inhalt einer quadrierbaren Menge.

iv) Auf eine weitere Diskussion sei hier verzichtet, **die im Folgenden auftretenden Menge sind stets quadrierbar.**

Was ist $\int_M f(\underline{\mathbf{x}}) dV$ für eine quadrierbare Menge M ?

Definition 19.1.4

Es sei $M \subset \mathbb{R}^n$ quadrierbar und $f: M \rightarrow \mathbb{R}$ beschränkt. Dann ist die **kanonische Fortsetzung** der Funktion f von M auf den \mathbb{R}^n definiert via

$$f_M(\underline{\mathbf{x}}) := \begin{cases} f(\underline{\mathbf{x}}) & \text{für } \underline{\mathbf{x}} \in M, \\ 0 & \text{für } \underline{\mathbf{x}} \in \mathbb{R}^n - M. \end{cases}$$

Definition 19.1.5

Es sei $M \subset \mathbb{R}^n$ quadrierbar und $f: M \rightarrow \mathbb{R}$ beschränkt. Dann heißt f **auf M integrierbar**, wenn die kanonische Fortsetzung f_M von f auf einer (und damit auf jeder)

Zelle C des \mathbb{R}^n mit $\overline{M} \subset \overset{\circ}{C}$ integrierbar ist. Das *Riemannsche Integral* $\int_M f \, dV$ ist definiert als

$$\int_M f \, dV := \int_M f(\underline{\mathbf{x}}) \, dV := \int_C f_M(\underline{\mathbf{x}}) \, dV .$$

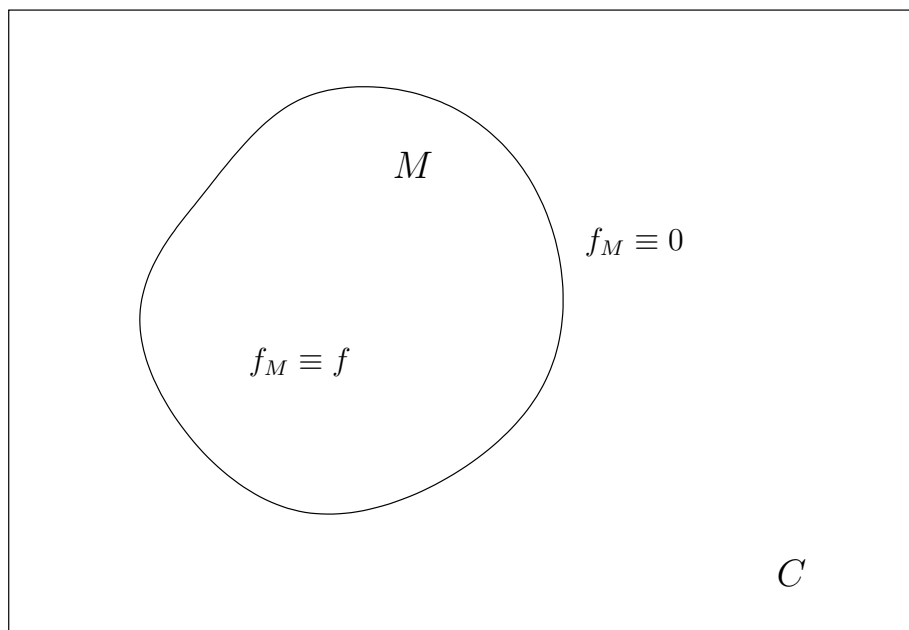


Abbildung 19.7: Zur Integrierbarkeit auf einer quadrierbaren Menge.

Bemerkung.

- i)* Diese Definition ist **unabhängig von der speziellen Wahl der Zelle**.
- ii)* Die Bedingung der Integrierbarkeit ist eine **Forderung an M und f** .
- iii)* Beschränkte, stetige Funktionen sind auf einer quadrierbaren Menge integrierbar (vgl. Satz 19.1.1).

Wieder übertragen sich die bekannten Eigenschaften, d.h. es gilt:

Satz 19.1.3

Es seien $f, g: M \subset \mathbb{R}^n \rightarrow \mathbb{R}$ integrierbar auf M .

i) Für beliebige $\alpha, \beta \in \mathbb{R}$ gilt

$$\int_M (\alpha f + \beta g) \, dV = \alpha \int_M f \, dV + \beta \int_M g \, dV .$$

ii) Es sind auch $fg, |f|, |g|$ integrierbar mit

$$\left| \int_M f \, dV \right| \leq \int_M |f| \, dV ,$$

$$\left| \int_M fg \, dV \right| \leq \sup_M |f| \int_M |g| \, dV ,$$

$$\left| \int_M fg \, dV \right|^2 \leq \int_M |f|^2 \, dV \int_M |g|^2 \, dV .$$

iii) Im Fall $f \leq g$ auf M gilt die Ungleichung

$$\int_M f \, dV \leq \int_M g \, dV .$$

iv) Es gilt insbesondere

$$\int_M 1 \, dV = |M| .$$

v) Ist f auch über einer quadrierbaren Menge $\tilde{M} \subset \mathbb{R}^n$ (definiert und) integrierbar, $M \cap \tilde{M} = \emptyset$, so ist

$$\int_{M \cup \tilde{M}} f \, dV = \int_M f \, dV + \int_{\tilde{M}} f \, dV .$$

Satz 19.1.3 erleichtert aber in den meisten Fällen nicht die **konkrete Berechnung** eines Integrals.

Gilt eine analoge Version von Satz 19.1.2?

Zur Klärung dieser Frage muss Definition 19.1.6 vorangestellt werden. Es ist nämlich zu beachten, dass bei quadrierbaren Mengen im Gegensatz zu einer Zelle **nicht klar ist, zwischen welchen Grenzen die Variablen laufen sollen**. Damit eine Argumentation ähnlich zu Satz 19.1.2 Sinn ergibt, muss die Klasse der betrachteten Mengen eingeschränkt werden.

Definition 19.1.6

i) Eine Menge $M \subset \mathbb{R}^2$ heißt **x_2 -projizierbar** oder **Normalbereich in x_2 -Richtung**, falls M von der Form ist

$$M = \{ \underline{\mathbf{x}} \in \mathbb{R}^2 : x_1 \in [a, b], \varphi_1(x_1) \leq x_2 \leq \varphi_2(x_1) \} ,$$

wobei φ_1, φ_2 zwei auf $[a, b]$ stetige Funktionen seien.

ii) Eine Menge $M \subset \mathbb{R}^2$ heißt **x_1 -projizierbar** oder **Normalbereich in x_1 -Richtung**, falls M von der Form ist

$$M = \{ \underline{\mathbf{x}} \in \mathbb{R}^2 : x_2 \in [c, d], \psi_1(x_2) \leq x_1 \leq \psi_2(x_2) \} ,$$

wobei ψ_1, ψ_2 zwei auf $[c, d]$ stetige Funktionen seien.

iii) Eine Menge $M \subset \mathbb{R}^n$ heißt *Normalbereich in x_j -Richtung*, $1 \leq j \leq n$, wenn sie die Form

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^n : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_j \leq \xi_2(\underline{\mathbf{y}})\}$$

hat, wobei

$$\underline{\mathbf{y}} := \begin{pmatrix} x_1 \\ \vdots \\ x_{j-1} \\ x_{j+1} \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n-1}$$

und wobei K eine kompakte, quadrierbare Menge des \mathbb{R}^{n-1} bezeichne. Die Funktionen $\xi_1, \xi_2: K \rightarrow \mathbb{R}$ seien stetig.

Bemerkungen.

- i) M.a.W. bedeutet etwa x_2 -Projizierbarkeit, dass M in $[a, b]$ von den Funktionen $\varphi_1(x_1)$ und $\varphi_2(x_1)$ begrenzt wird.
- ii) Insbesondere muss **auf dem ganzen Intervall $[a, b]$ die Bedingung $\varphi_1 \leq \varphi_2$ erfüllt sein** ($\xi_1 \leq \xi_2$ auf K analog).

Beispiele.

- i) Zweidimensionale Beispiele sind in den Abbildungen 19.8–19.11 dargestellt.
- ii) Man betrachte obiges dreidimensionales Beispiel der Zelle

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4\}.$$

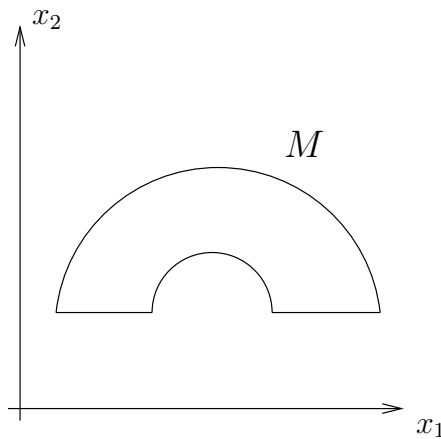


Abbildung 19.8: M ist x_2 projizierbar, nicht x_1 projizierbar.

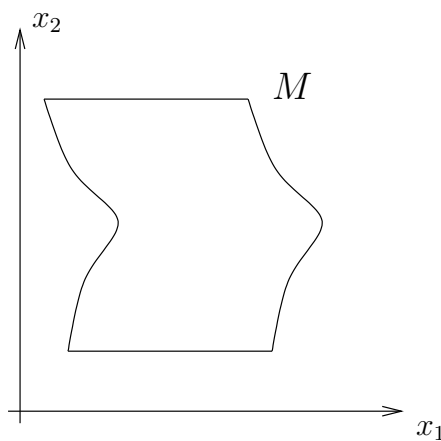


Abbildung 19.9: M ist x_1 projizierbar, nicht x_2 projizierbar.

Mit der Notation

$$\underline{\mathbf{y}} = \begin{pmatrix} x_2 \\ x_3 \end{pmatrix}, \quad \xi_1(\underline{\mathbf{y}}) \equiv 0, \quad \xi_2(\underline{\mathbf{y}}) \equiv 2,$$

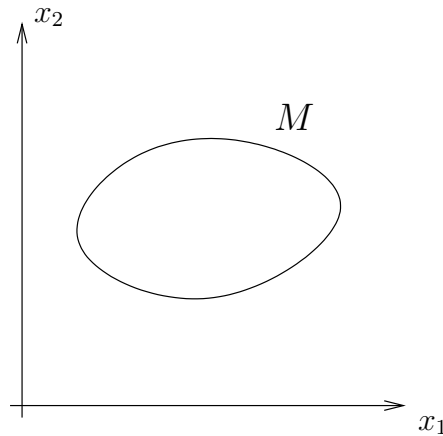
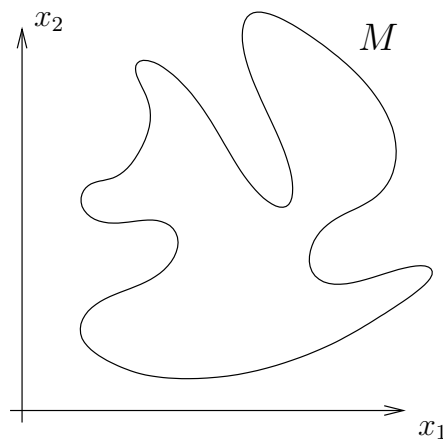
ist

$$K = C_{x_2, x_3} = \{\underline{\mathbf{y}} : 0 \leq x_2 \leq 1, 2 \leq x_3 \leq 4\}$$

und

$$C = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_1 \leq \xi_2(\underline{\mathbf{y}})\},$$

die Zelle ist also ein Normalbereich in x_1 -Richtung (selbstverständlich ebenso in x_2 - und x_3 -Richtung).

Abbildung 19.10: M ist x_1 projizierbar und x_2 projizierbar.Abbildung 19.11: M ist weder x_1 projizierbar noch x_2 projizierbar.

iii) Es sei $M \subset \mathbb{R}^3$ das "Raumstück", welches den beiden Zylindern $\{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1\}$ und $\{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_3^2 \leq 1\}$ gemeinsam ist,

$$M = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq 1\} \cap \{\underline{\mathbf{x}} \in \mathbb{R}^3 : x_1^2 + x_3^2 \leq 1\}.$$

Dann ist M Normalbereich in x_3 -Richtung (analog in x_2 -Richtung), denn mit

$$\underline{\mathbf{y}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in K := \{\underline{\mathbf{y}} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$$

und mit

$$\xi_1(\underline{\mathbf{y}}) = -\sqrt{1 - x_1^2}, \quad \xi_2(\underline{\mathbf{y}}) = \sqrt{1 - x_1^2} \quad \text{für alle } \underline{\mathbf{y}} \in K$$

ist

$$M = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{y}} \in K, \xi_1(\underline{\mathbf{y}}) \leq x_3 \leq \xi_2(\underline{\mathbf{y}}) \} .$$

iv) Es sei

$$M = \{ \underline{\mathbf{x}} \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1 \} .$$

Wieder ist M Normalbereich in x_3 -Richtung (andere Richtungen analog). Die Menge K findet man wie folgt. Zunächst muss auf K gelten: $x_1 \geq 0$ und $x_2 \geq 0$. Es ist weiter in K

$$\underline{\mathbf{y}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \xi_1(\underline{\mathbf{y}}) = 0, \quad \xi_2(\underline{\mathbf{y}}) = 1 - x_1 - x_2 .$$

Da die Bedingung $\xi_1(\underline{\mathbf{y}}) \leq \xi_2(\underline{\mathbf{y}})$ auf ganz K erfüllt sein muss, ergibt sich

$$K = \{ \underline{\mathbf{y}} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1 \} .$$

v) In den letzten beiden Beispielen schreibe man K selbst wieder als Normalbereich (Übung!).

Als Verallgemeinerung von Satz 19.1.2 kann schließlich mit Hilfe der Definition 19.1.6 das **Cavalierische² Prinzip** (Berechnung von Volumina, d.h. $f \equiv 1$) oder der **Satz von Fubini³** (der allgemeine Fall) formuliert werden.

Satz 19.1.4

Es sei $M \subset \mathbb{R}^n$ und $f: M \rightarrow \mathbb{R}$ stetig. Dann gilt mit den Bezeichnungen aus Definition 19.1.6:

i) Ist $n = 2$ und M ein Normalbereich in x_2 -Richtung, so gilt

$$\int_M f \, dV = \int_a^b \left[\int_{\varphi_1(x_1)}^{\varphi_2(x_1)} f(x_1, x_2) \, dx_2 \right] dx_1 .$$

²B. Cavalieri, 1598-1647; Mailand, Bologna.

³G. Fubini, 1879-1943; Catania, Genua, Turin, New York.

ii) Ist $n = 2$ und M ein Normalbereich in x_1 -Richtung, so gilt

$$\int_M f \, dV = \int_c^d \left[\int_{\psi_1(x_2)}^{\psi_2(x_2)} f(x_1, x_2) \, dx_1 \right] dx_2 .$$

iii) Im allgemeinen Fall $n \geq 2$, M Normalbereich in eine Richtung x_j , $1 \leq j \leq n$, gilt

$$\int_M f \, dV = \int_K \left[\int_{\xi_1(\underline{y})}^{\xi_2(\underline{y})} f(\underline{\mathbf{x}}) \, dx_j \right] dV_{n-1} ,$$

wobei dV_{n-1} das $(n - 1)$ -dimensionale Volumenelement auf $K \subset \mathbb{R}^{n-1}$ bezeichne.

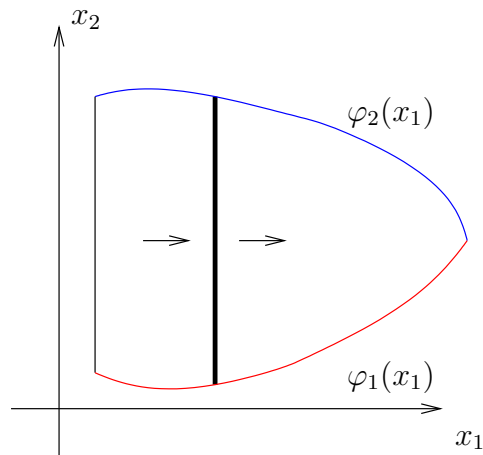
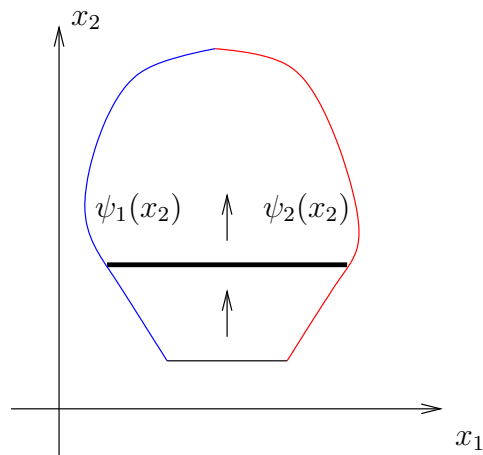
Bemerkungen.

- i) Der Satz ist ein wesentliches Hilfsmittel zur **konkreten Berechnung** von Integralen im \mathbb{R}^n als Mehrfachintegrale. Analog zu Satz 19.1.2 ist die anschauliche Vorstellung in den Abbildungen 19.12 und 19.13 wiedergegeben.
- ii) Ist M in verschiedene Richtungen ein Normalbereich, so kann natürlich die “einfachste” Berechnungsmethode gewählt werden.
- iii) Ist M kein Normalbereich, so kann versucht werden, **M in endlich viele Normalbereiche zu zerlegen**, um Satz 19.1.4 anzuwenden.

Beispiele.

- i) Es sei $R > 0$ fixiert und $M := \{\underline{\mathbf{x}} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq R^2, x_2 \geq 0\}$. Der Integrand $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sei gegeben durch

$$f(\underline{\mathbf{x}}) = x_1^2 x_2 .$$


 Abbildung 19.12: Iterierte Integration für einen x_2 -Normalbereich.

 Abbildung 19.13: Iterierte Integration für einen x_1 -Normalbereich.

1^{te} Möglichkeit: Man rechne nach Satz 19.1.4, i):

$$\begin{aligned}
 \int_M x_1^2 x_2 \, dV &= \int_{-R}^R \left[\int_0^{\sqrt{R^2 - x_1^2}} x_1^2 x_2 \, dx_2 \right] dx_1 \\
 &= \int_{-R}^R \left[\frac{x_1^2 x_2^2}{2} \right]_0^{\sqrt{R^2 - x_1^2}} dx_1 \\
 &= \frac{1}{2} \int_{-R}^R x_1^2 (R^2 - x_1^2) dx_1 \\
 &= \frac{1}{2} \left[\frac{x_1^3}{3} R^2 - \frac{x_1^5}{5} \right]_{-R}^R = \frac{2}{15} R^5.
 \end{aligned}$$

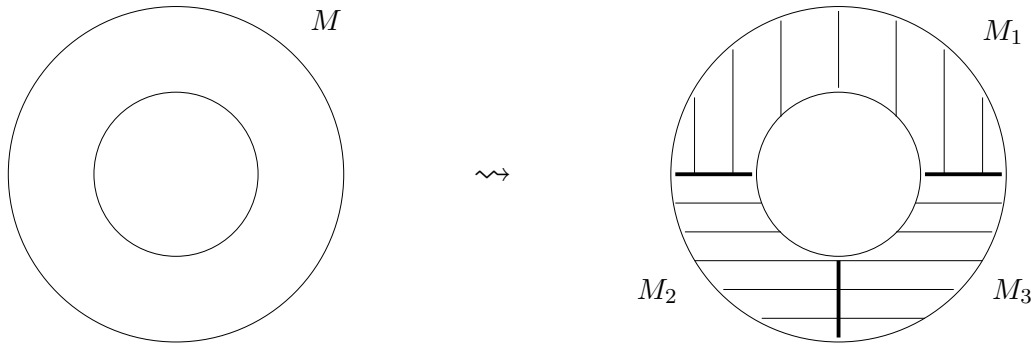
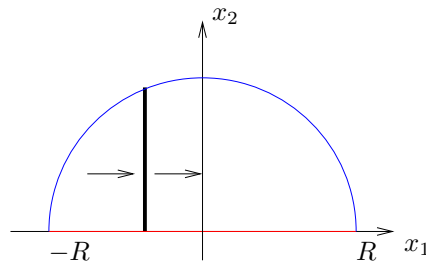


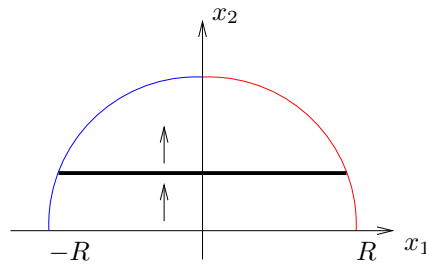
Abbildung 19.14: Eine Zerlegung in Normalbereiche.

Die Vorgehensweise ist in Abbildung 19.15 skizziert.

Abbildung 19.15: 1^{te} Möglichkeit.

2^{te} Möglichkeit: Man rechne nach Satz 19.1.4, *ii*) (vgl. Abbildung 19.16):

$$\begin{aligned}
 \int_M x_1^2 x_2 \, dV &= \int_0^R \left[\int_{-\sqrt{R^2-x_2^2}}^{\sqrt{R^2-x_2^2}} x_1^2 x_2 \, dx_1 \right] dx_2 \\
 &= \int_0^R \left[\frac{x_1^3}{3} x_2 \right]_{-\sqrt{R^2-x_2^2}}^{\sqrt{R^2-x_2^2}} dx_2 \\
 &= \frac{1}{3} \int_0^R 2x_2 (R^2 - x_2^2)^{\frac{3}{2}} dx_2 \\
 &= \frac{1}{3} \left[\left(-\frac{2}{5} \right) (R^2 - x_2^2)^{\frac{5}{2}} \right]_0^R = \frac{2}{15} R^5.
 \end{aligned}$$


 Abbildung 19.16: 2^{te} Möglichkeit.

ii) Gesucht: Das Volumen V des Tetraeders T mit den Ecken $O = \underline{\mathbf{0}}$,

$$A = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}, \quad a, b, c > 0 \text{ fixiert.}$$

Bemerkung. Zur Volumenbestimmung kann man einerseits den Graphen von

$$f(x_1, x_2) = c \left(1 - \frac{x_1}{a} - \frac{x_2}{b} \right)$$

über dem Dreieck mit den Ecken O, A, B in der (x_1, x_2) -Ebene betrachten (vgl. Abbildung 19.17). Ebenso kann T als Teilmenge des \mathbb{R}^3 aufgefasst werden und

$$\int_T 1 \, dV$$

berechnet werden, wobei dV das dreidimensionale Volumenelement bezeichne. Hier wird die zweite Sichtweise gewählt, die erste führt natürlich auf den gleichen Ansatz (ausprobieren!).

Um die richtigen Integrationsgrenzen zu finden, beachtet man, dass die Punkte A, B, C in der durch

$$\frac{x_1}{a} + \frac{x_2}{b} + \frac{x_3}{c} = 1$$

gegebenen Ebene liegen (vgl. wieder 19.17). Das gesuchte Volumen berechnet sich damit zu

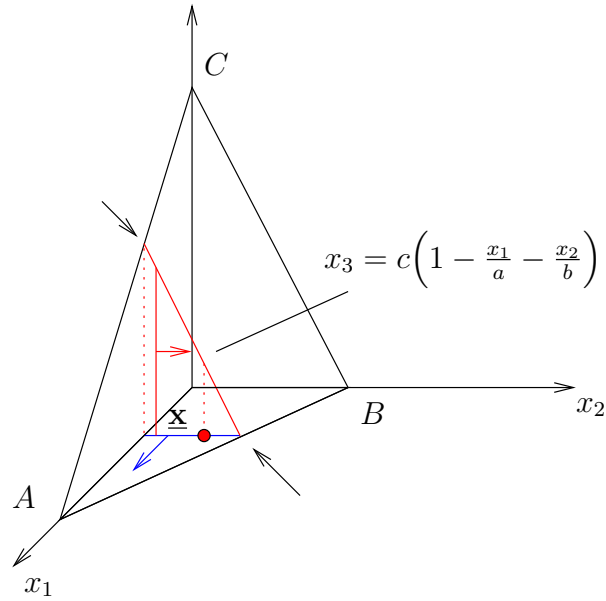


Abbildung 19.17: Zur Volumenberechnung des Tetraeders.

$$\begin{aligned}
 V &= \int_T 1 \, dV \\
 &= \int_0^a \left[\int_0^{b(1-x_1/a)} \left[\int_0^{c(1-x_1/a-x_2/b)} dx_3 \right] dx_2 \right] dx_1 \\
 &= \int_0^a \left[\int_0^{b(1-x_1/a)} c \left(1 - \frac{x_1}{a} - \frac{x_2}{b} \right) dx_2 \right] dx_1 \\
 &= \int_0^a c \left[\left(1 - \frac{x_1}{a} \right) x_2 - \frac{1}{2} \frac{x_2^2}{b} \right]_0^{b(1-x_1/a)} dx_1 \\
 &= \int_0^a \frac{bc}{2} \left(1 - \frac{x_1}{a} \right)^2 dx_1 = \frac{bc}{2} \left[-\frac{a}{3} \left(1 - \frac{x_1}{a} \right) \right]_0^a \\
 &= \frac{abc}{6}.
 \end{aligned}$$

19.2 Der Transformationssatz (Koordinatentransformation; Diffeomorphismus)

In diesem Abschnitt soll der Frage nachgegangen werden:

Kann analog zum Fall einer Funktion einer Variablen die Integration in bestimmten Situationen durch eine (**Substitution**) **Transformation** (d.h. durch eine geeignete **Koordinatentransformation**) vereinfacht werden?

Betrachtet sei dazu eine Abbildung g einer offenen Menge $U \subset \mathbb{R}^n$ in eine offene Menge $V \subset \mathbb{R}^n$,

$$U \ni \underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \mapsto g(\underline{\mathbf{u}}) = \underline{\mathbf{v}} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in V.$$

Die Abbildung g sei ein **Diffeomorphismus**, d.h. **bijektiv** und sowohl g als auch g^{-1} seien **stetig differenzierbar**. Insbesondere gelte für alle $\underline{\mathbf{u}} \in U$

$$\det Dg(\underline{\mathbf{u}}) = \left| \begin{pmatrix} \frac{\partial g_1}{\partial u_1} & \cdots & \frac{\partial g_1}{\partial u_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial u_1} & \cdots & \frac{\partial g_n}{\partial u_n} \end{pmatrix} \right| (\underline{\mathbf{u}}) \neq 0.$$

Unter dieser **Voraussetzung** gilt der **Transformationssatz**:

Satz 19.2.1

Ist $K \subset V$ **kompakt**, g wie oben und ist $f: V \rightarrow \mathbb{R}$ **stetig**, so gilt

$$\int_K f(\underline{\mathbf{v}}) \, dV(\underline{\mathbf{v}}) = \int_{g^{-1}(K)} f(g(\underline{\mathbf{u}})) |\det Dg(\underline{\mathbf{u}})| \, dV(\underline{\mathbf{u}}).$$

Dabei ist

$$g^{-1}(K) = \{\underline{\mathbf{u}} \in U : g(\underline{\mathbf{u}}) \in K\}$$

und im Volumenelement ist angedeutet, bzgl. welcher Variablen zu integrieren ist.

Beweisidee. Man untersuche die Frage, wie sich Volumina unter **linearen Abbildungen** $L: \mathbb{R}^n \rightarrow \mathbb{R}^n$ transformieren. \square

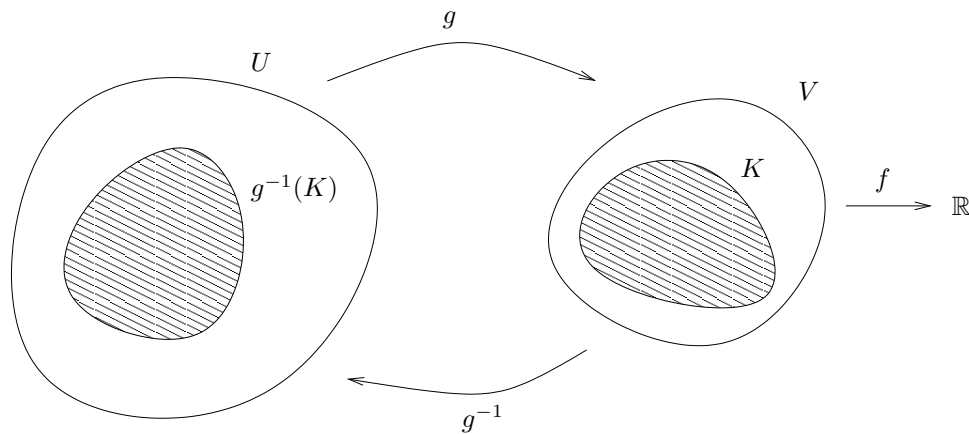


Abbildung 19.18: Zum Transformationssatz.

Bemerkungen.

- i)* Satz 19.2.1 hat formal die gleiche Gestalt wie die **Substitutionsregel** im Fall $n = 1$. Allerdings ist hier der **Betrag** der Jacobischen Funktionaldeterminante zu betrachten.
- ii)* Man beachte, dass die Kompaktheit von K zusammen mit der Stetigkeit von f insbesondere die Integrierbarkeit von f impliziert (warum ist das richtig?).

Beispiele.

- i)* Im \mathbb{R}^2 seien Polarkoordinaten betrachtet: Es ist

$$\underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} r \\ \varphi \end{pmatrix}, \quad 0 < r < \infty, \quad 0 < \varphi < 2\pi,$$

sowie

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g(r, \varphi) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}.$$

Die Menge

$$U = \left\{ \begin{pmatrix} r \\ \varphi \end{pmatrix} : 0 < r < \infty, 0 < \varphi < 2\pi \right\}$$

ist offen, die Bildmenge

$$V = \{\underline{x} \in \mathbb{R}^2 : x_2 \neq 0 \text{ falls } x \geq 0\}$$

ist die “aufgeschlitzte” (x_1, x_2) -Ebene.

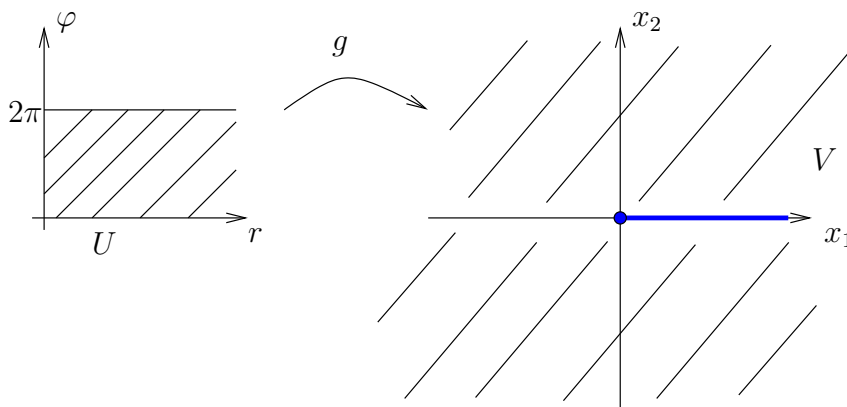


Abbildung 19.19: Polarkoordinaten im \mathbb{R}^2 .

Die Funktionaldeterminante der Transformation g ist

$$|\det Dg| = \left| \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{pmatrix} \right| = r \neq 0.$$

ii) Es sei $K = \{\underline{x} \in \mathbb{R}^2 : 1 \leq x_1^2 + x_2^2 \leq 4\}$. Gesucht ist

$$\int_K \sqrt{x_1^2 + x_2^2} \, dV.$$

An dieser Stelle sei exemplarisch etwas genauer als “üblich” argumentiert, zu beachten ist nämlich: K ist nicht vollständig in

der offenen Menge V aus i) enthalten.

Deshalb betrachtet man zu $0 < \alpha \ll 2\pi$ die in Abbildung 19.20 skizzierte kompakte Menge K_α :

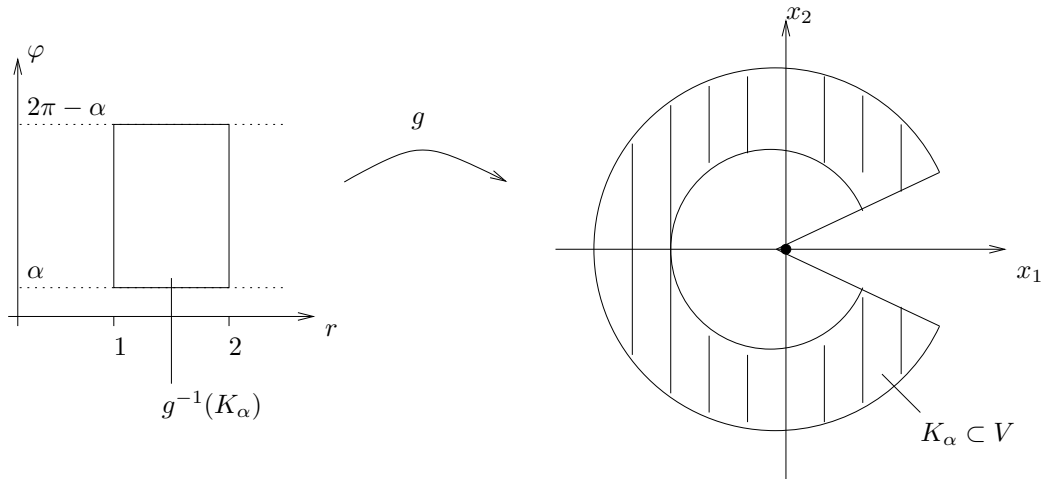


Abbildung 19.20: Die Menge K_α .

Als Abbildung $g: g^{-1}(K_\alpha) \rightarrow K_\alpha$ können jetzt Polarkoordinaten eingeführt werden, nach Satz 19.2.1 folgt

$$\begin{aligned} & \int_{K_\alpha} \sqrt{x_1^2 + x_2^2} \, dV(\underline{\mathbf{x}}) \\ &= \int_{g^{-1}(K_\alpha)} \sqrt{r^2 \cos^2(\varphi) + r^2 \sin^2(\varphi)} \, r \, dV(r, \varphi) \\ &= \int_\alpha^{2\pi-\alpha} \left[\int_1^2 r^2 \, dr \right] \, d\varphi = \int_\alpha^{2\pi-\alpha} \frac{7}{3} \, d\varphi = \frac{14}{3}(\pi - \alpha). \end{aligned}$$

Nun gilt

$$\begin{aligned} \int_K \sqrt{x_1^2 + x_2^2} \, dV(\underline{\mathbf{x}}) &= \int_{K_\alpha} \sqrt{x_1^2 + x_2^2} \, dV(\underline{\mathbf{x}}) \\ &\quad + \int_{K-K_\alpha} \sqrt{x_1^2 + x_2^2} \, dV(\underline{\mathbf{x}}) \end{aligned}$$

und weiterhin wegen $\sqrt{x_1^2 + x_2^2} \leq 2$ auf K

$$\int_{K-K_\alpha} \sqrt{x_1^2 + x_2^2} \, dV(\mathbf{x}) \leq 2|K - K_\alpha| \xrightarrow{\alpha \rightarrow 0} 0,$$

also liefert der Grenzwert $\alpha \rightarrow 0$

$$\int_K \sqrt{x_1^2 + x_2^2} \, dV = \frac{14}{3}\pi.$$

iii) Kugelkoordinaten im \mathbb{R}^3 : Hier wird die Transformation betrachtet

$$\underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} \xrightarrow{g} \underline{\mathbf{v}} = \underline{\mathbf{x}} = \begin{pmatrix} r \cos(\varphi) \cos(\theta) \\ r \sin(\varphi) \cos(\theta) \\ r \sin(\theta) \end{pmatrix}.$$

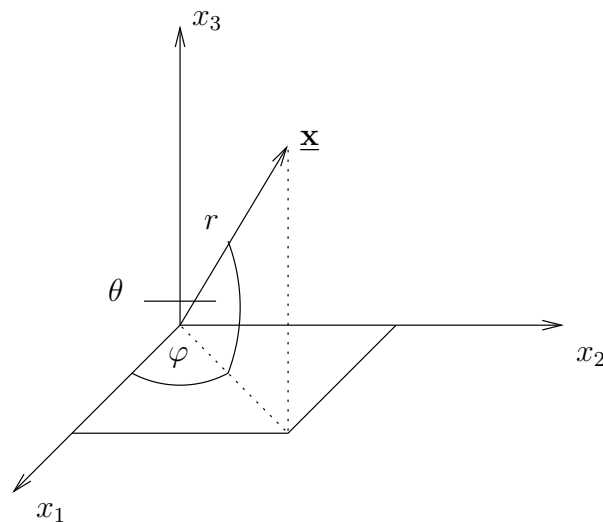


Abbildung 19.21: Kugelkoordinaten im \mathbb{R}^3 .

Es ist

$$U = \left\{ \begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} : r > 0, 0 < \varphi < 2\pi, -\frac{\pi}{2} < \theta < \frac{\pi}{2} \right\},$$

das Bild V ist der \mathbb{R}^3 ohne die nicht-negative x_1 -Achse und ohne die x_3 -Achse. Es gilt (Entwicklung nach der dritten Zeile, vgl. Satz 8.6.2)

$$\begin{aligned}
& |\det Dg| \\
&= \left| \begin{pmatrix} \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \cos(\theta) & -r \cos(\varphi) \sin(\theta) \\ \sin(\varphi) \cos(\theta) & r \cos(\varphi) \cos(\theta) & -r \sin(\varphi) \sin(\theta) \\ \sin(\theta) & 0 & r \cos(\theta) \end{pmatrix} \right| \\
&= \sin(\theta) (r^2 \sin^2(\varphi) \cos(\theta) \sin(\theta) + r^2 \cos^2(\varphi) \sin(\theta) \cos(\theta)) \\
&\quad + r \cos(\theta) (r \cos^2(\varphi) \cos^2(\theta) + r \sin^2(\varphi) \cos^2(\theta)) \\
&= \sin(\theta) r^2 \sin(\theta) \cos(\theta) + r \cos(\theta) r \cos^2(\theta) \\
&= r^2 \cos(\theta) \neq 0 \quad \text{für } 0 < r, \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}.
\end{aligned}$$

iv) Man betrachte jetzt

$$K = \{\mathbf{x} \in \mathbb{R}^3 : x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1^2 + x_2^2 + x_3^2 \leq 1\}$$

und die Funktion

$$f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Mittels der Transformation auf Kugelkoordinaten (ohne die genaue Argumentation wie in ii)) folgt

$$\begin{aligned}
\int_K \sqrt{x_1^2 + x_2^2 + x_3^2} \, dV(\mathbf{x}) &= \int_0^1 \left[\int_0^{\frac{\pi}{2}} \left[\int_0^{\frac{\pi}{2}} r r^2 \cos(\theta) \, d\varphi \right] d\theta \right] dr \\
&= \frac{\pi}{2} \int_0^1 \left[\int_0^{\frac{\pi}{2}} r^3 \cos(\theta) \, d\theta \right] dr \\
&= \frac{\pi}{2} \int_0^1 r^3 \, dr = \frac{\pi}{8}.
\end{aligned}$$

Bemerkung. Oft sind etwa auch Zylinderkoordinaten im \mathbb{R}^3 hilfreich (\rightsquigarrow Übungen).

19.3 Uneigentliche Integrale (reguläre Ausschöpfung; kompakt enthalten; Figur)

Wie im Fall $n = 1$ (Kapitel 12.4) sollen jetzt die **Restriktionen** “ M beschränkte Menge” und “ f beschränkt” fallen gelassen werden, die bisher immer vorausgesetzt waren.

M.a.W.: Das **uneigentliche Integral** $\int_{\Omega} f \, dV$ soll für **stetige Funktionen auf beliebigen offenen Mengen** Ω definiert werden (offene Mengen müssen nicht beschränkt sein, stetige Funktionen auf offenen Mengen müssen nicht beschränkt sein).

Definition 19.3.1

*Es seien $M, M_1, M_2, \dots, M_j, \dots$ Teilmengen des \mathbb{R}^n . Man nennt die Folge $\{M_j\}$ eine **reguläre Ausschöpfung** von M , Notation $M_j \nearrow M$, wenn gilt*

$$i) \overline{M}_j \in \overset{\circ}{M}_{j+1} \text{ für alle } j \in \mathbb{N},$$

$$ii) M = \bigcup_{j=1}^{\infty} M_j.$$

*Dabei bedeutet $A \in B$ (A **kompakt enthalten** in B): A ist beschränkt und es gilt $\overline{A} \subset B$.*

Beispiel. Man betrachte den \mathbb{R}^n und die Kugeln $B_k(0)$, $k \in \mathbb{N}$.

Bemerkung. Für jede offene Menge $\Omega \subset \mathbb{R}^n$ gibt es eine reguläre Ausschöpfung $F_j \nearrow \Omega$ durch sogenannte **Figuren** F_j (eine Figur ist die Vereinigung endlich vieler Zellen, die sich nicht überlappen).

Definition 19.3.2

*i) Es sei Ω eine offene Menge des \mathbb{R}^n , f stetig auf Ω und für **jede** reguläre Ausschöpfung $M_j \nearrow \Omega$ von Ω durch quadrierbare Mengen M_j sei die Folge der Zahlen $\int_{M_j} f \, dV$ konvergent. Dann heißt*

$$\int_{\Omega} f(\underline{\mathbf{x}}) \, dV := \lim_{j \rightarrow \infty} \int_{M_j} f(\underline{\mathbf{x}}) \, dV$$

das *uneigentliche Integral* von f über Ω , das *uneigentliche Integral existiert oder konvergiert* in diesem Fall.

ii) Das uneigentliche Integral heißt absolut konvergent, wenn das Integral $\int_{\Omega} |f(\underline{\mathbf{x}})| \, dV$ konvergiert.

Bemerkungen.

- i) Die Definition des uneigentlichen Integrals ist unabhängig von der speziellen Wahl der Ausschöpfung.*
- ii) Die Definition ist konsistent mit den vorherigen Abschnitten.*
- iii) Natürlich übertragen sich Eigenschaften wie die Linearität des Integrals sowie elementare Abschätzungen.*
- iv) Das uneigentliche Integral $\int_{\Omega} f(\underline{\mathbf{x}}) \, dV$ konvergiert, falls es absolut konvergiert.*

Wie kann in der Praxis die Konvergenz des uneigentlichen Integrals nachgeprüft werden?

Satz 19.3.1

Das uneigentliche Integral $\int_{\Omega} f(\underline{\mathbf{x}}) \, dV$ einer stetigen Funktion f ist genau dann absolut konvergent, wenn es eine Konstante $c \geq 0$ gibt, so dass

$$\int_M |f(\underline{\mathbf{x}})| \, dV \leq c$$

für alle quadrierbaren Mengen $M \in \Omega$ gilt.

Beispiel. Es sei $\Omega = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : |\underline{\mathbf{x}}| > 1\}$. Es sei $\alpha \in \mathbb{R}$ fixiert, $f(\underline{\mathbf{x}}) = |\underline{\mathbf{x}}|^{-\alpha}$, $R > 1$ und $\Omega_R = \Omega \cap B_R(0)$. Dann gilt (mit Hilfe der Transformation auf Kugelkoordinaten)

$$\begin{aligned} \int_{\Omega_R} f(\underline{\mathbf{x}}) \, dV &= \int_{\Omega_R} |\underline{\mathbf{x}}|^{-\alpha} \, dV \\ &= \int_1^R \left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[\int_0^{2\pi} r^{-\alpha} r^2 \cos(\theta) \, d\varphi \right] d\theta \right] dr \\ &= 2\pi \int_1^R r^{2-\alpha} [\sin(\theta)]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} dr = 4\pi \int_1^R r^{2-\alpha} \, dr. \end{aligned}$$

Die rechte Seite strebt gegen unendlich, falls $\alpha \leq 3$ ist, das uneigentliche Integral konvergiert in diesem Fall nicht. (Genau gesagt: Ist $R_j = j$ und $M_j = \Omega_{R_j}$, so ist eine reguläre Ausschöpfung von Ω gefunden, für die $\int_{M_j} f(\underline{\mathbf{x}}) \, dV$ nicht konvergiert, das uneigentliche Integral ist nach Definition 19.3.2, *i*), nicht konvergent.)

Ist $\alpha > 3$, so gilt

$$\int_{\Omega_R} f(\underline{\mathbf{x}}) \, dV = 4\pi \left[\frac{1}{3-\alpha} r^{3-\alpha} \right]_1^R = \frac{4\pi}{\alpha-3} (1 - R^{3-\alpha}) < \frac{4\pi}{\alpha-3}.$$

Jede quadrierbare Menge $M \Subset \Omega$ ist aber für hinreichend großes R in Ω_R enthalten (quadrierbare Mengen sind per definitionem beschränkt), man hat

$$\int_M |f(\underline{\mathbf{x}})| \, dV \leq \frac{4\pi}{\alpha-3}$$

für jede quadrierbare Menge $M \Subset \Omega$ im Fall $\alpha > 3$. Nach Satz 19.3.1 ist in diesem Fall das uneigentliche Integral konvergent. Zur Berechnung des Integrals reicht es nach Definition 19.3.2 dann, eine spezielle reguläre Ausschöpfung von Ω zu betrachten (etwa M_j wie oben) und man erhält

$$\int_{\Omega} |\underline{\mathbf{x}}|^{-\alpha} \, dV = \frac{4\pi}{\alpha-3}.$$

Teil VIII

Integralsätze der Vektoranalysis

Kapitel 20

Der Gaußsche Integralsatz

Der Gaußsche Integralsatz ist einer der bedeutendsten Sätze in der Analysis. Es wird etwa im Verlaufe dieses Kapitels geklärt werden, wie eine [partielle Integration](#) im Falle von Funktionen mehrerer Veränderlicher zu verstehen ist. Ebenso wird deutlich werden, warum die Divergenz eines Vektorfeldes ein Maß für Quellen und Senken ist (man vergleiche die Bemerkungen nach Definition 17.2.5).

Der Gaußsche Integralsatz in der Ebene kann mit den bisher bekannten Begriffen vorgestellt werden, zum Verständnis des Gaußschen Integralsatzes im \mathbb{R}^3 sind hingegen einige geometrische Betrachtungen voranzustellen.

20.1 Der Gaußsche Integralsatz in der Ebene (positiv orientierter Rand; Greenscher Bereich; äußere Einheitsnormale; Fluss eines Vektorfeldes)

Im Fall einer Veränderlichen besagt der Hauptsatz der Differential- und Integralrechnung (12.2.1), dass das bestimmte Integral einer Funktion f als Differenz der Funktionswerte einer Stammfunktion **in den Randpunkten** gegeben ist.

In diesem Paragraphen wird **ein zweidimensionales Integral über eine Divergenz als Integral über eine Randkurve** geschrieben werden, welches als [Flussintegral](#) zu interpretieren ist.

Zur Formulierung dieses Satzes ist noch eine Vorzeichenbedingung zu

klären: Ist $M \subset \mathbb{R}^2$ ein Normalbereich etwa in x_2 -Richtung (vgl. Definition 19.1.6), so kann **der Rand** von M wie in Abbildung 20.1 angedeutet durch eine stückweise glatte Kurve γ parametrisiert werden, die Spur der Kurve ist $\text{spur}\gamma = \Gamma^{(1)} \cup \Gamma^{(2)} \cup \Gamma^{(3)} \cup \Gamma^{(4)}$.

Es wird stets angenommen, dass diese Kurve regulär ($\gamma' \neq 0$ auf den Teilkurven) sei und dass sie so orientiert sei, dass M beim Durchlaufen der Kurve zur Linken liege, d.h. ∂M sei **positiv orientiert**.

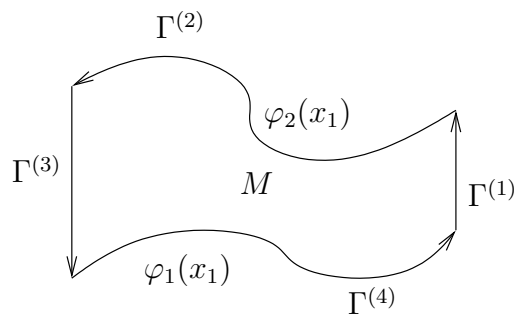


Abbildung 20.1: Der Rand ∂M .

Die Idee, den Gaußschen Integralsatz aus dem Satz von Fubini (Satz 19.1.4) und dem Hauptsatz der Differential- und Integralrechnung (Satz 12.2.1) abzuleiten, sei anhand des einfachsten Beispiels angedeutet.

Beispiel. Man betrachte das Quadrat $Q := \{\underline{x} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$ und überlege sich, dass es von den Kurven (jeweils auf $[0, 1]$

definiert)

$$\gamma^{(1)} : t \mapsto \begin{pmatrix} t \\ 0 \end{pmatrix},$$

$$\gamma^{(2)} : t \mapsto \begin{pmatrix} 1 \\ t \end{pmatrix},$$

$$\gamma^{(3)} : t \mapsto \begin{pmatrix} 1-t \\ 1 \end{pmatrix},$$

$$\gamma^{(4)} : t \mapsto \begin{pmatrix} 0 \\ 1-t \end{pmatrix}$$

positiv orientiert berandet wird. Für ein glattes Vektorfeld $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist nach den oben genannten Sätzen (mit der Notation $F(\underline{\mathbf{x}}) = F(x_1, x_2)$)

$$\begin{aligned} \int_Q \frac{\partial F_2}{\partial x_1} dV &= \int_0^1 \left[\int_0^1 \frac{\partial F_2}{\partial x_1} dx_1 \right] dx_2 \\ &= \int_0^1 [F_2(1, x_2) - F_2(0, x_2)] dx_2 \\ &= \int_{\gamma^{(2)}} \langle F, d\underline{\mathbf{x}} \rangle + \int_{\gamma^{(4)}} \langle F, d\underline{\mathbf{x}} \rangle. \end{aligned}$$

Analog berechnet sich

$$- \int_Q \frac{\partial F_1}{\partial x_2} dV = \int_{\gamma^{(1)}} \langle F, d\underline{\mathbf{x}} \rangle + \int_{\gamma^{(3)}} \langle F, d\underline{\mathbf{x}} \rangle.$$

Bezeichnet γ die aus $\gamma^{(1)}$ bis $\gamma^{(4)}$ zusammengesetzte stückweise glatte Kurve, so ist gezeigt

$$\int_Q \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = \int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle.$$

Allgemein gilt:

Satz 20.1.1 (*Gaußscher Integralsatz in der Ebene*)

Es sei $M \subset \mathbb{R}^2$ sowohl ein Normalbereich in x_1 -Richtung als auch ein Normalbereich in x_2 -Richtung. Der Rand ∂M sei positiv orientiert, stückweise glatt und regulär. Weiterhin sei $F: U \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld, definiert auf einer offenen Obermenge $U \supset M$. Schreibt man $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$ und ist γ wie oben, so folgt

$$\int_M \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = \int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle .$$

Bemerkungen.

- i)* Wie in Kapitel 18.1 nachgerechnet wurde, hängt das Kurvenintegral auf der rechten Seite nicht von der speziellen Wahl der Parametrisierung ab.
- ii)* Der Gaußsche Integralsatz in der Ebene gilt beispielsweise auch für [Greensche¹ Bereiche](#), das sind endliche Vereinigungen von Normalbereichen wie oben. Dabei ist auf die Orientierung der einzelnen Randkurven zu achten (vgl. Abbildung 20.2).

Beispiele.

i) Es sei $F(\underline{\mathbf{x}}) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$. Dann ist

$$\int_M \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dV = 2 \int_M dV = 2A(M) ,$$

¹G. Green, 1793-1841; Cambridge

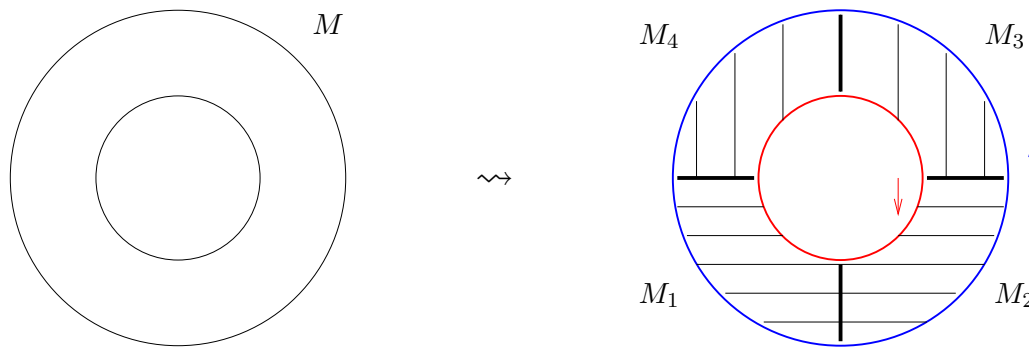


Abbildung 20.2: Ein Greenscher Bereich.

wobei $A(M)$ den Flächeninhalt von M bezeichne. Es gilt somit

$$A(M) = \frac{1}{2} \int_{\gamma} \langle F, d\underline{\mathbf{x}} \rangle, \quad F = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}.$$

ii) Ist speziell M die Ellipse

$$M = \left\{ \underline{\mathbf{x}} \in \mathbb{R}^2 : \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1 \right\}, \quad 0 \neq a, b \in \mathbb{R},$$

so ist

$$\gamma(t) := \begin{pmatrix} a \cos(t) \\ b \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi],$$

eine Parametrisierung von ∂M wie in Satz 20.1.1 gefordert, es folgt

$$\begin{aligned} A(M) &= \frac{1}{2} \int_0^{2\pi} \left\langle \begin{pmatrix} -b \sin(t) \\ a \cos(t) \end{pmatrix}, \begin{pmatrix} -a \sin(t) \\ b \cos(t) \end{pmatrix} \right\rangle dt \\ &= \frac{1}{2} \int_0^{2\pi} ab \, dt = \pi ab. \end{aligned}$$

Bemerkung. Es existiere eine reguläre C^1 -Parametrisierung $\gamma(t)$, $t \in [a, b]$, des positiv orientierten Randes ∂M . Dann gilt

$$\begin{pmatrix} \gamma_2'(t) \\ -\gamma_1'(t) \end{pmatrix} \perp \gamma'(t) = \begin{pmatrix} \gamma_1'(t) \\ \gamma_2'(t) \end{pmatrix}$$

und man bezeichnet

$$\underline{\mathbf{n}}(t) = \frac{1}{\|\gamma'(t)\|} \begin{pmatrix} \gamma'_2(t) \\ -\gamma'_1(t) \end{pmatrix}$$

als die **äußere Einheitsnormale an ∂M** (zum Parameterwert t).

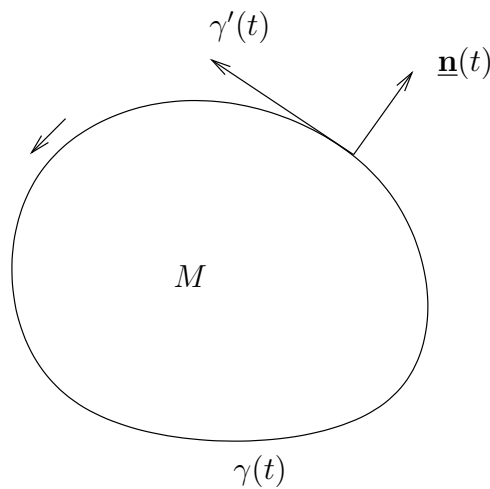


Abbildung 20.3: Die äußere Einheitsnormale an ∂M .

Ist $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}$ ein stetig differenzierbares Vektorfeld und setzt man

$\tilde{F} = \begin{pmatrix} -F_2 \\ F_1 \end{pmatrix}$, so folgt aus dem Gaußschen Satz für \tilde{F} :

$$\begin{aligned} \int_M \operatorname{div} F \, dV &= \int_M \left(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} \right) dV = \int_M \left(\frac{\partial \tilde{F}_2}{\partial x_1} - \frac{\partial \tilde{F}_1}{\partial x_2} \right) dV \\ &= \int_a^b \langle \tilde{F}(\gamma(t)), \gamma'(t) \rangle dt \\ &= \int_a^b \left(-F_2(\gamma(t))\gamma'_1(t) + F_1(\gamma(t))\gamma'_2(t) \right) dt \\ &= \int_a^b \langle F(\gamma(t)), \underline{\mathbf{n}}(t) \rangle \|\gamma'(t)\| dt . \end{aligned}$$

Bemerkungen.

- i)* Das Integral auf der rechten Seite heißt auch der **Fluss des Vektorfeldes F durch ∂M** . Man beachte, dass hier nur die Komponente von F senkrecht zu ∂M zu berücksichtigen ist.
- ii)* In dieser Form lässt sich Satz 20.1.1 auf den \mathbb{R}^3 verallgemeinern. Dazu ist zunächst ein kurzer Exkurs notwendig.

20.2 Flächen im \mathbb{R}^3 (Gebiet; zusammenhängend; Parameterdarstellung einer Fläche; Kurven auf Flächen; Tangentialebene; Normaleneinheitsvektor; Oberflächeninhalt; Oberflächenintegral)

Aus Kapitel 19.1 ist bekannt, was der Inhalt einer quadrierbaren Menge im \mathbb{R}^2 ist und wie über eine solche Menge zu integrieren ist.

Nun werden zweidimensionale (im Allgemeinen nicht-ebene) Flächen im \mathbb{R}^3 betrachtet und die Begriffe Inhalt der Fläche und Integral über diese Fläche eingeführt.

In diesem Abschnitt ist $U \subset \mathbb{R}^2$ stets ein **Gebiet**, das heißt U ist offen und U ist **zusammenhängend**. Zusammenhängend wiederum bedeutet, dass je zwei Punkte aus U durch eine stückweise glatte Kurve verbunden werden können, die **vollständig in U** verläuft. Die typische Vorstellung ist in Abbildung 20.4 angedeutet.

Definition 20.2.1

Es sei $U \subset \mathbb{R}^2$ ein quadrierbares Gebiet und $X: \bar{U} \rightarrow \mathbb{R}^3$,

$$\bar{U} \ni \underline{\mathbf{u}} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \mapsto X(\underline{\mathbf{u}}) = \begin{pmatrix} X_1(\underline{\mathbf{u}}) \\ X_2(\underline{\mathbf{u}}) \\ X_3(\underline{\mathbf{u}}) \end{pmatrix},$$

eine *injektive* Abbildung der Klasse C^1 . Ferner habe die Jacobi-Matrix

$$DX = \begin{pmatrix} \frac{\partial X_1}{\partial u_1} & \frac{\partial X_1}{\partial u_2} \\ \frac{\partial X_2}{\partial u_1} & \frac{\partial X_2}{\partial u_2} \\ \frac{\partial X_3}{\partial u_1} & \frac{\partial X_3}{\partial u_2} \end{pmatrix}$$

überall den *maximalen Rang 2*. Dann heißt die Punktmenge

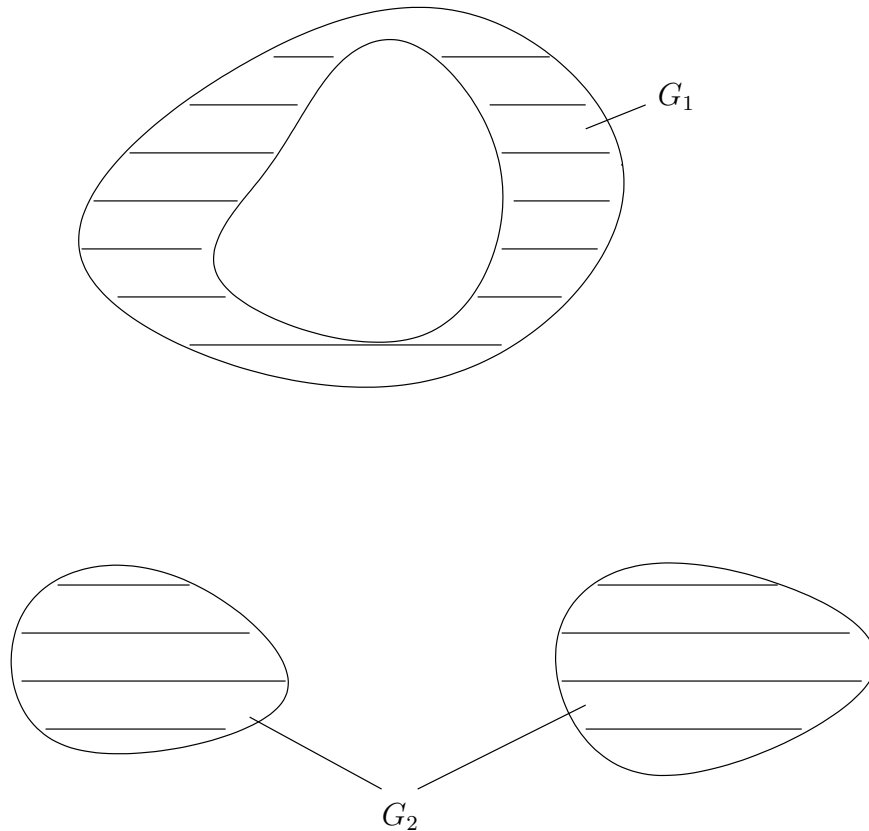
$$\mathbb{R}^3 \ni S = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = X(\underline{\mathbf{u}}), \underline{\mathbf{u}} \in U\}$$

eine (*reguläre, eingebettete*) *Fläche im \mathbb{R}^3* . Der *Rand von S* ist

$$\partial S = \{\underline{\mathbf{x}} \in \mathbb{R}^3 : \underline{\mathbf{x}} = X(\underline{\mathbf{u}}), \underline{\mathbf{u}} \in \partial U\}.$$

Bemerkungen.

- i) Analog zur Definition einer Kurve wird meist die Abbildung X als Fläche bezeichnet, S ist die Spur der Abbildung (X heißt auch **Parameterdarstellung von S**). Analog definiert man den Rand als Kurve $X|_{\partial U}: \underline{\mathbf{u}} \in \partial U \mapsto X(\underline{\mathbf{u}})$. Hier wird dieser Unterschied allerdings ignoriert, d.h. X und S werden identifiziert. Zu beachten ist


 Abbildung 20.4: G_1 ist zusammenhängend, G_2 ist es nicht.

aber: Es gibt unterschiedliche Parameterdarstellungen einer Fläche S (vgl. auch Parametertransformationen für Kurven)

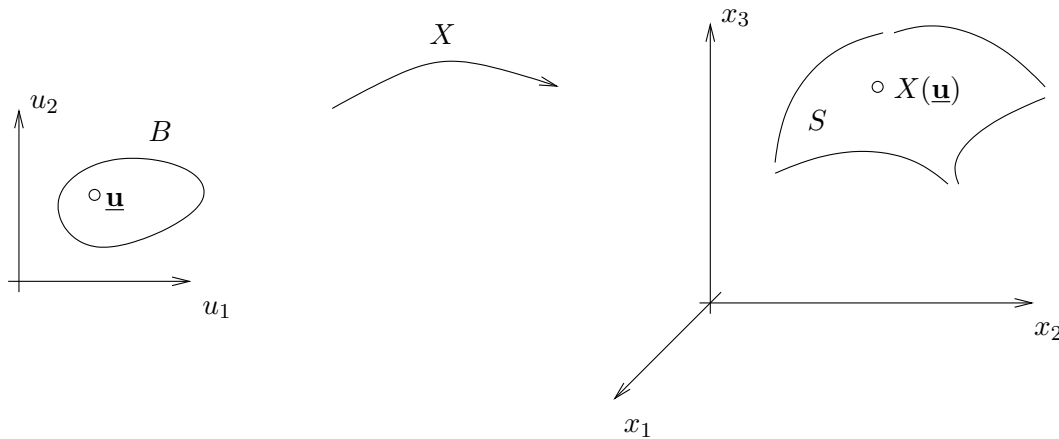
- ii)* Differenzierbarkeit wurde nur für offene Mengen definiert. Eine Abbildung der Klasse C^1 auf der abgeschlossenen Menge \bar{U} sei definiert und stetig differenzierbar auf einer offenen Obermenge $O \supset U$.

Beispiele.

- i)* Eine Kugel in \mathbb{R}^3 kann wie folgt parametrisiert sein: Zu fixiertem $r > 0$ sei

$$X(\underline{\mathbf{u}}) = \begin{pmatrix} r \cos(u_1) \cos(u_2) \\ r \sin(u_1) \cos(u_2) \\ r \sin(u_2) \end{pmatrix},$$

$$U = \left\{ \underline{\mathbf{u}} \in \mathbb{R}^2, 0 < u_1 < 2\pi, -\frac{\pi}{2} < u_2 < \frac{\pi}{2} \right\}.$$

Abbildung 20.5: Eine Fläche im \mathbb{R}^3 .

Es ist

$$DX = \begin{pmatrix} -r \sin(u_1) \cos(u_2) & -r \cos(u_1) \sin(u_2) \\ r \cos(u_1) \cos(u_2) & -r \sin(u_1) \sin(u_2) \\ 0 & r \cos(u_2) \end{pmatrix},$$

die Matrix hat vollen Rang 2.

ii) Der Spezialfall von **Graphenflächen** wurde in Kapitel 17.2 bereits ausführlich studiert.

Es sei nun eine Fläche S wie in Definition 20.2.1 gegeben. Weiterhin sei $\gamma: [a, b] \rightarrow U$ eine reguläre, glatte Kurve in U . Dann ist

$$c: [a, b] \rightarrow \mathbb{R}^3, \quad c := X \circ \gamma,$$

eine reguläre, glatte **Kurve auf der Fläche S** .

Es seien weiter $\underline{\mathbf{u}}^{(0)} \in U$ und $t_0 \in (a, b)$ betrachtet mit

$$\gamma(t_0) = \underline{\mathbf{u}}^{(0)}, \quad X(\gamma(t_0)) = X(\underline{\mathbf{u}}^{(0)}) = c(t_0).$$

Nach der Kettenregel gilt

$$c'(t_0) = X_{u_1}(\underline{\mathbf{u}}^{(0)})\gamma'_1(t_0) + X_{u_2}(\underline{\mathbf{u}}^{(0)})\gamma'_2(t_0),$$

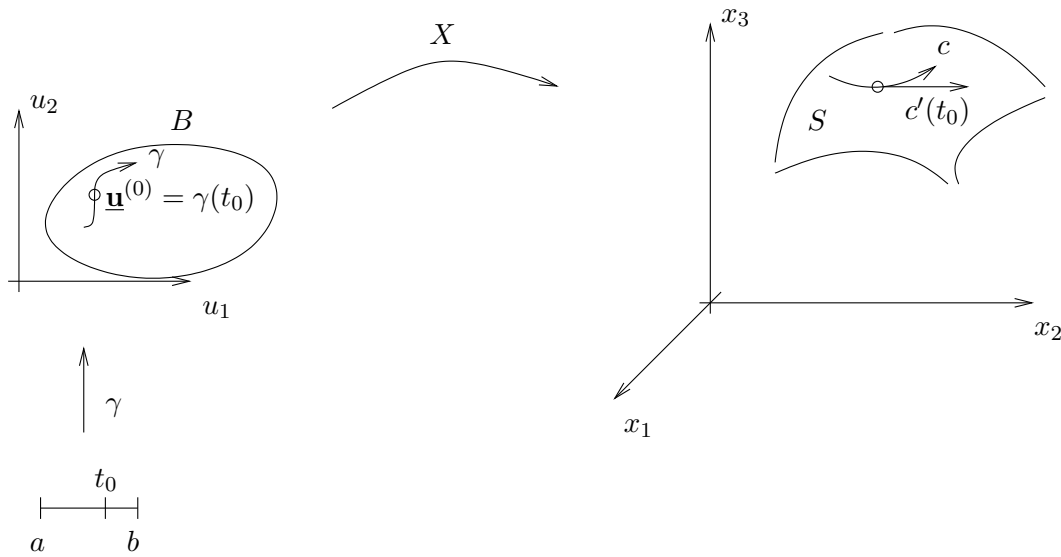


Abbildung 20.6: Eine Kurve auf S .

wobei die Notation benutzt wird

$$X_{u_i}(\mathbf{u}) := \begin{pmatrix} \frac{\partial X_1}{\partial u_i} \\ \frac{\partial X_2}{\partial u_i} \\ \frac{\partial X_3}{\partial u_i} \end{pmatrix}(\mathbf{u}), \quad i = 1, 2.$$

Der Vektor $c'(t_0)$ gibt die Richtung der Tangente an die Kurve c auf der Fläche S im Punkt $c(t_0)$. Alle solchen **Tangentenvektoren** bestimmen die **Tangentialebene** $T_{\mathbf{u}^{(0)}}X$. Da X_{u_1} und X_{u_2} nach Voraussetzung linear unabhängig sind, ist $T_{\mathbf{u}^{(0)}}X$ tatsächlich eine Ebene, gegeben durch

$$T_{\mathbf{u}^{(0)}}X = \{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \lambda X_{u_1}(\mathbf{u}^{(0)}) + \mu X_{u_2}(\mathbf{u}^{(0)}), \lambda, \mu \in \mathbb{R}\}.$$

Senkrecht zur Tangentialebene hat man

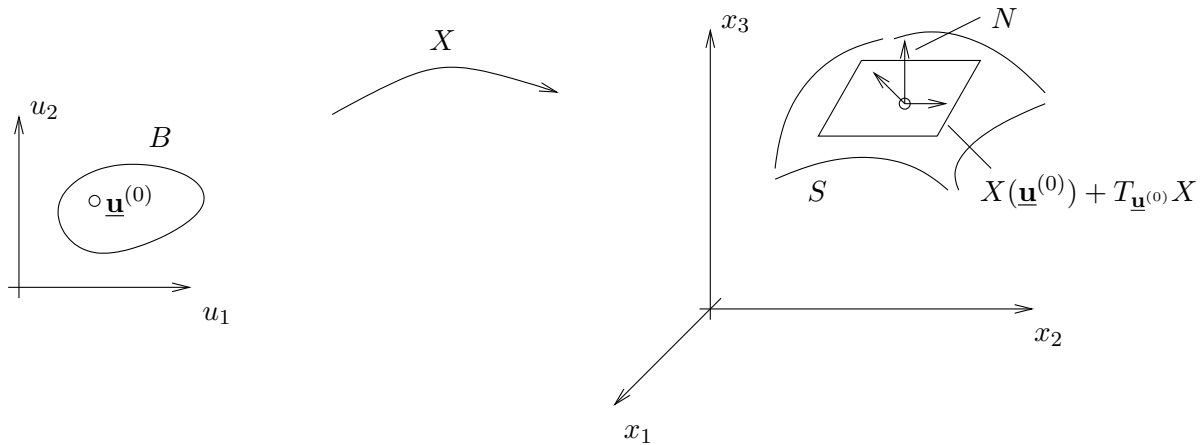


Abbildung 20.7: Normale und Tangentialebene.

Definition 20.2.2

Es seien S bzw. X wie in Definition 20.2.1 gegeben. Dann heißt

$$N(\underline{\mathbf{u}}) = \frac{X_{u_1}(\underline{\mathbf{u}}) \times X_{u_2}(\underline{\mathbf{u}})}{\|X_{u_1}(\underline{\mathbf{u}}) \times X_{u_2}(\underline{\mathbf{u}})\|}, \quad \underline{\mathbf{u}} \in U,$$

Normaleneinheitsvektor der Fläche in $X(\underline{\mathbf{u}})$ bzw. zum Parameterwert $\underline{\mathbf{u}}$.

Bemerkung. Streng genommen muss unterschieden werden, ob N (und analog die Tangentialebene) als Funktion von $\underline{\mathbf{u}}$ angesehen wird ($N: U \rightarrow B_1(0)$) oder ob man sich die Normale N als Vektor auf der Fläche vorstellt. Ohne auf formale Details einzugehen (und mit nicht ganz korrekter Notation) wird hier N als Symbol für beide Situationen benutzt (formal ist bei den folgenden Oberflächenintegralen N zu ersetzen durch $\tilde{N}(\underline{\mathbf{x}}) = N \circ X^{-1}(\underline{\mathbf{x}})$, insbesondere gilt dann $\tilde{N}(X(\underline{\mathbf{u}})) = N(\underline{\mathbf{u}})$).

Beispiel. Im obigen Beispiel

$$X(\underline{\mathbf{u}}) = \begin{pmatrix} r \cos(u_1) \cos(u_2) \\ r \sin(u_1) \cos(u_2) \\ r \sin(u_2) \end{pmatrix}$$

ist

$$X_{u_1} \times X_{u_2} = \begin{pmatrix} r^2 \cos(u_1) \cos^2(u_2) \\ r^2 \sin(u_1) \cos^2(u_2) \\ r^2 \sin(u_2) \cos(u_2) \end{pmatrix}, \quad \|X_{u_1} \times X_{u_2}\| = r^2 \cos(u_2),$$

also ist $N(\underline{\mathbf{u}})$ der (radiale Vektor)

$$N(\underline{\mathbf{u}}) = \begin{pmatrix} \cos(u_1) \cos(u_2) \\ \sin(u_1) \cos(u_2) \\ \sin(u_2) \end{pmatrix}.$$

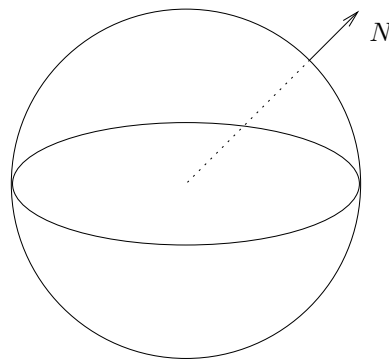


Abbildung 20.8: Einheitsnormale an die Sphäre.

Wie wird der **Oberflächeninhalt** von S gemessen bzw. **über S integriert**?

Definition 20.2.3

Es seien X, S wie oben. Dann heißt

$$A(S) := \int_U \|X_{u_1}(\underline{\mathbf{u}}) \times X_{u_2}(\underline{\mathbf{u}})\| \, dV$$

*der **Oberflächeninhalt** von S .*

Bemerkung. Hier ist die Unabhängigkeit der Definition von der speziellen Parametrisierung zu überprüfen (Transformationssatz).

Beispiel. Die Oberfläche der Kugel vom Radius r errechnet sich zu

$$\begin{aligned} A(S) &= \int_U r^2 \cos(u_2) \, dV \\ &= r^2 \int_0^{2\pi} \left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(u_2) \, du_2 \right] du_1 \\ &= r^2 \int_0^{2\pi} 2 \, du_1 = 4\pi r^2. \end{aligned}$$

Definition 20.2.4

Es seien X, S wie oben. Weiterhin sei $f: \mathbb{R}^3 \supset S \rightarrow \mathbb{R}$ stetig auf S . Dann heißt

$$\int_U f(X(\underline{\mathbf{u}})) \|X_{u_1}(\underline{\mathbf{u}}) \times X_{u_2}(\underline{\mathbf{u}})\| \, dV =: \int_S f \, dA$$

das *Oberflächenintegral* von f über S .

Beispiel. Man betrachte wieder die Kugel und $f(\underline{\mathbf{x}}) = x_1^2$. Dann ist

$$\begin{aligned} \int_S f \, dA &= \int_U (r^2 \cos^2(u_1) \cos^2(u_2)) r^2 \cos(u_2) \, dV \\ &= r^4 \int_0^{2\pi} \left[\underbrace{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3(u_2) \, du_2}_{=[\sin(u_2) - \frac{1}{3} \sin^3(u_2)]_{-\pi/2}^{\pi/2} = 4/3} \right] \cos^2(u_1) \, du_1 \\ &= \frac{4}{3} r^4 \left[\frac{1}{2} (\sin(u_1) \cos(u_1) + u_1) \right]_0^{2\pi} = \frac{4}{3} \pi r^4. \end{aligned}$$

Bemerkung. Oft ist es notwendig, kompliziertere “Flächen” zu betrachten, die nicht das Bild einer **einzig** Parametrisierung sind. Beispielsweise kann man sich S als Vereinigung von endlich vielen

Flächenstücken $S^{(1)}, \dots, S^{(k)}$ vorstellen, die sich gegenseitig nicht “überlappen”. Dann sind der Oberflächeninhalt und das Oberflächenintegral stückweise definiert.

20.3 Der Gaußsche Integralsatz im \mathbb{R}^3 (äußere Einheitsnormale; Massenfluss; Kontinuitätsgleichung; partielle Integration)

Im Folgenden ist M ein Normalbereich im \mathbb{R}^3 (projizierbar in x_1, x_2 und x_3 -Richtung) mit **äußerer Einheitsnormalen N an ∂M** (“zeigt aus M heraus”). Der Rand von M sei eine **Vereinigung von endlich vielen Flächenstücken nach Definition 20.2.1** (siehe Beispiel Würfel). (Endliche Vereinigungen solcher Normalbereiche können wieder “stückweise” untersucht werden.)

Beispiele.

- i) Es sei an Abbildung 20.8 erinnert.
- ii) Man betrachte den Würfel

$$M := \{\mathbf{x} \in \mathbb{R}^3 : 0 \leq x_i \leq 1, i = 1, 2, 3\}.$$

Die äußere Normale an ∂M ist in Abbildung 20.9 angedeutet.

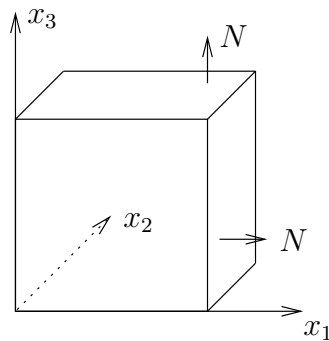


Abbildung 20.9: Äußere Normale an den Würfel.

Ähnlich wie in Kapitel 20.1 folgt

Satz 20.3.1 (*Gaußscher Integralsatz im \mathbb{R}^3*)

Es seien M , ∂M und N wie oben und $F: G \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$, G offen, $M \subset G$, ein stetig differenzierbares Vektorfeld. Dann gilt

$$\int_M \operatorname{div} F \, dV = \int_{\partial M} \langle F, N \rangle \, dA .$$

Beispiele.

- i)* Es sei $M = B_r(\mathbf{0})$ die Kugel vom Radius $r > 0$ um den Nullpunkt. Mit der obigen Parametrisierung $X(\mathbf{u})$ von $\partial B_r(\mathbf{0})$ ist bereits nachgerechnet

$$N(\mathbf{u}) = \begin{pmatrix} \cos(u_1) \cos(u_2) \\ \sin(u_1) \cos(u_2) \\ \sin(u_2) \end{pmatrix}, \quad \|X_{u_1} \times X_{u_2}\| = r^2 \cos(u_2) .$$

Man betrachte nun das Vektorfeld $F(\mathbf{x}) = \mathbf{x}$, d.h. $\operatorname{div} F \equiv 3$. Dann gilt

$$F(X(\mathbf{u})) = \begin{pmatrix} r \cos(u_1) \cos(u_2) \\ r \sin(u_1) \cos(u_2) \\ r \sin(u_2) \end{pmatrix}, \quad \langle F(X(\mathbf{u})), N(\mathbf{u}) \rangle = r .$$

Also folgt

$$\begin{aligned} \int_{\partial B_r(\mathbf{0})} F \cdot N \, dA &= \int_0^{2\pi} \left[\int_{-\pi/2}^{\pi/2} r r^2 \cos(u_2) \, du_2 \right] du_1 \\ &= 4\pi r^3 . \end{aligned}$$

Andererseits ist

$$\int_{B_r(\mathbf{0})} \operatorname{div} F \, dV = 3V ,$$

wobei V das Volumen der Kugel bezeichne. Nach dem Gaußschen Integralsatz im \mathbb{R}^3 gilt somit

$$V = \frac{4}{3}\pi r^3 .$$

ii) **Massenerhaltung in der Fluid-Mechanik.** Man betrachte eine strömende Flüssigkeit in drei Raumdimensionen. Notation:

$\underline{\mathbf{x}} \in \mathbb{R}^3$: Räumliche Variable;

$t \in [0, \infty)$: Zeit;

$v = v(\underline{\mathbf{x}}, t) = \begin{pmatrix} v_1(\underline{\mathbf{x}}, t) \\ v_2(\underline{\mathbf{x}}, t) \\ v_3(\underline{\mathbf{x}}, t) \end{pmatrix}$: Geschwindigkeitsfeld der Flüssigkeit;

$\rho = \rho(\underline{\mathbf{x}}, t)$: Massendichte.

Man betrachte nun ein beliebiges “Testvolumen” $\Omega \subset \mathbb{R}^3$. Dann gilt

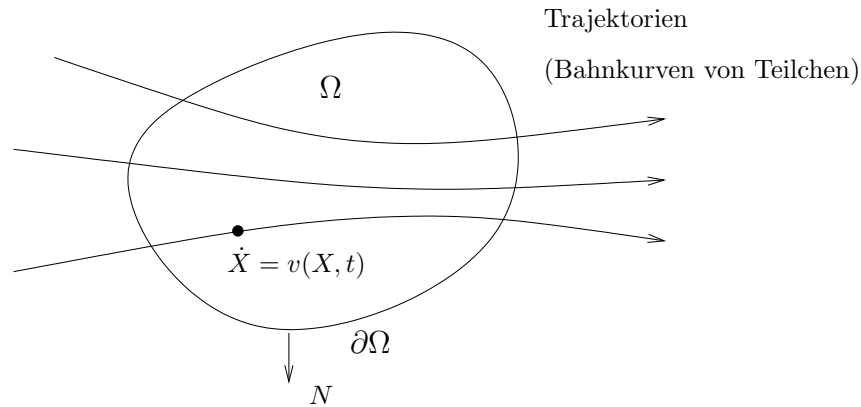
Massenänderung in Ω = **Massenfluss** durch $\partial\Omega$, d.h.

$$\begin{aligned} \int_{\Omega} \frac{\partial \rho(\underline{\mathbf{x}}, t)}{\partial t} dV &= - \underbrace{\int_{\partial\Omega} \langle \rho(\underline{\mathbf{x}}, t)v(\underline{\mathbf{x}}, t), N \rangle dA}_{\text{Fluß des Vektorfeldes } \rho v \text{ durch } \partial\Omega} \\ &= - \int_{\Omega} \operatorname{div} (\rho(\underline{\mathbf{x}}, t)v(\underline{\mathbf{x}}, t)) dx . \end{aligned}$$

(Hier ist die Divergenz bzgl. der räumlichen Variablen zu interpretieren.)

Da Ω beliebig gewählt werden kann folgt (“relativ leicht”) als **Kontinuitätsgleichung**:

$$\frac{\partial \rho(\underline{\mathbf{x}}, t)}{\partial t} + \operatorname{div} (\rho(\underline{\mathbf{x}}, t)v(\underline{\mathbf{x}}, t)) = 0 .$$

Abbildung 20.10: Trajektorie $t \mapsto X(t) \in \mathbb{R}^3$.

Ist die Flüssigkeit **inkompressibel** mit $\rho(\underline{\mathbf{x}}, t) = \textit{konstant}$, so reduziert sich die Kontinuitätsgleichung auf

$$\operatorname{div} v(\underline{\mathbf{x}}, t) = 0 .$$

iii) **Partielle Integration.** Die Voraussetzungen aus Satz 20.3.1 seien erfüllt und es seien $f, g: M \rightarrow \mathbb{R}$ stetig differenzierbar auf \overline{M} . Als Vektorfeld F wird zunächst gewählt:

$$F(\underline{\mathbf{x}}) = \begin{pmatrix} f(\underline{\mathbf{x}})g(\underline{\mathbf{x}}) \\ 0 \\ 0 \end{pmatrix} ,$$

es gilt also

$$\begin{aligned} \operatorname{div} F(\underline{\mathbf{x}}) &= D_1(f(\underline{\mathbf{x}})g(\underline{\mathbf{x}})) + D_2 0 + D_3 0 \\ &= (D_1 f(\underline{\mathbf{x}}))g(\underline{\mathbf{x}}) + f(\underline{\mathbf{x}})D_1 g(\underline{\mathbf{x}}) . \end{aligned}$$

Ist

$$N = \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}$$

die äußere Einheitsnormale an ∂M , so folgt

$$\int_M ((D_1 f(\underline{x}))g(\underline{x}) + f(\underline{x})D_1 g(\underline{x})) dV = \int_{\partial M} fgN_1 dA .$$

Für die anderen Komponenten wird analog argumentiert, es gilt:

$$\int_M f(D_i g) dV = - \int_M (D_i f)g dV + \int_{\partial M} fgN_i dA , \quad i = 1, 2, 3 .$$

Interpretation des Gaußschen Integralsatzes. Ist $\operatorname{div} F = 0$, so verschwindet der Fluß $\int_{\partial M} \langle F, N \rangle dA$ des Vektorfeldes F durch die Randfläche ∂M . Man bezeichnet F dann als **quellenfrei**.

Kapitel 21

Der Integralsatz von Stokes

21.1 Der Satz (orientierte Fläche; Zirkulation)

Der Integralsatz von Stokes stellt im \mathbb{R}^3 eine Verbindung her zwischen dem Fluss eines Vektorfeldes $\operatorname{rot}F$ durch eine **orientierte Fläche** und der **Zirkulation** längs des Randes.

Die Situation wird in diesem Kapitel nur skizzenhaft umrissen.

Satz 21.1.1 (*Integralsatz von Stokes*)

Es bezeichne S (mit Parametrisierung $X: \bar{U} \rightarrow \mathbb{R}^3$) eine Fläche im \mathbb{R}^3 gemäß Definition 20.2.1. Hierbei sei \bar{U} sowohl x_1 als auch x_2 -Normalbereich mit glattem, positiv orientierten Rand ∂U , der durch eine Kurve γ gegeben sei. Die Kurve $c = X \circ \gamma$ ist also eine Parameterdarstellung von ∂S .

Es sei weiter $S \subset V$, $V \subset \mathbb{R}^3$ offen, $F: V \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld und schließlich bezeichne N die (positiv orientierte) Normale auf S . Dann gilt

$$\int_S \langle \operatorname{rot}F, N \rangle \, dA = \int_c \langle F, d\underline{x} \rangle .$$

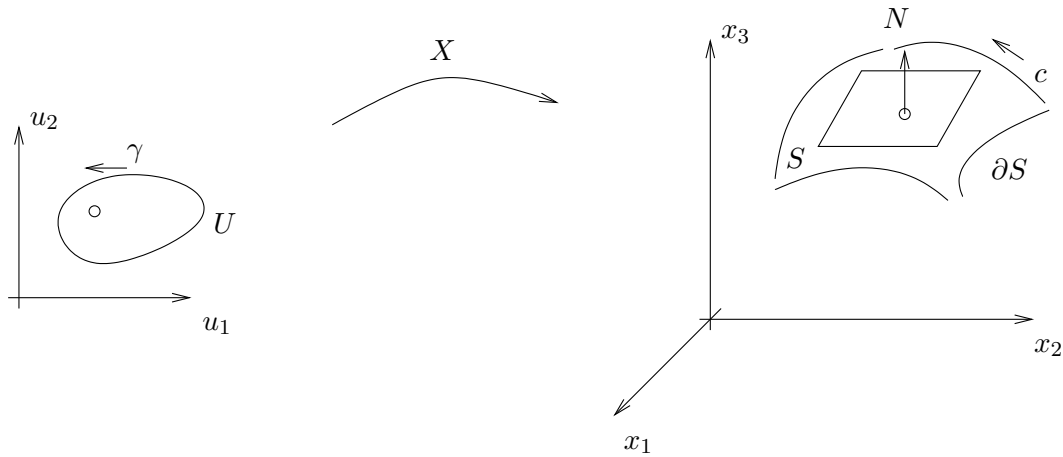


Abbildung 21.1: Zum Integralsatz von Stokes.

Bemerkung. Ist ∂S wie oben durch $c = X \circ \gamma$ parametrisiert, γ positiv orientiert, so zeigt die positiv orientierte Normale auf S in Richtung $X_{u_1} \times X_{u_2}$.

Beispiel. Es sei F das Vektorfeld

$$F(\underline{\mathbf{x}}) = \begin{pmatrix} -x_2 \\ x_1 \\ 1 \end{pmatrix}$$

und S die obere Halbsphäre vom Radius $r > 0$ um den Ursprung. Als Parametrisierung von S sei betrachtet:

$$X : U = \{\underline{\mathbf{u}} \in \mathbb{R}^2 : \|\underline{\mathbf{u}}\|^2 < r^2\}, \quad \underline{\mathbf{u}} \mapsto \begin{pmatrix} u_1 \\ u_2 \\ \sqrt{r^2 - u_1^2 - u_2^2} \end{pmatrix}.$$

Ist

$$\gamma(t) = \begin{pmatrix} r \cos(t) \\ r \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi],$$

als Parametrisierung von ∂U gewählt, so ergibt sich

$$c(t) = X \circ \gamma(t) = \begin{pmatrix} r \cos(t) \\ r \sin(t) \\ 0 \end{pmatrix}, \quad t \in [0, 2\pi]$$

als zu wählende Parametrisierung von ∂S .

i) Somit gilt

$$\int_c \langle F, d\mathbf{x} \rangle = \int_0^{2\pi} \left\langle \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 1 \end{pmatrix}, \begin{pmatrix} -r \sin(t) \\ r \cos(t) \\ 0 \end{pmatrix} \right\rangle dt = 2\pi r^2.$$

ii) Es gilt

$$\operatorname{rot} F(\mathbf{x}) \equiv \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}.$$

Als positiv orientierte Normale an S ist zu wählen

$$\begin{aligned} N(\mathbf{u}) &= \frac{1}{\|X_{u_1} \times X_{u_2}\|} \begin{pmatrix} 1 \\ 0 \\ -\frac{u_1}{\sqrt{r^2 - u_1^2 - u_2^2}} \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ -\frac{u_2}{\sqrt{r^2 - u_1^2 - u_2^2}} \end{pmatrix} \\ &= \frac{1}{\|X_{u_1} \times X_{u_2}\|} \begin{pmatrix} \frac{u_1}{\sqrt{r^2 - u_1^2 - u_2^2}} \\ \frac{u_2}{\sqrt{r^2 - u_1^2 - u_2^2}} \\ 1 \end{pmatrix}. \end{aligned}$$

Es folgt

$$\langle \operatorname{rot} F(X(\mathbf{u})), N(\mathbf{u}) \rangle = \frac{2}{\|X_{u_1} \times X_{u_2}\|},$$

und mit der Definition des Oberflächenintegrals ergibt sich schließlich

$$\begin{aligned}\int_S \langle \operatorname{rot} F, N \rangle \, dA &= \iint_U \frac{2}{\|X_{u_1} \times X_{u_2}\|} \|X_{u_1} \times X_{u_2}\| \, du_1 \, du_2 \\ &= 2\pi r^2 .\end{aligned}$$

Damit ist der Integralsatz von Stokes in diesem Beispiel verifiziert.

Interpretation des Integralsatzes von Stokes. Der Fluß eines Vektorfeldes $\operatorname{rot} F$ durch die orientierte Fläche S ist gleich der Zirkulation $\int_c \langle F, dx \rangle$ längs der Randkurve c .

Verswindet $\operatorname{rot} F$ identisch, so ist die Zirkulation längs jeder geschlossenen Kurve, die eine Fläche berandet, Null. Man nennt F wirbelfrei. (Vgl. auch die Diskussion konservativer Vektorfelder in Kapitel 18.)

Teil IX
Funktionentheorie

Kapitel 22

Einführung in die Funktionentheorie

In Kapitel 17 wurde die Differentialrechnung von Funktionen $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ mehrerer Veränderlicher besprochen. Der Ableitungsbegriff war dabei nicht als Verallgemeinerung der eindimensionalen Diskussion evident: Für Funktionen einer Variablen kann die Ableitung als Grenzwert des Differenzenquotienten definiert werden. Im höherdimensionalen Fall ist das nicht möglich, da nicht “durch einen Vektor geteilt werden kann”. Deshalb waren partielle Ableitungen, Richtungsableitungen und die totale Ableitung zu unterscheiden.

Die Situation ändert sich auch nicht, wenn Abbildungen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ zu untersuchen sind.

Die Situation ändert sich allerdings **dramatisch**, wenn der \mathbb{R}^2 als Gaußsche Zahlenebene mit der komplexen Multiplikation versehen wird, d.h. beim Studium von Abbildungen $f: \mathbb{C} \rightarrow \mathbb{C}$. Bzgl. der komplexen Multiplikation existiert ein Inverses, “durch komplexe Zahlen kann geteilt werden”. Damit ist es möglich analog zum eindimensionalen Fall eine Ableitung als Grenzwert von Differenzenquotienten zu definieren.

In der Funktionentheorie geht es um die (auf den ersten Blick) wirklich erstaunlichen Konsequenzen dieser Tatsache, die kein Analogon in der reellen Analysis haben.

Zum Verständnis des Folgenden werden die Betrachtungen aus Kapitel 7 vorausgesetzt (einige Stichworte sind: Der Körper der kom-

plexen Zahlen, komplexe Potenzreihen, Exponentialfunktion, Eulersche Formeln, die Gaußsche Zahlenebene).

Topologische Begriffe wie offene und abgeschlossene Mengen wurden dort ebenfalls aus dem \mathbb{R}^2 abgeleitet.

Besonders betont sei nochmals: Der Konvergenzbegriff spielt hier wie auch in der reellen Analysis die zentrale Rolle.

22.1 Holomorphe Funktionen (komplexe Differenzierbarkeit; höhere Ableitungen; Rechenregeln)

In diesem Paragraphen wird der zentrale Begriff in der Funktionentheorie, die **komplexe Differenzierbarkeit** eingeführt.

Die Notation ist dabei wie üblich: $z = x + iy$, $x, y \in \mathbb{R}$, $f(z): \mathbb{C} \supset U \rightarrow f(U) \subset \mathbb{C}$,

$$f(z) = u(x, y) + iv(x, y),$$

mit reellwertigen Funktionen $u(x, y)$, $v(x, y)$, U ist stets offen.

Die skizzenhafte Veranschaulichung ist in Abbildung 22.1 wiedergegeben.

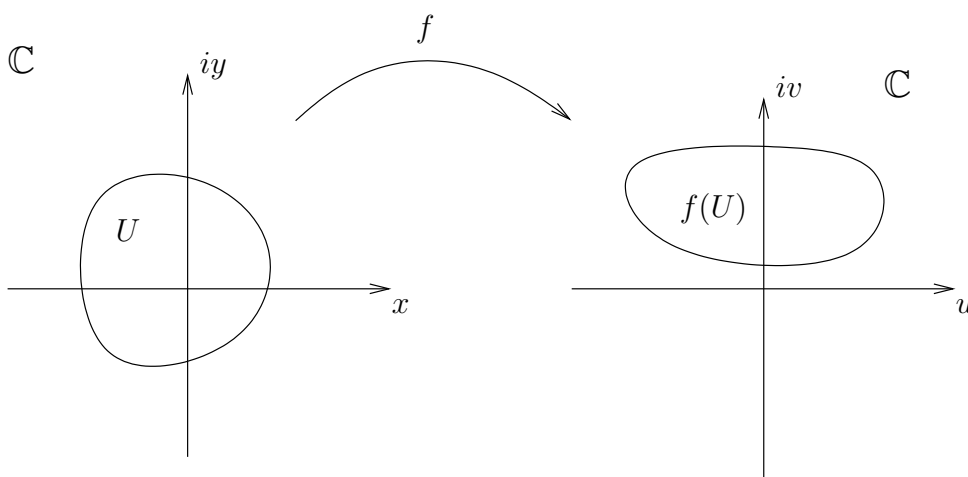


Abbildung 22.1: Eine Funktion $f: \mathbb{C} \rightarrow \mathbb{C}$.

Definition 22.1.1

i) Es sei $U \subset \mathbb{C}$ offen, $f: U \rightarrow \mathbb{C}$. Die Funktion f heißt im Punkt $z_0 \in \mathbb{C}$ (**komplex**) differenzierbar, falls der Grenzwert (die (**komplexe**) Ableitung $f'(z_0)$)

$$\lim_{z \rightarrow z_0, z \neq z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert (in \mathbb{C}).

ii) Die Funktion f heißt (**komplex**) differenzierbar auf U (oder: **holomorph auf U** , oder: **regulär auf U**), falls f in jedem Punkt $z_0 \in U$ differenzierbar ist.

Notation:

$$\lim_{z \rightarrow z_0, z \neq z_0} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0) = \frac{df}{dz}(z_0) = \frac{df}{dz}|_{z=z_0}.$$

Wie üblich wird hier die komplexe Ableitung aufgefasst als Funktion $f': U \rightarrow \mathbb{C}$.

Bemerkungen.

i) Die Definition der Ableitung erfolgt analog zum Fall einer Funktion einer reellen Variablen als Grenzwert von Differenzenquotienten. Fasst man f lediglich als Funktion $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ auf (ohne komplexe Multiplikation), so ergibt eine solche Definition keinen Sinn.

ii) Rekursiv werden **höhere Ableitungen** definiert:

$$f'' := (f')', \quad f''' := (f'')' \quad \dots, \quad f^{(n)} := (f^{(n-1)})'.$$

iii) Eine in z_0 komplex differenzierbare Funktion ist dort stetig.

Beispiele.

i) Konstante Funktionen sind holomorph auf \mathbb{C} , ist nämlich

$$f(z) = c = a + ib, \quad c \in \mathbb{C}, \quad a, b \in \mathbb{R},$$

so folgt für $z \neq z_0$

$$\frac{f(z) - f(z_0)}{z - z_0} = 0, \quad \text{also} \quad \lim_{z \rightarrow z_0, z \neq z_0} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0) = 0.$$

ii) Es sei $f(z) = z$ und ein beliebiges $z_0 \in \mathbb{C}$ sei fixiert. Es folgt

$$\lim_{z \rightarrow z_0, z \neq z_0} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{z \rightarrow z_0, z \neq z_0} \frac{z - z_0}{z - z_0} = 1,$$

also $f'(z_0) = 1$ für alle $z_0 \in \mathbb{C}$, d.h. $f'(z) \equiv 1$.

iii) Es sei jetzt $f(z) = \bar{z}$ und $z_0 = x_0 + iy_0 \in \mathbb{C}$ fixiert. Ist speziell $z = x + iy_0$, so folgt

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{(x - iy_0) - (x_0 - iy_0)}{(x + iy_0) - (x_0 + iy_0)} = 1,$$

ist andererseits $z = x_0 + iy$, so folgt

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{(x_0 - iy) - (x_0 - iy_0)}{(x_0 + iy) - (x_0 + iy_0)} = -1.$$

Die Funktion $f(z) = \bar{z}$ ist nicht holomorph!

Nach dem letzten Beispiel ist die Klasse der holomorphen Funktionen nicht so groß wie man es zunächst vielleicht erwartet hätte. Um einzusehen (ohne die Definition heranzuziehen), dass zumindest Polynome etc. holomorph sind, werden wie üblich Rechenregeln benötigt:

Satz 22.1.1

Es seien $U, V \subset \mathbb{C}$ offen.

i) Summe und Produkt zweier (in $z_0 \in U$) komplex differenzierbarer Funktionen f, g sind komplex differenzierbar. Es gilt

$$(f + g)'(z_0) = f'(z_0) + g'(z_0),$$

$$(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0).$$

ii) Es sei $f: U \rightarrow \mathbb{C}$ in z_0 komplex differenzierbar und $f(z_0) \neq 0$. Dann ist $\frac{1}{f}$ in einer Umgebung von z_0 wohl definiert, in z_0 komplex differenzierbar und es gilt

$$\left(\frac{1}{f}\right)'(z_0) = -\frac{f'(z_0)}{f^2(z_0)}.$$

iii) Es seien $f: U \rightarrow V$ und $g: V \rightarrow \mathbb{C}$ in $z_0 \in U$ bzw. in $w_0 = f(z_0)$ komplex differenzierbar. Dann ist die Verkettung $g \circ f$ in z_0 komplex differenzierbar mit

$$(g \circ f)'(z_0) = g'(w_0)f'(z_0).$$

Beispiele.

i) Es sei $f(z) = z^2$. Dann gilt

$$f'(z) = 1z + z1 = 2z,$$

allgemein folgt $n = 1, 2, 3, \dots$

$$\frac{d}{dz}z^n = nz^{n-1}.$$

ii) Es sei $f(z) = \frac{1}{z}$, $z \neq 0$. Dann gilt

$$f'(z) = -\frac{1}{z^2},$$

allgemein folgt für $k = \pm 1, \pm 2, \pm 3, \dots$ ($z \neq 0$ im Fall $k < 0$)

$$\frac{d}{dz}z^k = kz^{k-1}.$$

Ruft man sich die Notation

$$f(z) = u(x, y) + iv(x, y), \quad z = x + iy,$$

mit zwei Funktionen $u, v: \mathbb{R}^2 \rightarrow \mathbb{R}$ ins Gedächtnis, so stellt sich an dieser Stelle die natürliche Frage:

Wie hängt die **reelle Differenzierbarkeit von u und v** mit der **komplexen Differenzierbarkeit von f** zusammen?

Zur Erinnerung: Die Funktion $f(z) = \bar{z} = x + i(-y)$ ist **nicht** komplex differenzierbar, obwohl in diesem Fall u, v beliebig glatt sind.

22.2 Die Cauchy- Riemanschen Differentialgleichungen (komplexe Differenzierbarkeit versus reelle Differenzierbarkeit)

Zur Beantwortung obiger Frage wird die zusätzliche Notation

$$\begin{aligned} \frac{\partial u}{\partial x} &= u_x, & \frac{\partial u}{\partial y} &= u_y, \\ \frac{\partial v}{\partial x} &= v_x, & \frac{\partial v}{\partial y} &= v_y. \end{aligned}$$

eingeführt.

Heuristische Idee. Es sei f als Funktion $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ differenzierbar. Man schreibt also auch

$$f(z) = f(x, y) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}.$$

Nach Kapitel 17.2 gilt (für fixiertes z_0 mit $f_x = \begin{pmatrix} u_x \\ v_x \end{pmatrix} \in \mathbb{R}^2$ und

$$f_y = \begin{pmatrix} u_y \\ v_y \end{pmatrix} \in \mathbb{R}^2$$

$$\begin{aligned} f(z) &= f(z_0) + \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}(z_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \dots \\ &= f(z_0) + f_x(z_0)(x - x_0) + f_y(z_0)(y - y_0) + \dots \end{aligned}$$

Wegen (vgl. Kapitel 7.1)

$$\begin{aligned} x - x_0 &= \frac{1}{2}((z - z_0) + \overline{(z - z_0)}), \\ y - y_0 &= -\frac{i}{2}((z - z_0) - \overline{(z - z_0)}) \end{aligned}$$

folgt

$$\begin{aligned} f(z) &= f(z_0) + \frac{1}{2}f_x(z_0)((z - z_0) + \overline{(z - z_0)}) \\ &\quad - \frac{i}{2}f_y(z_0)((z - z_0) - \overline{(z - z_0)}) + \dots \\ &= f(z_0) + (z - z_0) \left[\frac{1}{2}(f_x(z_0) - if_y(z_0)) \right] \\ &\quad + \overline{(z - z_0)} \left[\frac{1}{2}(f_x(z_0) + if_y(z_0)) \right] + \dots \end{aligned} \quad (1)$$

Ist **andererseits f komplex differenzierbar**, so muss gelten

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \dots$$

Der Vergleich mit (1) zeigt

$$f'(z_0) = \frac{1}{2} [f_x(z_0) - if_y(z_0)], \quad (2)$$

$$0 = f_x(z_0) + if_y(z_0). \quad (3)$$

Die Gleichungen (2) stellen die **komplexe Ableitung in Termen der reellen partiellen Ableitungen** dar.

Die Gleichungen (3) heißen die **Cauchy- Riemannschen Differentialgleichungen**, wegen

$$\begin{aligned} f_x + if_y &= (u_x + iv_x) + i(u_y + iv_y) \\ &= (u_x - v_y) + i(v_x + u_y) \end{aligned}$$

schreibt man sie in der Form

$$u_x = v_y, \quad u_y = -v_x. \quad (*)$$

Eine genaue Argumentation liefert tatsächlich:

Satz 22.2.1

*Für eine Funktion $f: U \rightarrow \mathbb{C}$, $U \subset \mathbb{C}$ offen sind die folgenden Aussagen **äquivalent**:*

- i) f ist in $z_0 \in U$ komplex differenzierbar.*
- ii) f ist in $z_0 \in U$ reell differenzierbar und es gelten die Cauchy- Riemannschen Differentialgleichungen (*).*

Beispiele.

i) Es sei

$$\begin{aligned} f(z) &= 2i + 3z + 4z^2 \\ &= \underbrace{3x + 4x^2 - 4y^2}_{=u(x,y)} + i \underbrace{(2 + 3y + 8xy)}_{=v(x,y)}. \end{aligned}$$

Dann gilt

$$u_x = 3 + 8x = v_y,$$

$$u_y = -8y = -v_x,$$

die Cauchy- Riemannschen Differentialgleichung sind also erfüllt, f ist komplex differenzierbar.

ii) Es sei

$$f(z) = \bar{z} = x + i(-y) .$$

Hier gilt

$$u_x = 1 \neq -1 = v_y ,$$

die Cauchy- Riemannsches Differentialgleichungen sind also **nicht erfüllt**, wie bereits bekannt ist f nicht komplex differenzierbar.

iii) Es sei $f(z) = \operatorname{Re} z$. Auch hier können die Cauchy- Riemannsches Differentialgleichungen **nicht gelten**.

22.3 Kurvenintegrale (Das komplexe Integral) (Integrationsweg; Wegunabhängigkeit; Stammfunktion)

Identifiziert man die komplexe Ebene mit dem \mathbb{R}^2 , $\mathbb{C} \cong \mathbb{R}^2$, so überträgt sich der Begriff eine Kurve (Definition 17.1.1) direkt auf Kurven in der komplexen Ebene. Gleiches gilt für alle weiteren Begriffe aus Kapitel 17.1 (insbesondere den der Parametertransformation).

Notation.

i) Mit $\dot{\gamma}$ wird im Folgenden die Ableitung einer Kurve in \mathbb{C} nach dem reellen Parameter (der Zeit) bezeichnet. Ist also $\gamma: I = [a, b] \subset \mathbb{R} \rightarrow \mathbb{C}$,

$$\gamma(t) = \varphi(t) + i\psi(t) \cong \begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} ,$$

so ist

$$\dot{\gamma}(t) = \frac{d\gamma}{dt}(t) = \dot{\varphi}(t) + i\dot{\psi}(t) \cong \begin{pmatrix} \dot{\varphi}(t) \\ \dot{\psi}(t) \end{pmatrix} .$$

ii) Eine stückweise glatte Kurve (vgl. Bemerkungen nach Definition 18.1.1) $\gamma: [a, b] \rightarrow U \subset \mathbb{C}$ heißt im Folgenden ein **Integrationsweg in $U \subset \mathbb{C}$** .

iii) Für stetiges $\xi: [a, b] \subset \mathbb{R} \rightarrow \mathbb{C}$ ist

$$\int_a^b \xi(t) dt := \int_a^b \operatorname{Re} \xi(t) dt + i \int_a^b \operatorname{Im} \xi(t) dt .$$

Definition 22.3.1

Es sei $\gamma: [a, b] \rightarrow \mathbb{C}$ ein Integrationsweg und $f: \gamma(I) \rightarrow \mathbb{C}$ eine stetige Funktion. Dann ist das (komplexe) **Kurvenintegral** (**Wegintegral**) längs γ definiert als

$$\int_{\gamma} f(z) \, dz := \int_a^b f(\gamma(t)) \dot{\gamma}(t) \, dt .$$

Bemerkungen.

- i) Die Bildung des komplexen Integrals erfolgt analog zu der des reellen Kurvenintegrals, wobei **das Skalarprodukt durch die komplexe Multiplikation zu ersetzen ist**.
- ii) Die Invarianz unter orientierungserhaltenden Parametertransformationen und der Vorzeichenwechsel bei orientierungsumkehrenden Parametertransformationen sind genau wie in Kapitel 17.1 zu zeigen. Dementsprechend kann wieder von Wegen gesprochen werden, auf die genaue Unterscheidung wird im Folgenden wie üblich nicht immer eingegangen.
- iii) Es gelten wieder die bekannten Regeln (**Linearität, Beschränktheit, Wegadditivität**).
- iv) Ist $\gamma(t) = \varphi(t) + i\psi(t)$, $f(z) = u(x, y) + iv(x, y)$, so gilt

$$\begin{aligned} \int_{\gamma} f(z) \, dz &= \int_a^b [u(\varphi(t), \psi(t)) + iv(\varphi(t), \psi(t))] [\dot{\varphi}(t) + i\dot{\psi}(t)] \, dt \\ &= \int_a^b [u(\varphi(t), \psi(t))\dot{\varphi}(t) - v(\varphi(t), \psi(t))\dot{\psi}(t)] \, dt \\ &\quad + i \int_a^b [v(\varphi(t), \psi(t))\dot{\varphi}(t) + u(\varphi(t), \psi(t))\dot{\psi}(t)] \, dt . \end{aligned}$$

Beispiele.

- i) Es seien $z_0 \in \mathbb{C}$ und $r > 0$ fixiert. Man betrachte den Integrationsweg (**positiv orientierte Kreislinie**, vgl. Abbildung 22.2)

$$\gamma: [0, 2\pi] \rightarrow \mathbb{C}, \quad t \mapsto z_0 + re^{it} .$$

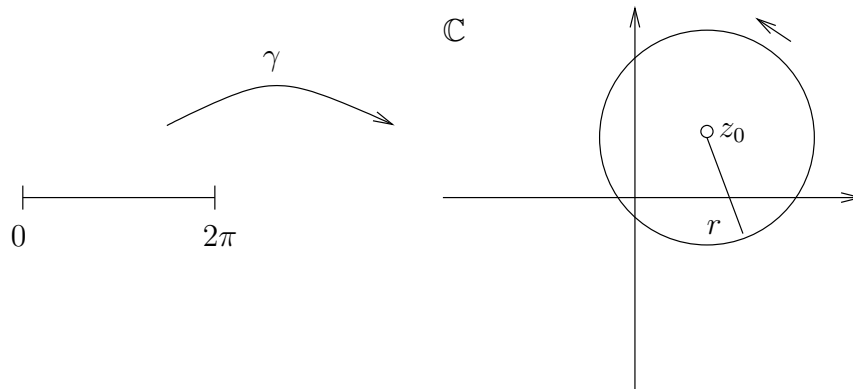


Abbildung 22.2: Eine positiv orientierte Kreislinie.

Es gilt

$$\begin{aligned} \gamma(t) &= z_0 + r(\cos(t) + i \sin(t)) , \\ \dot{\gamma}(t) &= -r \sin(t) + ir \cos(t) = ire^{it} . \end{aligned}$$

ii) Es sei $f(z) = |z|$ (insbesondere ist f nicht holomorph). Betrachtet sei zunächst der Integrationsweg

$$\gamma : [0, \pi] \rightarrow \mathbb{C} , \quad t \mapsto e^{i(\pi-t)} .$$

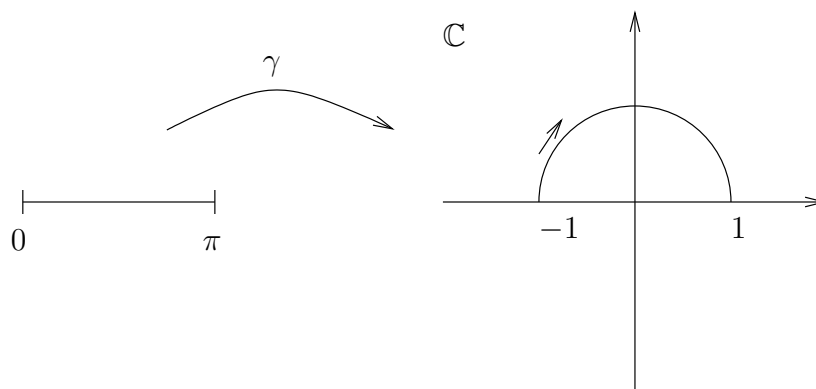


Abbildung 22.3: Der Integrationsweg γ aus Beispiel ii).

Mit der Notation $\gamma = \varphi + i\psi$ folgt nach dem Hauptsatz der Differential- und Integralrechnung für Funktionen einer reellen

Veränderlichen

$$\begin{aligned}
 \int_{\gamma} |z| \, dz &= \int_0^{\pi} 1 \dot{\gamma}(t) \, dt \\
 &= \int_0^{\pi} \dot{\varphi}(t) \, dt + i \int_0^{\pi} \dot{\psi}(t) \, dt \\
 &= \varphi(\pi) - \varphi(0) + i(\psi(\pi) - \psi(0)) \\
 &= \gamma(\pi) - \gamma(0) = 2.
 \end{aligned}$$

Betrachtet sei jetzt der Integrationsweg

$$\tilde{\gamma} : [-1, 1] \rightarrow \mathbb{C}, \quad t \mapsto t.$$

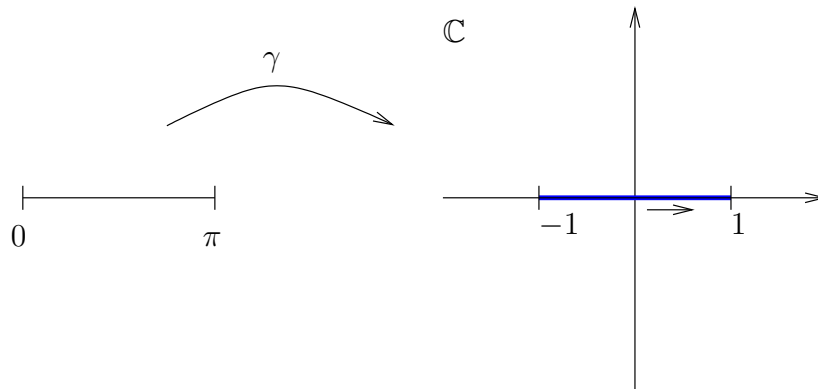


Abbildung 22.4: Der Integrationsweg $\tilde{\gamma}$ aus Beispiel ii).

Hier gilt

$$\int_{\tilde{\gamma}} |z| \, dz = \int_{-1}^1 |t| \, dt = 1 \neq \int_{\gamma} |z| \, dz.$$

Wieder hängen Kurvenintegrale i.A. nicht nur vom Anfangs- und Endpunkt des Integrationsweges ab.

Das wichtigste Beispiel ist das Folgende:

Satz 22.3.1

Es seien $z_0 \in \mathbb{C}$, $r > 0$ fixiert und

$$\gamma : [0, 2\pi] \rightarrow \mathbb{C}, \quad t \mapsto z_0 + re^{it}.$$

Dann gilt

$$\int_{\gamma} z^k dz = \begin{cases} 0 & \text{für } k \in \mathbb{Z} - \{-1\}, \\ 2\pi i & \text{für } k = -1. \end{cases}$$

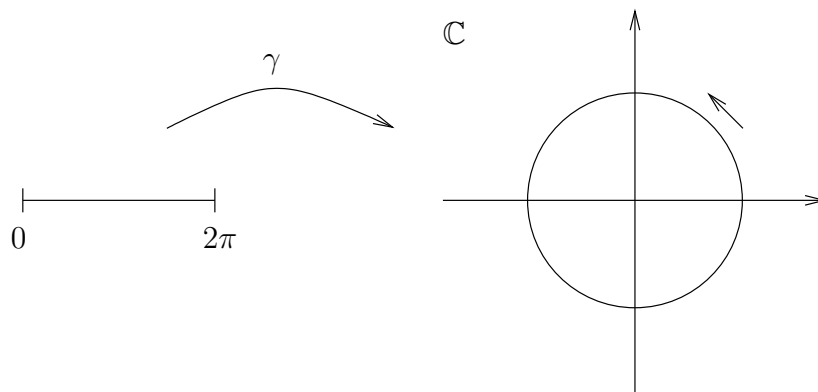


Abbildung 22.5: Zu Satz 22.3.1.

Beweis. O.E. sei $z_0 = 0$.

i) Es sei zunächst $k \in \mathbb{N}_0$. Dann gilt

$$\begin{aligned} \int_{\gamma} z^k dz &= \int_0^{2\pi} r^k e^{ikt} i r e^{it} dt \\ &= i r^{k+1} \int_0^{2\pi} e^{i(k+1)t} dt \\ &= i r^{k+1} \left[\int_0^{2\pi} \cos((k+1)t) dt + i \int_0^{2\pi} \sin((k+1)t) dt \right] \\ &= i r^{k+1} \left[\left[\frac{\sin((k+1)t)}{k+1} \right]_0^{2\pi} + \left[-i \frac{\cos((k+1)t)}{k+1} \right]_0^{2\pi} \right] \\ &= 0. \end{aligned}$$

ii) Exakt die gleiche Rechnung liefert den Satz im Fall $k = -2, -3, -4 \dots$

iii) Es sei schließlich $k = -1$.

$$\int_{\gamma} \frac{1}{z} dz = \int_0^{2\pi} r^{-1} e^{-it} i r e^{it} dt = i \int_0^{2\pi} dt = 2\pi i .$$

□

Bemerkung. Obwohl die **Singularität** von beispielsweise z^{-2} im Ursprung “schlimmer aussieht” als die von z^{-1} , verschwindet obiges Kurvenintegral über z^{-2} , das über z^{-1} verschwindet nicht.

Kriterien zur **Wegunabhängigkeit** des Kurvenintegrals?

Das erste Kriterium ist ein Analogon zu Satz 18.1.1. Man benötigt dazu:

Definition 22.3.2

*Es sei $U \subset \mathbb{C}$ offen und $f: U \rightarrow \mathbb{C}$ stetig. Eine Funktion $F: U \rightarrow \mathbb{C}$ heißt **Stammfunktion** von f , falls F holomorph ist und $F' = f$ gilt.*

Bemerkung. Ist eine Funktion h auf einem **Gebiet** G holomorph mit $h' \equiv 0$, so ist h konstant. Konsequenz: Auf einem Gebiet G sind Stammfunktionen (**falls existent**) bis auf Konstanten eindeutig bestimmt. (Warum kann die Aussage nicht für beliebige offene Mengen gelten?)

Satz 22.3.2

i) Es sei $f: U \rightarrow \mathbb{C}$ eine stetige Funktion, die **eine Stammfunktion F besitze**. $\gamma: [a, b] \rightarrow U$ sei ein Integrationsweg in U von $\gamma(a) = z_0$ nach $\gamma(b) = z_1$. Dann ist

$$\int_{\gamma} f(z) \, dz = F(z_1) - F(z_0) .$$

ii) Ist insbesondere γ **geschlossen** ($z_0 = z_1$), so folgt

$$\int_{\gamma} f(z) \, dz = 0 .$$

Beweis. Der Beweis folgt leicht aus der folgenden Proposition, einer Art Kettenregel für Kurven in der komplexen Ebene. Man beachte den Unterschied zu Satz 22.1.1, iii). \square

Proposition 22.3.1

Es sei $F: U \rightarrow \mathbb{C}$ holomorph auf der offenen Menge U und $\gamma: [a, b] \rightarrow U$ eine glatte Kurve in U . Dann gilt

$$\frac{d}{dt} F(\gamma(t)) = F'(\gamma(t)) \dot{\gamma}(t) .$$

Bemerkung. Es sei nochmals betont, dass F' die **komplexe Ableitung** bezeichnet, wohingegen d/dt und $\dot{\gamma}$ **Ableitungen nach der reellen Variablen t** bezeichnen.

Beweis der Proposition. Nach Formel (2) aus Kapitel 22.2 ist ($F = u(x, y) + iv(x, y)$)

$$\begin{aligned} F' &= \frac{1}{2} \{ (u_x + iv_x) - i(u_y + iv_y) \} \\ &= \frac{1}{2} \{ (u_x + v_y) + i(v_x - u_y) \} . \end{aligned}$$

Es folgt

$$\begin{aligned} F'(\gamma(t))\dot{\gamma}(t) &= \frac{1}{2} \left\{ [(u_x + v_y)\dot{\gamma}_1 - (v_x - u_y)\dot{\gamma}_2] \right. \\ &\quad \left. + i[(u_x + v_y)\dot{\gamma}_2 + (v_x - u_y)\dot{\gamma}_1] \right\} \\ &= [u_x\dot{\gamma}_1 + u_y\dot{\gamma}_2] + i[v_y\dot{\gamma}_2 + v_x\dot{\gamma}_1], \end{aligned}$$

wobei die Cauchy-Riemannschen Differentialgleichungen ausgenutzt wurden und wobei die Ableitungen von u und v an der Stelle $\gamma(t)$ auszuwerten sind.

Andererseits ist nach der (reellen) Kettenregel (Satz 17.2.1)

$$\begin{aligned} \frac{d}{dt}F(\gamma(t)) &= \frac{d}{dt}[u(\gamma_1(t), \gamma_2(t)) + iv(\gamma_1(t), \gamma_2(t))] \\ &= [u_x\dot{\gamma}_1 + u_y\dot{\gamma}_2] + i[v_x\dot{\gamma}_1 + v_y\dot{\gamma}_2] \end{aligned}$$

und die Proposition ist bewiesen. \square

Es gilt auch die Umkehrung von Satz 22.3.2 im Sinne von

Satz 22.3.3

Es sei f auf einem Gebiet G stetig. Für jeden geschlossenen Integrationsweg γ in G gelte

$$\int_{\gamma} f(z) dz = 0.$$

Dann hat f auf G eine Stammfunktion.

Beispiele.

i) Die Funktion $f(z) = z^k$, $k \in \mathbb{Z}$, $k \neq -1$, hat auf \mathbb{C} bzw. $\mathbb{C} - \{0\}$ im Fall $k < 0$ die Stammfunktion

$$F(z) = \frac{1}{k+1} z^{k+1} \quad (+\text{konst.}),$$

also folgt

$$\int_{\gamma} z^k dz = \frac{1}{k+1} (z_1^{k+1} - z_0^{k+1}),$$

wobei z_0, z_1 den Anfangs- bzw. Endpunkt von γ bezeichnen.

ii) Ist $\gamma: [0, 2\pi] \rightarrow \mathbb{C}$, $\gamma(t) = re^{it}$, $r > 0$ fixiert, so gilt nach Satz 22.3.1:

$$\int_{\gamma} \frac{1}{z} dz = 2\pi i .$$

Die auf $\mathbb{C} - \{0\}$ holomorphe Funktion z^{-1} kann dort keine Stammfunktion besitzen (vgl. das Beispiel “unendlich langer Leiter” aus Kapitel 18). Dies ist ein wesentlicher Unterschied zu der Funktion z^{-n} , $n > 1$.

Kapitel 23

Cauchys Integralsatz und Cauchys Integralformel

23.1 Der Cauchysche Integralsatz (einfach zusammenhängend; einfache geschlossene Kurven; Fresnelsche Integrale)

Wird die Voraussetzung “ f habe eine Stammfunktion” weggelassen, so zeigt der Cauchysche Integralsatz, dass für holomorphe Funktionen die Frage nach der Wegunabhängigkeit des Kurvenintegrals zurückgeführt werden kann auf eine Eigenschaft des zugrunde liegenden Gebietes. Es ist herauszustellen, dass eine Analogon für Kurvenintegrale differenzierbarer Funktionen in der reellen Analysis nicht gelten kann.

Definition 23.1.1

Ein Gebiet G (offen und zusammenhängend) heißt einfach zusammenhängend, wenn jeder in G verlaufende geschlossene (Anfangs- gleich Endpunkt) doppelpunktfreie ($\gamma(t_1) \neq \gamma(t_2)$ für alle $t_1 \neq t_2$ mit $a \leq t_1, t_2 < b$) Polygonzug $\gamma: [a, b] \rightarrow G$ nur Punkte von G umschließt.

Bemerkungen.

- i) Die Begriffe “geschlossen” und “doppelpunktfrei” sind für Kurven bzw. Integrationswege natürlich analog definiert. Geschlossene, doppelpunktfreie Kurven (...) heißen im Folgenden einfache geschlossene Kurven (...).

- ii)* Da in der obigen Definition lediglich einfache geschlossene Polygonzüge betrachtet werden, bereitet der Begriff “umschließen” keine Schwierigkeiten.
- iii)* Anschaulich bedeutet die Definition: **“ G hat keine Löcher”**.

Beispiele.

- i)* Beispiele von geschlossenen bzw. doppelpunktfreien Integrationswegen sind in Abbildung 23.1 angedeutet.

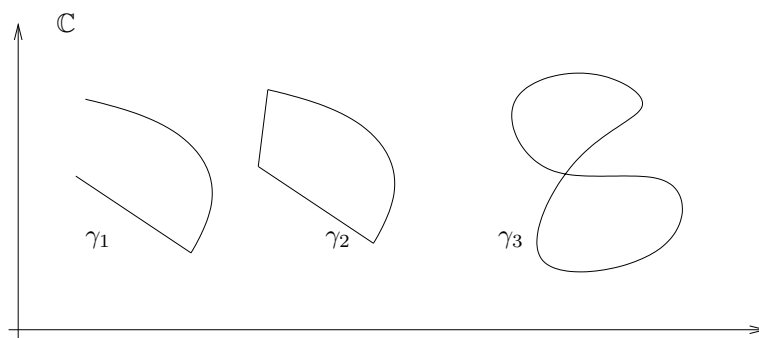


Abbildung 23.1: Integrationswege: Geschlossen sind γ_2 , γ_3 , doppelpunktfrei γ_1 , γ_2 .

- ii)* Die Menge G_1 aus Abbildung 23.2 ist einfach zusammenhängend, G_2 ist es nicht.

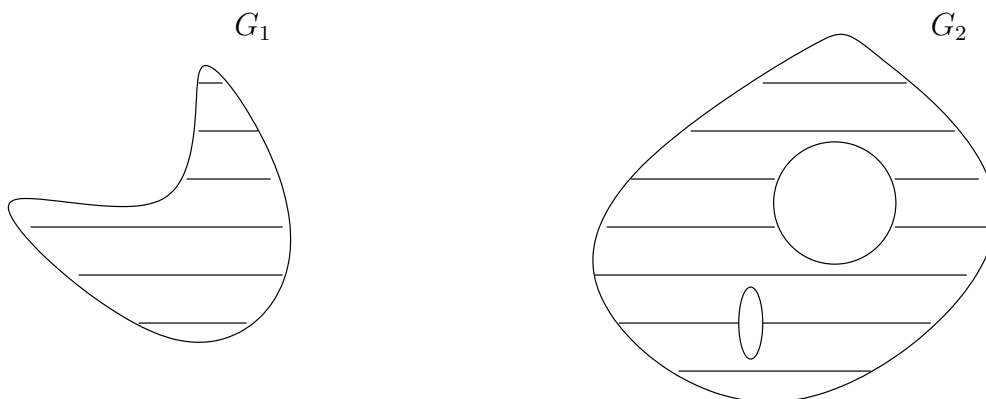


Abbildung 23.2: Zum Begriff “einfach zusammenhängend”.

- iii)* Die Einheitskreisscheibe $\{z \in \mathbb{C} : |z| < 1\}$ ist einfach zusammenhängend.

- iv) Das Gebiet $\{z \in \mathbb{C} : z \neq 0\}$ ist **nicht einfach zusammenhängend**. Auf dieses Beispiel wird noch oft zurückzukommen sein.

Satz 23.1.1 (*Der Cauchysche Integralsatz*)

*Es sei $G \subset \mathbb{C}$ ein **einfach zusammenhängendes Gebiet** und f **holomorph** auf G . Dann gilt für jeden **geschlossenen Integrationsweg in G** :*

$$\int_{\gamma} f(z) \, dz = 0 .$$

Beweisidee. Der Beweis von Satz 23.1.1 wird zunächst für den Fall ausgeführt, dass γ die Berandung eines Dreiecks parametrisiert, dann für geschlossene Polygonzüge und schließlich approximativ im allgemeinen Fall. \square

Bemerkungen.

- i) Ist wie oben $\gamma(t) = re^{it}$, so folgt aus

$$\int_{\gamma} \frac{1}{z} \, dz = 2\pi i$$

kein Widerspruch, da $\mathbb{C} - \{0\}$ **nicht einfach zusammenhängend** ist.

- ii) Unter den Voraussetzungen von Satz 23.1.1 folgt die Wegunabhängigkeit des Kurvenintegrals.

- iii) Mit Hilfe des Cauchyschen Integralsatzes können auch **reelle Integrale** berechnet werden.

Beispiel. (**Fresnelsche Integrale**) Es soll die Behauptung

$$\int_0^{\infty} \sin(t^2) \, dt = \int_0^{\infty} \cos(t^2) \, dt = \frac{1}{2} \sqrt{\pi/2}$$

mit Hilfe des Cauchyschen Integralsatzes verifiziert werden. Betrachtet wird dazu das Gebiet $G = \mathbb{C}$ sowie das Kurvenintegral

$$\int_{\gamma} e^{-z^2} \, dz ,$$

wobei der geschlossene Integrationsweg stückweise gegeben ist durch ($R > 0$ fixiert)

$$\gamma_1(t) = t \quad \text{für } t \in [0, R] \quad \text{d.h.} \quad \dot{\gamma}_1(t) = 1 ,$$

$$\gamma_2(t) = Re^{it} \quad \text{für } t \in [0, \pi/4] \quad \text{d.h.} \quad \dot{\gamma}_2(t) = iRe^{it} ,$$

$$\gamma_3(t) = -te^{i\pi/4} \quad \text{für } t \in [-R, 0] \quad \text{d.h.} \quad \dot{\gamma}_3(t) = -e^{i\pi/4} .$$

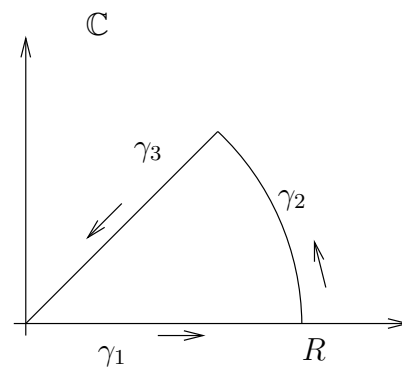


Abbildung 23.3: Der zusammengesetzte Integrationsweg γ .

Die Funktion $f(z) = e^{-z^2}$ ist holomorph, G ist einfach zusammenhängend, der aus γ_1 , γ_2 und γ_3 zusammengesetzte Weg γ ist ein geschlossener Integrationsweg (stückweise glatt), also folgt aus dem Cauchyschen Integralsatz

$$\int_{\gamma} e^{-z^2} dz = 0 .$$

Mit Hilfe der Wegadditivität erhält man

$$\begin{aligned} 0 &= \int_{\gamma} e^{-z^2} dz \\ &= \int_0^R e^{-t^2} 1 dt + \int_0^{\pi/4} e^{-R^2 e^{2it}} iRe^{it} dt - \int_{-R}^0 e^{-t^2 e^{i\pi/2}} e^{i\pi/4} dt , \end{aligned}$$

das ergibt

$$\begin{aligned}
 0 &= \underbrace{\int_0^R e^{-t^2} dt}_{=: I_1(R)} + iR \underbrace{\int_0^{\frac{\pi}{4}} e^{-R^2[\cos(2t)+i\sin(2t)]+it} dt}_{=: I_2(R)} \\
 &\quad - \underbrace{e^{i\pi/4} \int_{-R}^0 e^{-it^2} dt}_{=: I_3(R)} .
 \end{aligned}$$

Es gilt (zum Beweis vgl. etwa Hildebrandt, Analysis)

$$\lim_{R \rightarrow \infty} I_1(R) = \frac{1}{2} \sqrt{\pi} .$$

Mit Hilfe von $|e^{i\alpha}| = 1$ kann man weiter zeigen:

$$\lim_{R \rightarrow \infty} |I_2(R)| = 0 .$$

Wegen (Substitution $\tau = -t$)

$$\int_{-R}^0 e^{-it^2} dt = \int_0^R e^{-it^2} dt$$

ist also

$$0 = \frac{1}{2} \sqrt{\pi} - \lim_{R \rightarrow \infty} \frac{1}{2} (\sqrt{2} + i\sqrt{2}) \int_0^R (\cos(t^2) - i \sin(t^2)) dt .$$

Die Trennung in Real- und Imaginärteil ergibt

$$\begin{aligned}
 0 &= \frac{1}{2} \sqrt{\pi} - \frac{1}{2} \sqrt{2} \int_0^\infty (\cos(t^2) + \sin(t^2)) dt , \\
 0 &= \int_0^\infty (\cos(t^2) - \sin(t^2)) dt ,
 \end{aligned}$$

d.h. die Behauptung. □

Bemerkungen.

- i)* Es gibt auch **Verallgemeinerungen des Cauchyschen Integralsatzes für nicht einfach zusammenhängende Gebiete**. Man betrachte etwa die in Abbildung 23.4 dargestellte Situation:

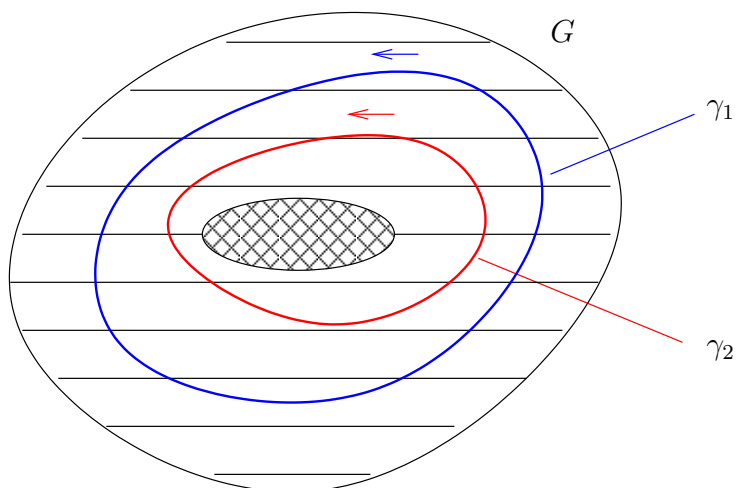


Abbildung 23.4: Zur Verallgemeinerung des Cauchyschen Integralsatzes.

Dann gilt die Behauptung

$$\int_{\gamma_1} f(z) dz = \int_{\gamma_2} f(z) dz .$$

Zu beachten ist: Die einzelnen Integrale **verschwinden i.A. nicht**.

Die obige Gleichheit sieht man mit einer **typischen Idee** ein: Gegeben seien zwei Integrationswege γ_1, γ_2 wie auf der linken Seite von Abbildung 23.5 angedeutet.

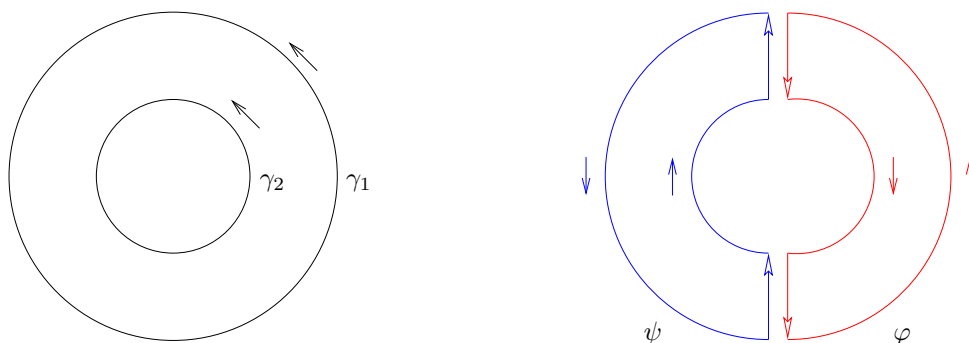


Abbildung 23.5: Zerlegung in einfach zusammenhängende Gebiete.

Mit vertikalen “Schnitten” wird das (nicht einfach zusammenhängende) Gebiet in zwei einfach zusammenhängende Teilgebiete zerlegt, die von der roten Kurve φ bzw. von der blauen Kurve ψ (siehe rechte Seite von Abbildung 23.5) berandet sind. Nun kann

der Cauchysche Integralsatz angewendet werden:

$$\int_{\varphi} f(z) \, dz = 0, \quad \int_{\psi} f(z) \, dz = 0.$$

Anschließend überlegt man sich:

- Die beiden inneren Halbkreise entsprechen zusammen der Kurve $-\gamma_2$.
- Die beiden äußeren Halbkreise entsprechen zusammen der Kurve γ_1 .
- Die Integrale über die “vertikalen Schnitte” heben sich gegenseitig auf.

Insgesamt ist so gezeigt:

$$\begin{aligned} 0 &= \int_{\varphi} f(z) \, dz + \int_{\psi} f(z) \, dz \\ &= \int_{\gamma_1} f(z) \, dz - \int_{\gamma_2} f(z) \, dz, \end{aligned}$$

also die Behauptung. □

ii) Eine wichtige Konsequenz aus i) ist: Es sei $G = \mathbb{C} - \{0\}$ und

$$f(z) = \frac{1}{z}.$$

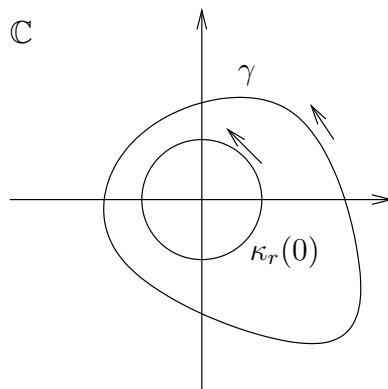
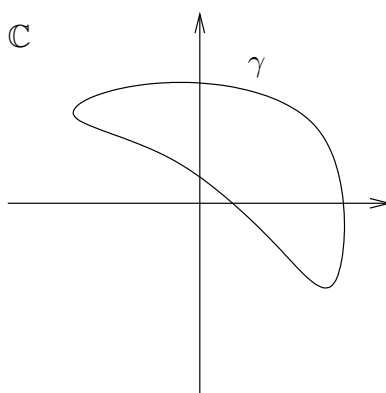
Dann gilt für **jeden (positiv orientierten) einfachen geschlossenen Integrationsweg γ , der den Nullpunkt im Inneren enthält** (Zur Erinnerung: Positiv orientiert bedeutet, dass beim Durchlaufen der Kurve das Innere auf der linken Seite liegt):

$$\int_{\gamma} \frac{1}{z} \, dz = \int_{\kappa_r(0)} \frac{1}{z} \, dz = 2\pi i.$$

Dabei ist $\kappa_r(0)$ wie in Abbildung 23.6 angedeutet eine Parametrisierung der positiv orientierten Kreislinie vom Radius r um 0.

Enthält γ den Nullpunkt **nicht im Innern**, so gilt (warum?)

$$\int_{\gamma} \frac{1}{z} \, dz = 0.$$

Abbildung 23.6: γ enthält den Nullpunkt im Innern.Abbildung 23.7: γ enthält den Nullpunkt nicht im Innern.

23.2 Die Cauchysche Integralformel (Existenz beliebiger Ableitungen holomorpher Funktionen; Harmonizität von Real- und Imaginärteil holomorpher Funktionen; Satz von Liouville)

Aus dem Cauchyschen Integralsatz und dessen Verallgemeinerung folgt eine Aussage, die **in der reellen Analysis völlig falsch** ist:

Allein aus der Kenntnis einer holomorphen Funktion f auf dem Rand einer Kugel $\partial B_r(z_0)$ ergeben sich die Funktionswerte auf der ganzen Kugel $B_r(z_0)$. M.a.W.:

Die kleinste Änderung von f im Innern von $B_r(z_0)$ bei festgehaltenen Werten auf $\partial B_r(z_0)$ zerstört unmittelbar die Differenzierbarkeit von f .

Präzise lautet der Satz für die Kreisscheibe:

Satz 23.2.1 (*Cauchys Integralformel für die Kreisscheibe*)

Es sei f holomorph auf einer offenen Menge $U \subset \mathbb{C}$ und es gelte $B_r(z_0) \Subset U$. Dann ist

$$f(w) = \frac{1}{2\pi i} \int_{\kappa_r(z_0)} \frac{f(z)}{z - w} dz \quad \text{für alle } w \in B_r(z_0).$$

Hier und im folgenden bezeichnet $\kappa_r(z_0)$ die Parametrisierung $\kappa_r(z_0)(t) = z_0 + re^{it}$, $t \in [0, 2\pi]$, der Kreislinie vom Radius r um den Punkt z_0 .

Bemerkungen.

i) Die Voraussetzungen des Satzes sind in Abbildung 23.8 skizziert.

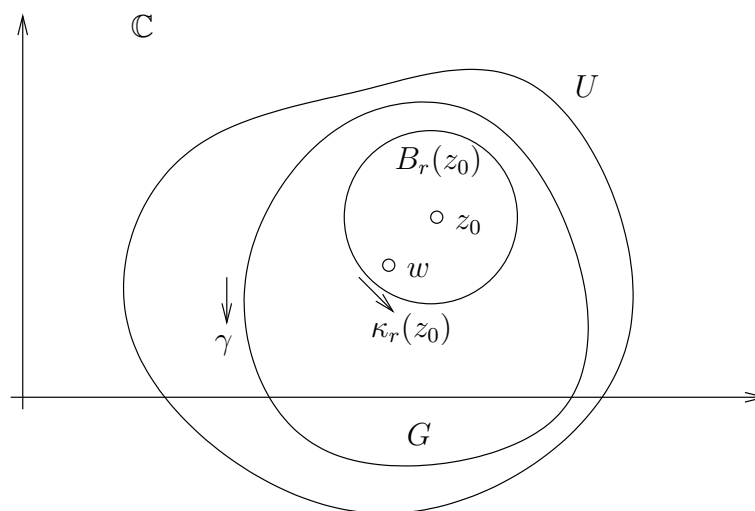


Abbildung 23.8: Zur Cauchyschen Integralformel.

ii) Der Cauchysche Integralsatz gilt nicht nur für Kreisscheiben, er kann analog für einfach zusammenhängende Gebiete $G \Subset U$ formuliert werden (vgl. ebenfalls Abbildung 23.8).

iii) Für $w \in U - \overline{B_r(z_0)}$ gilt natürlich (warum?)

$$\frac{1}{2\pi i} \int_{\kappa_r(z_0)} \frac{f(z)}{z - w} dz = 0.$$

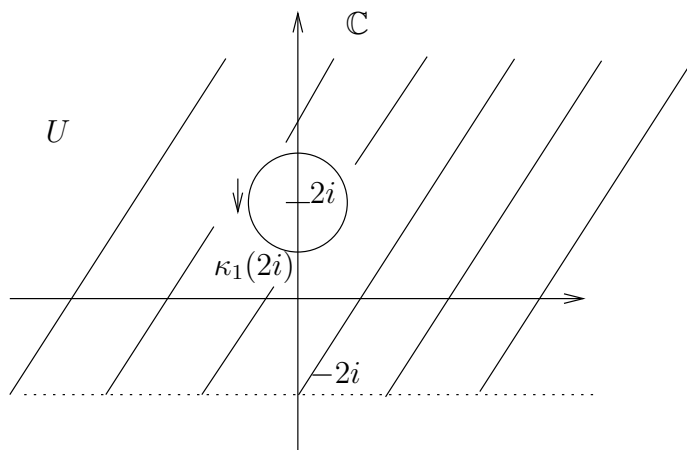


Abbildung 23.9: Ein Beispiel zur Cauchyschen Integralformel.

Beispiel. Es sei $U = \{z : \operatorname{Im} z > -2\}$, $z_0 = 2i$ und $r = 1$, $w = 2i$. Weiterhin sei

$$f(z) = \frac{1}{z + 2i}.$$

Aus

$$\frac{1}{z^2 + 4} = \frac{1}{z + 2i} = \frac{f(z)}{z - w}$$

folgt

$$\int_{\kappa_1(2i)} \frac{1}{z^2 + 4} dz = \int_{\kappa_1(2i)} \frac{f(z)}{z - 2i} dz = 2\pi i f(2i) = 2\pi i \frac{1}{4i} = \frac{\pi}{2}.$$

Differenziert man in Satz 23.2.1 formal unter dem Integralzeichen, so ergibt sich sogar:

Satz 23.2.2

Eine holomorphe Funktion $f: U \rightarrow \mathbb{C}$ ($U \subset \mathbb{C}$ offen) ist **beliebig oft differenzierbar**. Damit sind alle komplexen Ableitungen f' , f'' , \dots , $f^{(n)}$, \dots in U holomorphe Funktionen und für jede Kreisscheibe $B_r(z_0) \Subset U$ und alle $w \in B_r(z_0)$ gilt

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_{\kappa_r(z_0)} \frac{f(z)}{(z - w)^{n+1}} dz.$$

Bemerkungen.

- i) Wieder ist auch eine Version für einfach zusammenhängende Gebiete richtig.
- ii) Eine auf U komplex differenzierbare Funktion ist also automatisch beliebig oft differenzierbar. Wieder wird der Unterschied zur reellen Analysis deutlich.
- iii) Ist $f = u + iv$ holomorph auf U , so ist f nach Satz 23.2.2 beliebig oft differenzierbar und die Cauchy- Riemannschen Differentialgleichungen

$$u_x = v_y, \quad u_y = -v_x$$

können differenziert werden. Wird die erste Gleichung nach x differenziert, die zweite nach y , so folgt

$$u_{xx} = v_{yx}, \quad u_{yy} = -v_{xy}$$

und aus der Vertauschbarkeit der Ableitungen folgt

$$u_{xx} + u_{yy} = \Delta u = 0.$$

Analog sieht man

$$v_{xx} + v_{yy} = \Delta v = 0$$

ein, d.h.: **Real- und Imaginärteil holomorpher Funktionen sind harmonisch.**

Beispiele.

- i) Es seien $w = z_0 \in \mathbb{C}$ fixiert, $f(z) \equiv 1$. Dann ist (für alle $r > 0$)

$$f'(z_0) = 0 = \frac{1}{2\pi i} \int_{\kappa_r(z_0)} \frac{1}{(z - z_0)^2} dz.$$

- ii) Es sei $w = z_0 = 0$ und

$$f(z) = e^{iz}, \quad \text{also} \quad f''(z) = -e^{iz}.$$

Es folgt

$$f''(0) = -1 = \frac{2!}{2\pi i} \int_{\kappa_1(0)} \frac{e^{iz}}{z^3} dz, \quad \text{also} \quad \int_{\kappa_1(0)} \frac{e^{iz}}{z^3} dz = -\pi i.$$

Satz 23.2.1 bzw. Satz 23.2.2 haben zahlreiche wichtige Folgerungen exemplarisch sei genannt

Satz 23.2.3 (*Satz von Liouville*)

*Ist f auf ganz \mathbb{C} holomorph (Bezeichnung: f ist eine **ganze Funktion**) und beschränkt (d.h. $|f| \leq \text{konst.}$), so ist f konstant.*

Bemerkung. Man vergleiche Satz 23.2.3 wieder mit dem reellen Fall, d.h. man vergleiche etwa \sin und \cos im Reellen und im Komplexen.

Kapitel 24

Entwicklungen holomorpher Funktionen

Reihenentwicklungen spielen in der Funktionentheorie eine ganz besondere Rolle. Im Reellen wurden Potenzreihen in Kapitel 5.2 besprochen, das komplexe Gegenstück wurde in Kapitel 7.2 behandelt. Auf die dort gelegten Grundlagen wird hier aufgebaut.

24.1 Taylor-Reihen (Potenzreihen und holomorphe Funktionen; Differentiation von Potenzreihen)

Cauchys Beispiel aus Kapitel 13.1 zeigt, dass im Reellen selbst unendlich oft differenzierbare Funktion nicht unbedingt durch ihre Taylor-Reihe dargestellt werden. Auch diesbezüglich sieht die Situation in der komplexen Analysis anders aus, wie der vorliegende Paragraph zeigen wird.

Zunächst sei jedoch an die wichtigste Potenzreihe, die **geometrische Reihe** erinnert:

Beispiel. Man betrachte die Reihe (vgl. Kapitel 7.2)

$$\sum_{n=0}^{\infty} z^n .$$

Für die Partialsummen gilt

$$\sum_{n=0}^N z^n = \frac{1 - z^{N+1}}{1 - z} , \quad \text{falls } z \neq 1 .$$

Für $|z| < 1$ ist $\lim_{N \rightarrow \infty} z^{N+1} = 0$, also folgt

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z} \quad \text{für } |z| < 1.$$

Für $|z| \geq 1$ ist $\{z^n\}$ keine Nullfolge, die geometrische Reihe **divergiert**.

Nun soll die Ableitung der Funktion $1/(1-z)$ mit Hilfe des Cauchy-Produktes für Reihen (Satz 5.3.1) analysiert werden: Für $|z| < 1$ ist

$$\begin{aligned} \left(\frac{1}{1-z}\right)' &= \frac{1}{(1-z)^2} = \frac{1}{1-z} \frac{1}{1-z} \\ &= \sum_{n=0}^{\infty} z^n \sum_{n=0}^{\infty} z^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n z^k z^{n-k}\right) \\ &= \sum_{n=0}^{\infty} (n+1)z^n. \end{aligned}$$

Mit $j = n + 1$ ist gezeigt

$$\frac{1}{(1-z)^2} = \left(\frac{1}{1-z}\right)' = \sum_{j=1}^{\infty} j z^{j-1} = \sum_{n=1}^{\infty} n z^{n-1} = \sum_{n=0}^{\infty} (z^n)'$$

In Verallgemeinerung des Beispiels besagt der folgende Satz, dass durch Potenzreihen holomorphe Funktionen definiert sind und dass sich die Ableitung durch gliedweise Differentiation berechnen lässt (vgl. auch Satz 11.1.7).

Satz 24.1.1 (*Differentiation von Potenzreihen*)

Besitzt eine Potenzreihe $P(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ den Konvergenzradius $r > 0$, so ist dadurch in der Kreisscheibe $B_r(z_0)$ eine **holomorphe Funktion** $f(z)$ mit

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}$$

als Ableitung gegeben. Diese Potenzreihe hat **denselben Konvergenzradius wie $P(z)$** .

Beispiel. Man betrachte

$$f(z) = e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} .$$

Es ist $r = \infty$ und

$$f'(z) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!} = \sum_{k=0}^{\infty} \frac{z^k}{k!} = e^z .$$

Nach Satz 24.1.1 sind durch Potenzreihen holomorphe Funktionen definiert. Umgekehrt stellt sich die Frage, ob holomorphe Funktionen immer als Potenzreihe, d.h. durch ihre **Taylor-Reihe**, dargestellt werden können.

Satz 24.1.2

*Es sei $f(z)$ holomorph für $|z - z_0| < r$ mit $z_0 \in \mathbb{C}$ und $r > 0$. Dann ist f **eindeutig darstellbar durch eine Potenzreihe** um den Entwicklungspunkt z_0 :*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n .$$

Die Reihe konvergiert in $B_r(z_0)$, und die Koeffizienten sind gegeben durch

$$a_n = \frac{f^{(n)}(z_0)}{n!} .$$

Bemerkungen.

i) Nach Satz 23.2.2 gilt für $0 < \rho < r$

$$a_n = \frac{1}{2\pi i} \int_{\kappa_\rho(z_0)} \frac{f(z)}{(z - z_0)^{n+1}} dz .$$

Hier kann $\kappa_\rho(z_0)$ auch durch einen beliebigen, einfachen geschlossenen Integrationsweg ersetzt werden, der im mathematisch positiven Sinne durchlaufen wird, ganz in $B_r(z_0)$ liegt und z_0 im Innern enthält.

- ii) Ist f in einer offenen Menge U holomorph, $z_0 \in U$, so **konvergiert die Taylor-Reihe in der größten Kreisscheibe um z_0 , die noch in U liegt.**

Typisches Beispiel. Für $z \in U = \mathbb{C} - \{1, -2\}$ sei

$$f(z) = \frac{3}{(1-z)(2+z)}.$$

Gesucht sind die Koeffizienten der Taylor-Reihe um $z_0 = i$, die nach Bemerkung ii) für $|z - i| < \sqrt{2}$ konvergiert.

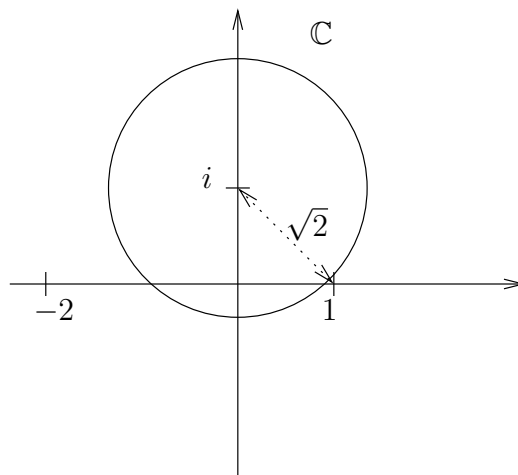


Abbildung 24.1: Zur Taylorreihe von f um i .

Die Koeffizienten können im Prinzip nach Bemerkung i) berechnet werden, es gilt aber auch (vgl. die Partialbruchzerlegung aus Kapitel 12.3)

$$f(z) = \frac{3}{(1-z)(2+z)} = \frac{1}{1-z} + \frac{1}{2+z}.$$

Für $|z - i| < \sqrt{2}$ gilt $|\frac{z-i}{1-i}| < 1$, und mit Hilfe der geometrischen Reihe findet man

$$\begin{aligned} \frac{1}{1-z} &= \frac{1}{1-i-(z-i)} = \frac{1}{1-i} \frac{1}{1-\frac{z-i}{1-i}} \\ &= \frac{1}{1-i} \sum_{n=0}^{\infty} \left(\frac{z-i}{1-i}\right)^n. \end{aligned}$$

Analog gilt wegen $|(z - i)/(2 + i)| < 1$

$$\begin{aligned} \frac{1}{2+z} &= \frac{1}{2+i+(z-i)} = \frac{1}{2+i} \frac{1}{1+\frac{z-i}{2+i}} \\ &= \frac{1}{2+i} \sum_{n=0}^{\infty} (-1)^n \left(\frac{z-i}{2+i}\right)^n. \end{aligned}$$

Man erhält die Taylor-Entwicklung um $z_0 = i$

$$f(z) = \sum_{n=0}^{\infty} \underbrace{\left[\frac{1}{(1-i)^{n+1}} + (-1)^n \frac{1}{(2+i)^{n+1}} \right]}_{=a_n} (z-i)^n.$$

24.2 Laurent-Reihen (holomorphe Funktionen auf Kreisringen; Haupt- und Nebenteil einer Laurent-Reihe; Singularität; gelochte Kreisscheibe; Charakterisierung von isolierten Singularitäten)

Betrachtet man Funktionen wie $f(z) = g(z)z^{-1}$, $g(z)$ holomorph, so ist der Punkt $z = 0$ von besonderer Bedeutung und bietet sich als Entwicklungspunkt einer Reihenentwicklung an. Die Funktion f ist jedoch in diesem Punkt **nicht definiert**, f ist **nicht holomorph auf einer Kreisscheibe** um 0.

Es ist f jedoch **holomorph auf Kreisringen** um den Nullpunkt. Diese Situation soll nun studiert werden: Bei der Entwicklung von f sind dann negative Potenzen mitzunehmen, man spricht von der sogenannten **Laurent¹-Entwicklung** der Funktion. Im obigen Beispiel etwa werden die Exponenten in der Taylor-Entwicklung von g um eins vermindert werden.

Im Folgenden sei für $0 \leq r_1 < r_2$ und für $z_0 \in \mathbb{C}$

$$A_{r_1, r_2}(z_0) := \{z \in \mathbb{C} : r_1 < |z - z_0| < r_2\}.$$

¹P.A. Laurent, 1813–1854; Le Havre.

Definition 24.2.1

Mit $a_k \in \mathbb{C}$, $k \in \mathbb{Z}$ und $z_0 \in \mathbb{C}$ heißt

$$\sum_{k=-\infty}^{\infty} a_k (z - z_0)^k := \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n}$$

eine *Laurent-Reihe* um den *Entwicklungspunkt* z_0 . Es heißt weiter

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

der *Nebenteil* (oder der *Regulärteil*) und

$$\sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n}$$

der *Hauptteil* der *Laurent-Reihe*.

Konvergenz einer Laurent-Reihe?

Bei der Frage nach der möglichen Konvergenz einer Laurent-Reihe ist zunächst zu beachten, dass der Nebenteil eine Potenzreihe im üblichen Sinne ist. Es sei angenommen, dass diese die im Konvergenzreis $B_{r_2}(z_0)$ konvergiere.

Zur Analyse des Hauptteils setzt man $w = (z - z_0)^{-1}$. Mit dieser Substitution kann der Hauptteil geschrieben werden als

$$\sum_{n=1}^{\infty} a_{-n} (z - z_0)^{-n} = \sum_{n=1}^{\infty} a_{-n} w^n .$$

Dies wiederum ist eine Potenzreihe in w , man nehme an, sie konvergiere für $|w| < \rho$, also für

$$|z - z_0| > \frac{1}{\rho} =: r_1 .$$

Ist nun $0 < r_1 < r_2$, so konvergiert die Laurent-Reihe also auf dem Kreisring $A_{r_1, r_2}(z_0)$. Ist $r_1 = r_2$, so kann die Reihe höchstens für $|z| = r_1 = r_2$ konvergieren, ist $r_1 > r_2$, so kann die Reihe nirgends konvergieren.

Satz 24.2.1

Für eine Laurent-Reihe gilt eine der drei folgenden *Alternativen*:

- i) Die Reihe konvergiert *nicht*.
- ii) Die Reihe konvergiert *für gewisse Punkte einer Kreislinie um z_0* .
- iii) Die Reihe konvergiert *auf einem Kreisring um z_0* .

Kann eine auf einem Kreisring holomorphe Funktion durch eine Laurent-Reihe dargestellt werden?

Satz 24.2.2

Es sei f *holomorph auf einem Kreisring* $A_{r_1, r_2}(z_0)$, $0 \leq r_1 < r_2$, $z_0 \in \mathbb{C}$. Dann ist f auf $A_{r_1, r_2}(z_0)$ *eindeutig darstellbar als Laurent-Reihe* (Laurent-Entwicklung um den Entwicklungspunkt z_0)

$$f(z) = \sum_{k=-\infty}^{\infty} a_k (z - z_0)^k ,$$

wobei die Koeffizienten für alle $k \in \mathbb{Z}$ gegeben sind durch ($r_1 < \rho < r_2$)

$$a_k = \frac{1}{2\pi i} \int_{\kappa_\rho(z_0)} \frac{f(z)}{(z - z_0)^{k+1}} dz .$$

Bemerkungen.

- i)* Wieder können die a_k auch durch Integration über einen einfachen, geschlossenen Integrationsweg berechnet werden, der ganz in $A_{r_1, r_2}(z_0)$ liegt und im positiven Sinne durchlaufen wird.
- ii)* Eine Taylor-Reihe ist eine Laurent-Reihe mit $a_{-n} = 0$ für alle $n \in \mathbb{N}$.

Beispiel. Es sei wieder

$$f(z) = \frac{3}{(1-z)(2+z)} = \frac{1}{1-z} + \frac{1}{2+z}.$$

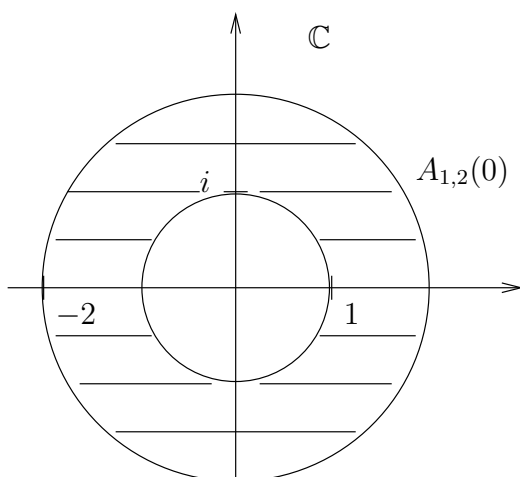


Abbildung 24.2: f ist holomorph auf $A_{1,2}(0)$.

Mit $z_0 = 0$ ist f holomorph auf $A_{1,2}(z_0)$. Wie oben sieht man mit Hilfe der Konvergenz der geometrischen Reihe und wegen $|z^{-1}|, |z/2| < 1$ auf $A_{1,2}(0)$ (man beachte: Auf $A_{1,2}(0)$ ist $|z| \geq 1$ und $1/(1-z)$ kann nicht als konvergente geometrische Reihe geschrieben werden.):

$$\frac{1}{1-z} = -\frac{1}{z} \frac{1}{1-\frac{1}{z}} = -\frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^n,$$

$$\frac{1}{2+z} = \frac{1}{2} \frac{1}{1+\frac{z}{2}} = \frac{1}{2} \sum_{n=0}^{\infty} (-1)^n \left(\frac{z}{2}\right)^n.$$

Die Laurent-Reihe von $f(z)$ um 0 ist demnach für $1 < |z| < 2$:

$$f(z) = - \sum_{n=1}^{\infty} z^{-n} + \sum_{n=0}^{\infty} \frac{(-1)^n}{2^{n+1}} z^n .$$

Mit Hilfe von Satz 24.2.2 können **Singularitäten** einer Funktion charakterisiert werden.

Notation. Zu $z_0 \in \mathbb{C}$, $r > 0$, bezeichne $B'_r(z_0)$ die **gelochte Kreisscheibe** vom Radius r um z_0 :

$$B'_r(z_0) := \{z \in \mathbb{C} : 0 < |z - z_0| < r\} .$$

Definition 24.2.2

*Ist f auf $B'_r(z_0)$ definiert und differenzierbar, so heißt z_0 eine **isolierte Singularität** von f . Dabei muss f nicht in z_0 definiert sein.*

Beispiel. Die Funktion

$$f(z) = \frac{1}{(1-z)(2+z)(z-i)} , \quad z \in \mathbb{C} - \{1, -2, i\} ,$$

hat isolierte Singularitäten in den Punkten $z_1 = 1$, $z_2 = -2$, $z_3 = i$.

Im folgenden werden stets isolierte Singularitäten betrachtet, deren Charakterisierung lautet:

Definition 24.2.3

Es sei f auf $B'_r(z_0)$ differenzierbar.

- i) Der Punkt z_0 heißt **hebbare Singularität**, wenn der Hauptteil der Laurent-Reihe **um z_0** verschwindet.*
- ii) Der Punkt z_0 heißt **Pol der Ordnung p** , $p \in \mathbb{N}$, wenn der Hauptteil der Laurent-Reihe **um z_0** von der Form ist*

$$\sum_{n=1}^p a_{-n} (z - z_0)^{-n} , \quad a_{-p} \neq 0 .$$

*Die Funktion f heißt in diesem Fall **meromorph**.*

iii) Der Punkt z_0 heißt *wesentliche Singularität*, wenn der Hauptteil der Laurent-Reihe *um z_0* unendlich viele nicht-verschwindende Glieder hat.

Beispiele.

i) Es sei

$$f(z) = \frac{\sin(z)}{z}, \quad z_0 = 0.$$

Dann ist f zwar in z_0 nicht definiert, es gilt aber:

$$\frac{\sin(z)}{z} = \frac{1}{z} \left[z - \frac{z^3}{3!} + \frac{z^5}{5!} \dots \right] = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} \dots$$

Der Hauptteil der Laurentreihe um 0 verschwindet, $\sin(z)/z$ lässt sich durch den Wert 1 holomorph in den Nullpunkt fortsetzen (heb-bare Singularität).

ii) Es sei

$$f(z) = \frac{3}{(1-z)(2+z)}, \quad z \in \mathbb{C} - \{1, -2\}.$$

Hier sind $z_1 = 1$ und $z_2 = -2$ Pole der Ordnung 1.

iii) Es sei

$$f(z) = \frac{1}{(z-i)^3}, \quad z \in \mathbb{C} - \{i\}.$$

In diesem Beispiel ist $z_0 = i$ ein Pol der Ordnung 3.

iv) Es sei

$$f(z) = e^{\frac{1}{z}}, \quad z \in \mathbb{C} - \{0\}.$$

Es gilt

$$f(z) = \sum_{n=0}^{\infty} \frac{(z^{-1})^n}{n!} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} z^{-n},$$

also ist $z_0 = 0$ eine wesentliche Singularität.

Kapitel 25

Der Residuensatz

25.1 Der Satz (Residuum)

Der **Residuensatz** ist eine **Verallgemeinerung des Cauchyschen Integralsatzes auf Funktionen mit isolierten Singularitäten**.

Definition 25.1.1

Es sei $z_0 \in \mathbb{C}$, $r > 0$, f differenzierbar auf $B_r'(z_0)$. Für $0 < \rho < r$ heißt

$$\frac{1}{2\pi i} \int_{\kappa_\rho(z_0)} f(z) \, dz =: \operatorname{res}_{z_0} f$$

*das **Residuum** von f in der isolierten Singularität z_0 .*

Bemerkungen.

- i)* Nach dem Cauchyschen Integralsatz (bzw. dessen Verallgemeinerung) hängt $\operatorname{res}_{z_0} f$ nicht von ρ ab. Außerdem kann wieder statt $\kappa_\rho(z_0)$ ein beliebiger, einfacher geschlossener, positiv orientierter Integrationsweg in $B_r'(z_0)$ betrachtet werden.
- ii)* Nach Satz 24.2.2 ist das Residuum von f in z_0 **genau der Koeffizient a_{-1} in der Laurent-Entwicklung von f um z_0 .**
- iii)* Bemerkung *ii)* kann wie folgt veranschaulicht werden: Schreibt man f lokal als Laurent-Reihe und vertauscht man Integration

und Summation, so verschwinden nach Satz 22.3.1 in der Summe alle Integrale über $a_k(z - z_0)^k$, $k \neq -1$. Lediglich das Integral mit $k = -1$ gibt einen Beitrag.

Beispiele.

i) Es sei

$$f(z) = \frac{3}{z - i}, \quad z \in \mathbb{C} - \{i\}.$$

Nach Satz 22.3.1 ist für $\rho > 0$

$$\int_{\kappa_\rho(i)} \frac{1}{z - i} dz = 2\pi i,$$

also folgt

$$\operatorname{res}_i f(z) = \frac{1}{2\pi i} [3 \cdot 2\pi i] = 3.$$

Zu beachten ist hier: $3/(z - i)$ ist bereits die Laurent-Reihe von f um den Punkt $z_0 = i$, also $a_{-1} = 3$.

ii) Es sei

$$f(z) = e^{\frac{1}{z}}, \quad z \in \mathbb{C} - \{0\}.$$

Es wurde bereits gezeigt:

$$f(z) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} z^{-n},$$

also ist $a_{-1} = 1$ und $\operatorname{res}_0 f = 1$.

Die Integraldarstellung des Residuums sowie die Entwicklung von f in eine Laurent-Reihe können in der Praxis allerdings Schwierigkeiten bereiten, d.h. es stellt sich die Frage nach:

Hilfsmittel zur Berechnung von Residuen?

Satz 25.1.1

i) Es sei $f(z) = \frac{g(z)}{h(z)}$, wobei g, h holomorph seien auf $B_r(z_0)$. Weiterhin sei $g(z_0) \neq 0$, $h(z_0) = 0$ und $h'(z_0) \neq 0$. Dann ist das Residuum von f in z_0

$$\operatorname{res}_{z_0} f = \operatorname{res}_{z_0} \frac{g}{h} = \frac{g(z_0)}{h'(z_0)}.$$

ii) Die meromorphe Funktion f habe in $z_0 \in \mathbb{C}$ einen Pol der Ordnung p . Dann gilt

$$\operatorname{res}_{z_0} f = \lim_{z \rightarrow z_0} \frac{1}{(p-1)!} \left[\frac{d^{p-1}}{dz^{p-1}} \left[(z-z_0)^p f(z) \right] \right].$$

Beispiele.

i) Es sei

$$f(z) = \frac{z^2 + 4}{\sin(z)}, \quad 0 < |z| < \frac{\pi}{2}, \quad z_0 = 0.$$

Satz 25.1.1, i), ist anwendbar mit $g(z) = z^2 + 4$, $h(z) = \sin(z)$.

$$\operatorname{res}_0 \frac{z^2 + 4}{\sin(z)} = \frac{4}{\cos(0)} = 4.$$

ii) Es sei

$$f(z) = \frac{e^{iz}}{z^3}, \quad z \in \mathbb{C} - \{0\}, \quad z_0 = 0.$$

Hier ist z_0 ein Pol der Ordnung 3 und es gilt nach Satz 25.1.1, ii),

$$\operatorname{res}_0 \frac{e^{iz}}{z^3} = \lim_{z \rightarrow z_0} \frac{1}{2!} \left[\frac{d^2}{dz^2} e^{iz} \right] = -\frac{1}{2}.$$

Der Hauptsatz dieses Kapitels lautet nun:

Satz 25.1.2 (*Der Residuensatz*)

Es sei γ ein *einfacher geschlossener* Integrationsweg in einem *einfach zusammenhängenden* Gebiet G , der im *mathematisch positiven Sinne* durchlaufen werde. Die Funktion f sei definiert und differenzierbar in G mit Ausnahme von *endlich vielen isolierten Singularitäten*. Die Spur von γ treffe keine dieser Singularitäten. Sind z_1, \dots, z_N die *Singularitäten*, die *von γ eingeschlossen* werden, so gilt

$$\int_{\gamma} f(z) \, dz = 2\pi i \sum_{k=1}^N \operatorname{res}_{z_k} f .$$

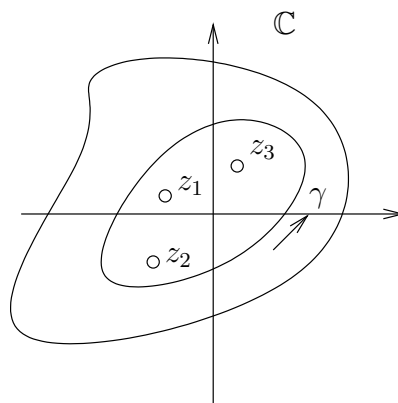


Abbildung 25.1: Zum Residuensatz.

Bemerkungen.

- i) Allein die Kenntnis der Residuen ermöglicht also die Berechnung des Kurvenintegrals.
- ii) Ist f holomorph auf G (keine Singularitäten) so erkennt man den Cauchyschen Integralsatz als Spezialfall.

Beispiele.

i) Betrachtet sei das Standardbeispiel $f(z) = z^{-1}$. Das Residuum $\text{res}_0 f$ von f in $z_0 = 0$ ist der Koeffizient $a_{-1} = 1$ der Laurent-Reihe von f um $z_0 = 0$. Es folgt (vgl. Satz 22.3.1)

$$\int_{\kappa_\rho(0)} \frac{1}{z} dz = 2\pi i \text{res}_0 f = 2\pi i .$$

Es sei $f(z) = z^{-2}$. Hier gilt $\text{res}_0 f = 0$ (wegen $a_{-1} = 0$) und (vgl. wieder Satz 22.3.1)

$$\int_{\kappa_\rho(0)} \frac{1}{z^2} dz = 0 .$$

ii) Es sei $G = \mathbb{C}$ und

$$f(z) = \frac{e^z}{1+z^2}, \quad z \neq \pm i .$$

Nach 25.1.1 gilt

$$\begin{aligned} \text{res}_i \frac{e^z}{1+z^2} &= \frac{e^z}{2z|_{z=i}} = \frac{e^i}{2i}, \\ \text{res}_{-i} \frac{e^z}{1+z^2} &= -\frac{e^i}{2i}. \end{aligned}$$

Es folgt

$$\int_{\kappa_3(i)} \frac{e^z}{1+z^2} dz = 2\pi i \underbrace{\left[\frac{e^i}{2i} - \frac{e^{-i}}{2i} \right]}_{\frac{2i \sin(1)}{2i}} = 2\pi i \sin(1) .$$

Der Residuensatz hilft auch bei der Berechnung reeller Integrale, beispielsweise gilt

Satz 25.1.3

Es sei $f(z) = \frac{P(z)}{Q(z)}$ eine rationale Funktion mit Polyno-

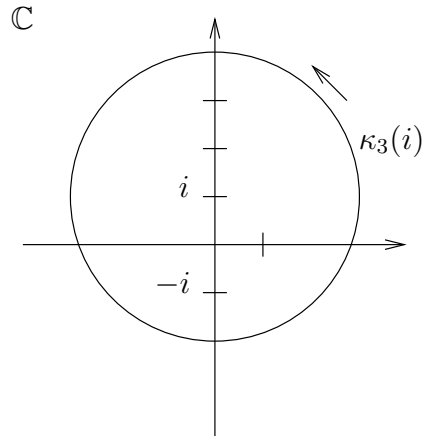


Abbildung 25.2: Ein Beispiel zum Residuensatz.

men $P(z)$, $Q(z)$. Der Grad von Q sei um *mindestens 2 höher* als der von P , Q habe keine *reellen* Nullstellen. Sind dann z_1, \dots, z_N die Nullstellen von Q in der oberen Halbebene $\text{Im } z > 0$, so gilt

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx = 2\pi i \sum_{k=1}^N \text{res}_{z_k} \frac{P(z)}{Q(z)} .$$

Bemerkung. Ist der Grad von Q nur um eins höher als der Grad von P , so divergiert das Integral auf der linken Seite.

Beispiel. Gesucht sei das reelle Integral

$$\int_{-\infty}^{\infty} \frac{1}{(1+x^2)^2} dx .$$

Mit $P(z) = 1$, $Q(z) = (1+z^2)^2$ sind die Voraussetzungen des Satzes erfüllt. In der oberen Halbebene liegt die Nullstelle $z = i$ mit der

Vielfachheit 2. Nach Satz 25.1.1 folgt

$$\begin{aligned}\operatorname{res}_i \frac{1}{(1+z^2)^2} &= \lim_{z \rightarrow i} \left[\frac{d}{dz} \left[(z-i)^2 \frac{1}{(1+z^2)^2} \right] \right] \\ &= \lim_{z \rightarrow i} \left[\frac{d}{dz} \left[(z-i)^2 \frac{1}{(z-i)^2(z+i)^2} \right] \right] \\ &= \lim_{z \rightarrow i} \left[\frac{d}{dz} \frac{1}{(z+i)^2} \right] \\ &= \lim_{z \rightarrow i} \left[-\frac{2}{(z+i)^3} \right] = \frac{1}{4i} = -\frac{i}{4}.\end{aligned}$$

Es gilt also

$$\int_{-\infty}^{\infty} \frac{1}{(1+x^2)^2} dx = 2\pi i \left(-\frac{i}{4} \right) = \frac{\pi}{2}.$$

Teil X

Integraltransformationen

Kapitel 26

Fourier-Reihen

26.1 Einführung (Spektrum; harmonische Analyse; Periode einer Funktion; trigonometrische Reihen; trigonometrische Polynome; gliedweise Integration; Integration und Grenzübergang; Fourier-Koeffizienten)

Bevor in den nachfolgenden Kapiteln der kontinuierliche Fall der [Fourier¹-Transformation](#) bzw. der [Laplace-Transformation](#) diskutiert wird, soll hier zunächst das [diskrete Spektrum periodischer Funktionen](#) studiert werden.

Eine ausführliche Schilderung der Bedeutung von Fourier-Reihen in der Entwicklung der Mathematik findet sich in [Hi1].

Frage. Das Beispiel einer schwingenden Saite motiviert die Frage: Lässt sich eine periodische Funktion als Überlagerung oder [Superposition](#) von [Grund-](#) und [Oberschwingungen](#) darstellen? ([harmonische Analyse](#))

Es geht also um das Studium [periodischer Vorgänge](#) wie die Erzeugung von Wechselstrom, mechanische Schwingungen, Analyse von Messkurven etc.

Dabei heißt eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ [periodisch mit der Periode \$T \neq 0\$](#) , falls für alle $x \in \mathbb{R}$ gilt: $f(x + T) = f(x)$.

¹J.B.J. Baron de Fourier, 1768–1830; Grenoble, Paris.

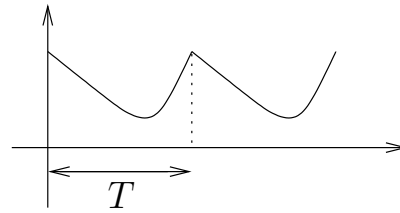


Abbildung 26.1: Eine periodische Funktion.

Beispiele.

- i) Konstante Funktionen, trigonometrische Funktionen,
- ii) Für beliebige Amplituden $A_1, A_2 \in \mathbb{R}$ ist die Funktion

$$A_1 \sin(\omega_1 x) + A_2 \sin(\omega_2 x)$$

periodisch, falls das Verhältnis der Kreisfrequenzen (und damit der Perioden) rational ist (\rightsquigarrow Übungen).

Im Folgenden werden zunächst **o.E. 2π -periodische Funktionen** betrachtet.

Fourier-Reihen sind nun **trigonometrische Reihen** der Form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)), \quad a_n, b_n \in \mathbb{R},$$

die Partialsummen

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx))$$

heißen **trigonometrische Polynome**.

Bemerkung. Die Funktion $A_1 \sin(x) + A_2 \sin(\omega_2 x)$ hat Periode 2π , falls (wie in obiger Übung zu zeigen) $\omega_2 \in \mathbb{N}$, d.h. trigonometrische Reihen mit obigen **Frequenzen** sind die Grundlage der harmonischen Analyse.

Fragen. Präzisiert lauten die entscheidenden Frage dieses Kapitels:

- i) Lässt sich eine 2π -periodische Funktion in eine trigonometrische Reihe entwickeln?
- ii) Wenn ja, in welchem Sinne konvergiert die Reihe?

Zur Vorbereitung benötigt man das folgende Lemma, welches als Übung nachzurechnen ist:

Lemma 26.1.1 (*Eigenschaften trigonometrischer Funktionen*)

i) Für $m, n \in \mathbb{N} \cup \{0\} =: \mathbb{N}_0$ gelten die *Orthogonalitätsbeziehungen*

$$\int_0^{2\pi} \cos(mx) \cos(nx) \, dx = 0 \quad \text{für } m \neq n ;$$

$$\int_0^{2\pi} \sin(mx) \sin(nx) \, dx = 0 \quad \text{für } m \neq n ;$$

$$\int_0^{2\pi} \cos(mx) \sin(nx) \, dx = 0 .$$

ii) Für $m, n \in \mathbb{N}_0$ gelten die *Normierungsbeziehungen*

$$\int_0^{2\pi} \cos^2(nx) \, dx = \begin{cases} \pi & \text{für } n \in \mathbb{N} , \\ 2\pi & \text{für } n = 0 ; \end{cases}$$

$$\int_0^{2\pi} \sin^2(nx) \, dx = \begin{cases} \pi & \text{für } n \in \mathbb{N} , \\ 0 & \text{für } n = 0 . \end{cases}$$

Zunächst wird *einfach angenommen*, dass eine formal gebildete *Fourier-Reihe gleichmäßig gegen eine Funktion f konvergiert*. Unter dieser Annahme folgt:

Satz 26.1.1

Konvergiert die Fourier-Reihe (die Folge der Partialsummen)

$$\frac{a_0}{2} + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx))$$

*im Intervall $[0, 2\pi]$ **gleichmäßig** gegen eine Funktion $f(x)$, so ist f stetig und es gilt*

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, \dots$$

Bemerkung. Aus der **angenommenen** gleichmäßigen Konvergenz der Reihe gegen eine Funktion f folgen Darstellungsformeln für die Koeffizienten a_n und b_n .

Zum Beweis von Satz 26.1.1 benötigt man den folgenden Satz, der ohne Beweis angegeben sei:

Satz 26.1.2 (*gliedweise Integration, Integration & Grenzübergang*)

Betrachtet sei auf einem Intervall I eine Folge von Funktionen $g_1(x), g_2(x), \dots$. Für alle $x \in I$ existiere der Grenzwert

$$g(x) := \lim_{N \rightarrow \infty} g_N(x).$$

*Sind die g_N auf dem Intervall $I = [a, b]$ stetig und konvergieren sie **gleichmäßig** gegen $g(x)$, so gilt*

$$\int_a^b g(x) \, dx = \int_a^b \lim_{N \rightarrow \infty} g_N(x) \, dx = \lim_{N \rightarrow \infty} \int_a^b g_N(x) \, dx.$$

Im Fall $g_N(x) = \sum_{n=1}^N h_n(x)$ folgt (mit obiger Notation)

$$\int_a^b \sum_{n=1}^{\infty} h_n(x) \, dx = \sum_{n=1}^{\infty} \int_a^b h_n(x) \, dx .$$

Beweis von Satz 26.1.1 Die Stetigkeit von

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

folgt aus der **angenommenen** gleichmäßigen Konvergenz. Betrachtet seien nun die Partialsummen

$$g_N(x) := \frac{a_0}{2} + \sum_{n=1}^N (a_n \cos(nx) + b_n \sin(nx)) , \quad N \in \mathbb{N} .$$

Für ein fixiertes $k \in \mathbb{N}_0$ gilt nach Annahme

$$g_N(x) \cos(kx) \xrightarrow{N \rightarrow \infty} f(x) \cos(kx)$$

($\sin(kx)$ analog). Also ist nach Satz 26.1.2 gliedweise Integration zulässig. Lemma 26.1.1 impliziert

$$\begin{aligned} & \int_0^{2\pi} f(x) \cos(kx) \, dx \\ &= \frac{a_0}{2} \int_0^{2\pi} \cos(kx) \, dx + \sum_{n=1}^{\infty} \int_0^{2\pi} (a_n \cos(nx) + b_n \sin(nx)) \cos(kx) \, dx \\ &= a_k \pi \end{aligned}$$

und der Satz ist bewiesen. □

Bemerkung. Der Satz beantwortet **nicht** die Frage nach der Darstellbarkeit einer gegebenen Funktion f als Fourier-Reihe (gleichmäßige Konvergenz war vorausgesetzt), er motiviert aber die nachfolgende Definition.

Definition 26.1.1

Die reellwertige Funktion f sei auf $[0, 2\pi]$ definiert und integrierbar.

i) Die Zahlen

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) \, dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx, \quad n = 1, 2, \dots$$

heißen *Fourier-Koeffizienten* von f .

ii) Die mit diesen Koeffizienten *formal* gebildete Reihe

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

heißt *Fourier-Reihe* von f .

Beispiele.

i) Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit der Periode 2π ,

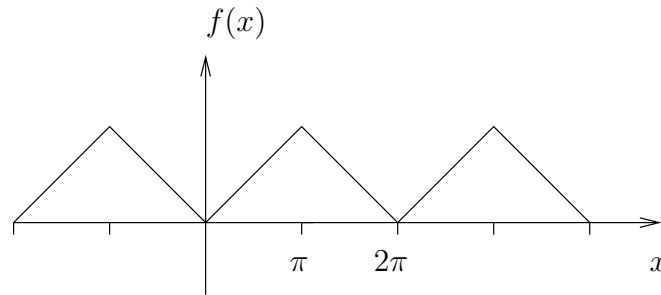
$$f(x) = \begin{cases} x, & 0 \leq x \leq \pi, \\ 2\pi - x, & \pi \leq x \leq 2\pi. \end{cases}$$

Die Fourier-Koeffizienten lauten

$$a_0 = \pi, \quad a_n = \begin{cases} 0 & \text{für } n = 2m, \\ -\frac{4}{\pi n^2} & \text{für } n = 2m + 1, \end{cases} \quad n \in \mathbb{N},$$

$$b_n = 0, \quad n \in \mathbb{N}.$$

Die Fourier-Reihe ist gleichmäßig konvergent (\rightsquigarrow Übungen).

Abbildung 26.2: Die Funktion $f(x)$.

ii) Ist f **gerade**, also $f(-x) = f(x)$ für alle $x \in \mathbb{R}$, so folgt (\rightsquigarrow Übung)

$$b_n = 0 \quad \text{für alle } n \in \mathbb{N},$$

$$a_n = \frac{2}{\pi} \int_0^\pi f(x) \cos(nx) \, dx \quad \text{für alle } n \in \mathbb{N}_0.$$

Ist f **ungerade**, also $f(-x) = -f(x)$ für alle $x \in \mathbb{R}$, so folgt

$$a_n = 0 \quad \text{für alle } n \in \mathbb{N}_0,$$

$$b_n = \frac{2}{\pi} \int_0^\pi f(x) \sin(nx) \, dx \quad \text{für alle } n \in \mathbb{N}.$$

26.2 Der Satz (Sägezahnfunktion; Gibbs-Phänomen; stückweise glatte Funktion; mittlere quadratische Abweichung; Parsevalsche Gleichung)

Musterbeispiel dieses Kapitels. Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit Periode 2π und

$$f(x) = \pi - x \quad \text{für } 0 \leq x < 2\pi \quad (\text{Sägezahnfunktion}).$$

In diesem Fall ist f ungerade, also $a_n = 0$ für alle $n \in \mathbb{N}_0$ (wie f in den Sprungstellen definiert ist, spielt für dieses Argument keine Rolle).

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) \, dx = \frac{1}{\pi} \int_0^{2\pi} (\pi - x) \sin(nx) \, dx \\ &= -\frac{1}{\pi} \int_0^{2\pi} x \sin(nx) \, dx \\ &= -\frac{1}{\pi} \left[x \left(-\frac{1}{n} \cos(nx) \right) \Big|_0^{2\pi} + \frac{1}{n} \int_0^{2\pi} \cos(nx) \, dx \right] = \frac{2}{n}. \end{aligned}$$

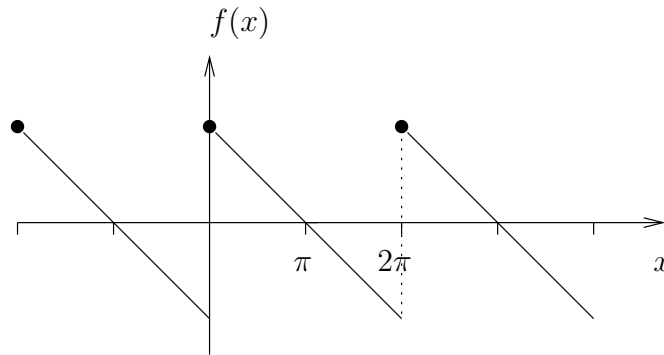


Abbildung 26.3: Die Sägezahnfunktion.

Die (formale) Fourier-Reihe lautet:

$$2\left\{\sin(x) + \frac{1}{2}\sin(2x) + \frac{1}{3}\sin(3x) + \dots\right\} = 2\sum_{n=1}^{\infty} \frac{\sin(nx)}{n}.$$

Zur Konvergenz der Reihe. Betrachtet seien dazu die Partialsummen

$$S_N(x) := 2\sum_{n=1}^N \frac{\sin(nx)}{n}$$

sowie die Differenz

$$R_N(x) := S_N(x) - f(x).$$

Man kann zeigen: Für $x \in (0, 2\pi)$ gilt

$$|R_N(x)| \leq \frac{4}{(2N+1)\sin(x/2)},$$

d.h.: Gleichmäßige Konvergenz gegen $f(x)$ gilt nur auf Teilintervallen der Form $[\varepsilon, 2\pi - \varepsilon]$ (weg von der Sprungstelle).

Für $x = 0, x = 2\pi, \dots$ ist $S_N = 0$ ("Mittelwert des Sprungs"), es liegt **keine Konvergenz der Reihe gegen die Funktion in den Sprungstellen** vor.

Man kann weiter zeigen: Für $x_N = \pi/(N + 1/2)$ und für alle N hinreichend groß gilt

$$R_N(x_N) \gtrsim 0.1789 \cdot \pi,$$

d.h. auch wenn N sehr groß wird **überschwingen** die Partialsummen den Wert π um ca. 18 % (**Gibbs²-Phänomen**).

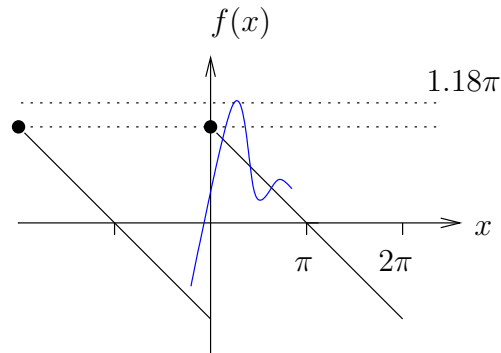


Abbildung 26.4: Das Gibbs-Phänomen.

Bevor nun das Hauptergebnis dieses Kapitels formuliert werden kann, ist eine geeignete Funktionenklasse einzuführen.

Definition 26.2.1

i) Eine **Sprungstelle** einer Funktion f ist eine Unstetigkeitsstelle x_0 von f , für die rechtsseitiger und linksseitiger Grenzwert

$$f(x_0^+) = \lim_{x \downarrow x_0} f(x), \quad f(x_0^-) = \lim_{x \uparrow x_0} f(x)$$

existieren. Die **Sprunghöhe** ist dabei $h = f(x_0^+) - f(x_0^-)$.

ii) Eine Funktion f heißt in $[a, b]$ **stückweise stetig**, wenn es eine Zerlegung $\mathcal{Z} = \{a = x_0 < x_1 < \dots < x_n = b\}$ gibt, sodass f in jedem Teilintervall (x_{j-1}, x_j) stetig ist und die einseitigen Grenzwerte $f(a^+)$, $f(b^-)$, $f(x_j^+)$, $f(x_j^-)$, $1 \leq j \leq n-1$ existieren.

²J.W. Gibbs, 1839–1903; New Haven.

iii) Eine Funktion f heißt in $[a, b]$ **stückweise glatt**, wenn f und f' in $[a, b]$ stückweise stetig sind.

Bemerkungen.

- i) Eine stückweise glatte Funktion hat höchstens endlich viele Sprungstellen und endlich viele “Ecken”.
- ii) Die Bedeutung stückweise glatter Funktionen in den Anwendungen wird deutlich, wenn man etwa an Einschaltvorgänge denkt.

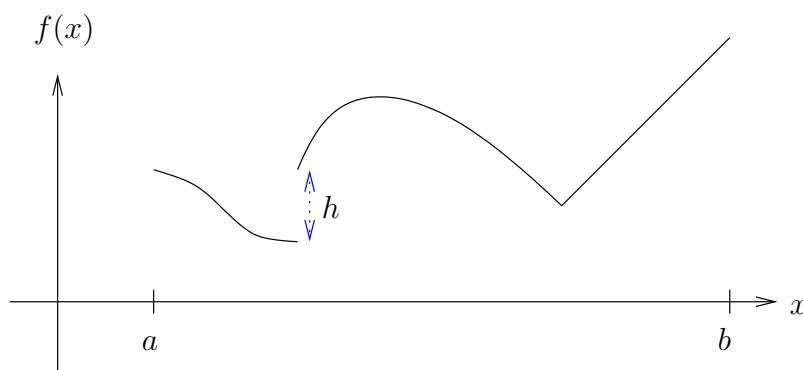


Abbildung 26.5: Eine stückweise glatte Funktion.

Satz 26.2.1

Die 2π -periodische Funktion f sei in $[0, 2\pi]$ stückweise glatt. Dann konvergiert die Fourier-Reihe von f für alle $x \in \mathbb{R}$. Ist f **stetig an der Stelle x** , so gilt

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) .$$

Ist x eine Sprungstelle von f , so gilt

$$\frac{f(x^+) + f(x^-)}{2} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) .$$

In jedem abgeschlossenen Intervall, in dem f stetig ist, ist die Reihe gleichmäßig konvergent.

Beispiel. Im obigen Beispiel *i*) ist f stückweise glatt und stetig, also

$$f(x) = \left\{ \begin{array}{ll} x, & 0 \leq x < \pi \\ 2\pi - x, & \pi \leq x < 2\pi \end{array} \right\} = \frac{\pi}{2} - \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{\cos((2m+1)x)}{(2m+1)^2} .$$

Insbesondere folgt mit der Wahl $x = 0$

$$\sum_{m=0}^{\infty} \frac{1}{(2m+1)^2} = \frac{\pi^2}{8} .$$

Bemerkungen.

- i*) Man überlege sich, warum hier viel schwächere Bedingungen an f zu stellen sind als es für die Entwicklung in Taylor-Reihen notwendig war.
- ii*) Die Diskussion von periodischen Funktionen f mit Periode $T > 0$, $T \neq 2\pi$, erfolgt analog (mittels der Transformation $u = 2\pi x/T$). Die Fourier-Reihe ist dann

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi}{T}nx\right) + b_n \sin\left(\frac{2\pi}{T}nx\right) \right)$$

mit den Fourier-Koeffizienten

$$a_n = \frac{2}{T} \int_0^T f(x) \cos\left(\frac{2\pi}{T}nx\right) dx, \quad n = 0, 1, 2, \dots,$$

$$b_n = \frac{2}{T} \int_0^T f(x) \sin\left(\frac{2\pi}{T}nx\right) dx, \quad n = 1, 2, \dots$$

iii) Statt des Intervalls $[0, T]$ können analog Intervalle der Form $[a, a+T]$, $a \in \mathbb{R}$, betrachtet werden.

iv) Approximiert man eine 2π -periodische Funktion mit einem trigonometrischen Polynom

$$P_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (\alpha_k \cos(kx) + \beta_k \sin(kx))$$

mit Koeffizienten derart, dass

$$Q_n := \int_0^{2\pi} [f(x) - P_n(x)]^2 dx$$

(mittlere quadratische Abweichung) minimal wird (Approximation im quadratischen Mittel), so erhält man genau die Fourier-Koeffizienten.

v) Gilt $\lim_{n \rightarrow \infty} Q_n = 0$ (konvergiert also die Fourier-Reihe im quadratischen Mittel gegen f), so gilt die Parsevalsche Gleichung

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_0^{2\pi} f^2(x) dx$$

(zu beachten: Konvergenz im quadratischen Mittel impliziert nicht punktweise Konvergenz).

Komplexe Notation. Man setzt im Fall einer 2π -periodischen Funktion

$$c_n := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx .$$

Dabei beobachtet man zunächst

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) \cos(nx) dx - \frac{i}{2\pi} \int_0^{2\pi} f(x) \sin(nx) dx ,$$

d.h. es gilt

$$c_n + c_{-n} = a_n \quad \text{und} \quad c_n - c_{-n} = -ib_n .$$

Man beobachtet weiter:

$$\begin{aligned}
 \sum_{n=-\infty}^{\infty} c_n e^{inx} &= \sum_{n=0}^{\infty} c_n (\cos(nx) + i \sin(nx)) \\
 &\quad + \sum_{n=1}^{\infty} c_{-n} (\cos(nx) - i \sin(nx)) \\
 &= c_0 \cos(x) + \sum_{n=1}^{\infty} (c_n + c_{-n}) \cos(nx) \\
 &\quad + \sum_{n=1}^{\infty} i(c_n - c_{-n}) \sin(nx) \\
 &= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) ,
 \end{aligned}$$

die Fourier-Reihe ist in komplexer Schreibweise dargestellt.

Hat f die Periode $T > 0$, so setzt man analog (in diesem Beispiel symmetrisch zum Ursprung)

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-in\frac{2\pi}{T}x} dx ,$$

und die Fourier-Reihe wird geschrieben als

$$\sum_{n=-\infty}^{\infty} c_n e^{in\frac{2\pi}{T}x} .$$

Bemerkung. Zum Abschluss dieses Kapitels sei nochmals festgehalten: Nach Satz 26.2.1 kann eine periodische Funktion **spektral zerlegt** werden, die Gesamtheit der harmonischen Schwingungen, die zur Superposition beitragen, heißt das Spektrum von f . Im Fall einer 2π -periodischen Funktion ist die Frequenz der Grundschwingung 1, die Frequenzen der Oberschwingungen sind die natürliche Zahlen größer als 1. Im Fall einer T -periodischen Funktion ist die Grundfrequenz $\omega = \frac{2\pi}{T}$, die Frequenzen der Oberschwingungen sind Vielfache davon. Die Spektralzerlegung kann etwa mit Hilfe eines **Amplitudenspektrums** visualisiert

werden. Dabei wird $|c_n|$ gegen die Frequenz aufgetragen (vgl. Abbildung 27.3 und Abbildung 27.5 des nächsten Kapitels: Wegen $c_{-n} = \overline{c_n}$ wird dabei nur die positive Achse berücksichtigt, das konstante Glied ist ebenfalls weggelassen.).

Kapitel 27

Fourier-Transformation

27.1 Einführung (periodische Fortsetzung; kontinuierliches Spektrum)

Die Integraltransformation im Mittelpunkt der hier ausgeführten Betrachtungen ist die Fourier-Transformationen. Es sei vorweggeschickt, dass mittels Integraltransformationen ein Problem derart transformiert werden soll, dass es in der neuen Form gelöst werden kann. Der wesentliche Punkt ist, dass anschließend eine **Rücktransformation** möglich sein muss, die auf die Lösung des ursprünglichen Problems führt. M.a.W. muss eine Integraltransformation **invertierbar** sein.

Für die folgenden Betrachtungen gibt es eine Vielzahl von Anwendungen etwa in der mathematischen Physik. Man denke vor Allem auch an die Informationstechnologie (Signale senden, übertragen, empfangen, analysieren, verarbeiten ...).

Idee. Man betrachte als Standardbeispiel dieses Kapitels die **nicht periodische** Funktion $f: [-\pi, \pi] \rightarrow \mathbb{R}$,

$$f(x) = \begin{cases} 1, & \text{falls } -\pi \leq x \leq \pi, \\ 0, & \text{falls } |x| > \pi. \end{cases}$$

Studiert wird also ein **typisches Ein- und Ausschaltsignal**.

Da die Funktion nicht-periodisch ist, kann die Theorie aus Kapitel 26 nicht angewandt werden. Deshalb betrachtet man zunächst die **Einschränkung** $f|_{[-2\pi, 2\pi]}$ der Funktion f auf das Intervall $[-2\pi, 2\pi]$, d.h. $f|_{[-2\pi, 2\pi]}$ ist lediglich definiert auf dem Intervall $[-2\pi, 2\pi]$ und es

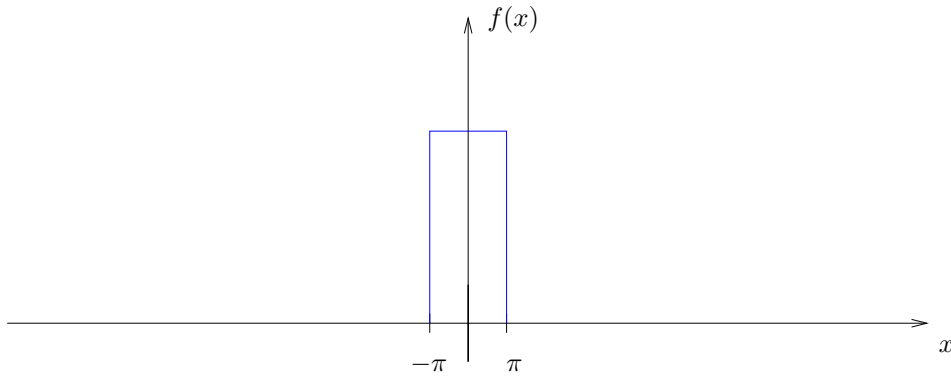
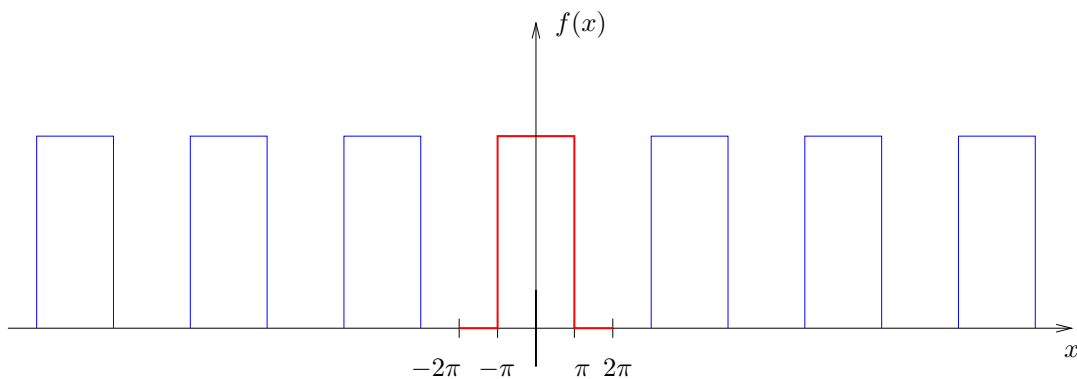


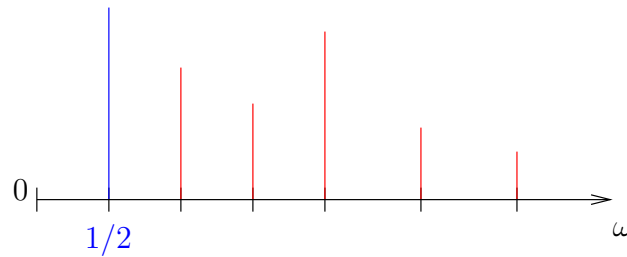
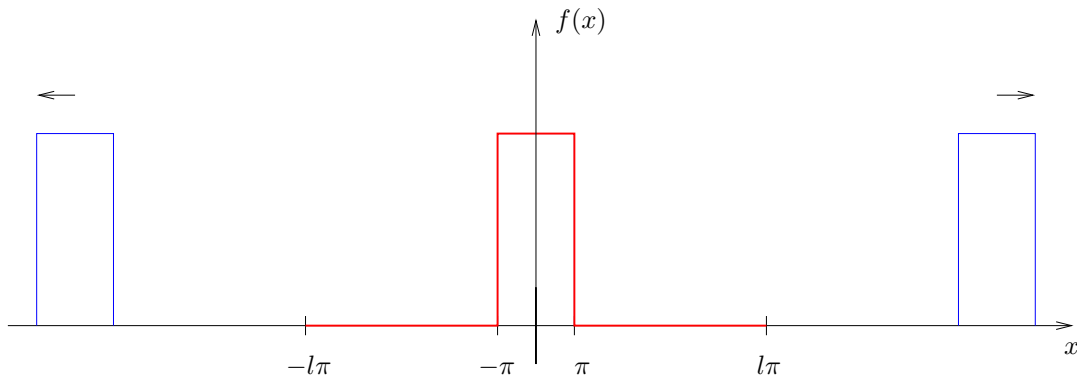
Abbildung 27.1: Das Signal.

gilt für alle x mit $|x| \leq 2\pi$: $f_{|[-2\pi, 2\pi]}(x) = f(x)$. Diese Funktion wird fortgesetzt als Funktion mit der Periode 4π , so wie es in Abbildung 27.2 dargestellt ist. Es ist eine periodische Funktion mit der Periode 4π entstanden, deren Spektrum nach Kapitel 26 analysiert werden kann. Es handelt sich dabei um ein diskretes Spektrum, die Grundschwingung hat die Frequenz $1/2$, die Oberschwingungen sind Vielfache davon (vgl. Abbildung 27.3).

Abbildung 27.2: Die 4π -periodische Fortsetzung.

In der obigen Konstruktion (Einschränkung und periodische Fortsetzung) ersetze man nun 2π durch $l\pi$, $l \in \mathbb{N}$. Die Situation ist in Abbildung 27.4 angedeutet. Die Grundschwingung hat nun die Frequenz $1/l$, die Oberschwingungen sind wiederum Vielfache davon. Man erkennt in Abbildung 27.5, dass die Eintragungen des Spektrums für große l immer näher zusammenrücken.

Abbildung 27.4 suggeriert nun einerseits, dass sich beim Grenzübergang $l \rightarrow \infty$ das Spektrum des nicht-periodischen Ein- und Ausschalt-

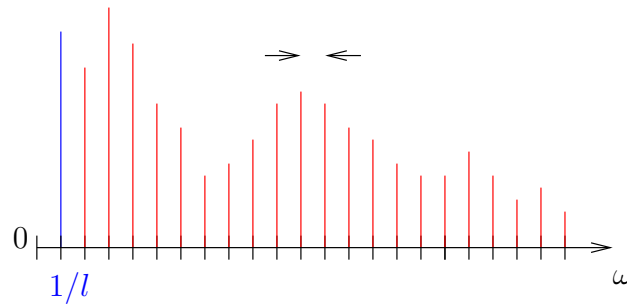

 Abbildung 27.3: Das Spektrum der 4π -periodischen Fortsetzung.

 Abbildung 27.4: Die $2l\pi$ -periodische Fortsetzung.

signals ergeben sollte.

Abbildung 27.5 suggeriert andererseits, dass beim Grenzübergang aus dem diskreten ein [kontinuierliches Spektrum](#) entstehen sollte.

Die nachfolgende “präzisierte Heuristik” ist aus zwei Gründen recht ausführlich gehalten:

- i) Sie bietet die schöne Möglichkeit, eine Vielzahl von Grundbegriffen der Analysis (Konvergenz, Konvergenz von Reihen, uneigentliche Integrale, Riemannsche Zwischensummen, Vertauschung von Grenzwerten, absolute Konvergenz, harmonische Reihe, Konvergenzkriterium von Leibniz) anhand eines wesentlichen Anwendungsbeispiels ins Gedächtnis zu rufen bzw. zu zeigen, wie diese ineinander greifen.
- ii) In der Literatur werden zwei unterschiedliche Zugänge gewählt: Entweder wird ein rigoroser mathematischer Beweis präsentiert, was hier nicht geschehen soll. Oder aber es wird versucht, die

Abbildung 27.5: Das Spektrum der $l\pi$ -periodischen Fortsetzung.

Fourier-Transformation mit Grenzwertbetrachtungen aus dem periodischen Fall herzuleiten, wie es bereits angedeutet wurde. Dabei werden leider oft in unzulässiger Weise Grenzwerte vertauscht und die entscheidende Idee geht aufgrund einer unsaubereren Notation verloren. In der Tat konnte für die folgenden Rechnungen kein Zitat in der gängigen Literatur gefunden werden.

“Präzisierte Heuristik”. Man betrachte den Ausdruck

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du ,$$

wobei im Folgenden die Existenz aller auftretenden uneigentlichen Integrale angenommen sei. Es sei nun $\varepsilon > 0$ fixiert. Für $K > 0$ hinreichend groß gilt

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du \\ &= \frac{1}{\sqrt{2\pi}} \int_{-K}^K e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du + \psi_1(K) . \end{aligned}$$

Hier ist $\psi_1 : \mathbb{R}_+ \rightarrow \mathbb{R}$ eine Funktion mit $\psi_1(K) \rightarrow 0$, falls $K \rightarrow \infty$.

Zu $l \in \mathbb{N}$ wird im nächsten Schritt das Intervall $[-K, K]$ äquidistant mit der Schrittweite $1/l$ zerlegt. Ist l hinreichend groß, so kann das

Integral über u mit einer Riemannschen Summe approximiert werden:

$$\begin{aligned}
 & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du \\
 &= \frac{1}{\sqrt{2\pi}} \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\frac{k}{l}t} dt \frac{1}{l} + \psi_2(l) + \psi_1(K) \\
 &= \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-l\pi}^{l\pi} f(t) e^{-i\frac{k}{l}t} dt \\
 &\quad + \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{l\pi}^{\infty} f(t) e^{-i\frac{k}{l}t} dt \\
 &\quad + \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-\infty}^{-l\pi} f(t) e^{-i\frac{k}{l}t} dt + \psi_2(l) + \psi_1(K) .
 \end{aligned}$$

Wie ψ_1 ist auch $\psi_2: \mathbb{R}_+ \rightarrow \mathbb{R}$ eine Funktion mit $\psi_2(l) \rightarrow 0$, falls $l \rightarrow \infty$.

Fällt $f(x)$ hinreichend schnell gegen Null, gilt also beispielsweise für ein $\alpha > 0$

$$\int_{l\pi}^{\infty} |f(t)| dt \leq l^{-\alpha} \quad \text{und} \quad \int_{-l\pi}^{\infty} |f(t)| dt \leq l^{-\alpha}$$

für l hinreichend groß, so folgt

$$\left| \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{l\pi}^{\infty} f(t) e^{-i\frac{k}{l}t} dt \right| \leq 2Kl \frac{1}{2\pi l} l^{-\alpha} \leq \frac{K}{\pi} l^{-\alpha} .$$

Eine analoge Abschätzung gilt für das Integral von $-\infty$ bis $-l\pi$. Insgesamt ist bisher gezeigt:

$$\begin{aligned}
 & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du \tag{*} \\
 &= \sum_{k=-Kl}^{Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-l\pi}^{l\pi} f(t) e^{-i\frac{k}{l}t} dt + \psi_1(K) + \psi_2(l) + \psi_3(K, l) ,
 \end{aligned}$$

wobei die Funktion $\psi_3: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$ abgeschätzt ist durch

$$|\psi_3(K, l)| \leq \frac{2K}{\pi} l^{-\alpha}.$$

Bisher wurde mit der Funktion f selbst gerechnet. Jetzt wird beobachtet: In der Summe über k auf der rechten Seite von (*) steht das **diskrete Spektrum der periodischen Fortsetzung von $f_{|[-l\pi, l\pi]}$ im Frequenzbereich $[-K, K]$** .

Wäre diese Summe eine **unendliche Summe**, so **wäre** diese für $|x| < l\pi$ nach Satz 26.2.1 gleich $f_{|[l\pi, l\pi]}(x) = f(x)$ und es **würde** folgen

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du = f(x) + \psi_1(K) + \psi_2(l) + \psi_3(K, l).$$

Für **festes K könnte** dann zunächst der Grenzwert $l \rightarrow \infty$ betrachtet werden, wobei insbesondere $\psi_3(K, l)$ gegen Null streben **würde**. Der **anschließende** Grenzübergang $K \rightarrow \infty$ **würde** liefern

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-iut} dt \right] du. \quad (**)$$

In der Tat gilt die Formel (**) unter geeigneten Voraussetzungen an f . Sie ist das Hauptergebnis dieses Kapitels.

Problem. In (*) steht nur eine **endliche** Summe. Um für ein **festes l** mit der Funktion $f_{|[-l\pi, l\pi]}$ argumentieren zu können und für diese wie oben angedeutet Satz 26.2.1 ausnutzen zu können, muss **zunächst der Grenzübergang $K \rightarrow \infty$** betrachtet werden. Es gilt jedoch

$$\lim_{K \rightarrow \infty} |\psi_3(K, l)| \leq \lim_{K \rightarrow \infty} \frac{2K}{\pi} l^{-\alpha} = \infty!$$

Dementsprechend beinhaltet die oben angedeutete Schlussweise **eine unzulässige Vertauschung von Grenzwerten**, die ohne wesentliche Zusatzinformationen **nicht zum Ziel führen kann**. Es sei angemerkt, dass die unzulässige Vertauschung von Grenzwerten eine der häufigsten Fehlerquellen in der Analysis ist.

Zur Präzisierung wird nun (*) zunächst geschrieben als:

$$\begin{aligned}
 & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-iut} dt \right] du \\
 &= \underbrace{\sum_{k=-\infty}^{\infty} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-l\pi}^{l\pi} f(t)e^{-i\frac{k}{l}t} dt}_{=f(x)} \\
 & \quad - \sum_{k=-\infty}^{-Kl} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-l\pi}^{l\pi} f(t)e^{-i\frac{k}{l}t} dt \\
 & \quad - \sum_{k=Kl}^{\infty} e^{i\frac{k}{l}x} \frac{1}{2\pi l} \int_{-l\pi}^{l\pi} f(t)e^{-i\frac{k}{l}t} dt \\
 & \quad + \psi_1(K) + \psi_2(l) + \psi_3(K, l) . \qquad (***)
 \end{aligned}$$

Bemerkungen.

i) Die obigen Reihen existieren nach Satz 26.2.1 als Fourier-Reihen von $f|_{[-l\pi, l\pi]}$.

ii) Für $|x| < l\pi$ ist der erste Ausdruck auf der rechten Seite von (***) nach Satz 26.2.1 gleich $f|_{[-l\pi, l\pi]} = f(x)$.

In der Formel (***) setzt man

$$\psi_4(K, l) := - \sum_{k=-\infty}^{-Kl} \dots - \sum_{k=Kl}^{\infty} \dots$$

Dabei ist anzumerken, dass der Parameter l sowohl in den **Summationsgrenzen** als auch in den **Summanden** auftaucht. Dementsprechend ist das Verhalten der Funktion $\psi_4(K, l)$ **a priori völlig ungewiss**.

Behauptung. Es existiert eine reelle Konstante $c > 0$, sodass für alle $K > 0$ und für alle l hinreichend groß gilt

$$|\psi_4(K, l)| \leq \frac{c}{K} .$$

Konsequenz. Ist die Behauptung gezeigt, so folgt aus (***)

$$\left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-iut} dt \right] du - f(x) \right|$$

$$\leq |\psi_1(K)| + |\psi_2(l)| + |\psi_3(K, l)| + \frac{c}{K}.$$

Ist K zunächst fixiert und $l \gg K$ hinreichend groß, so gilt (mit dem fixierten $\varepsilon > 0$)

$$|\psi_2(l)| + |\psi_3(K, l)| \leq \frac{\varepsilon}{2}.$$

Ist anschließend K hinreichend groß gewählt mit

$$|\psi_1(K)| + \frac{c}{K} \leq \frac{\varepsilon}{2},$$

so folgt

$$\left| \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-iut} dt \right] du - f(x) \right| \leq \varepsilon.$$

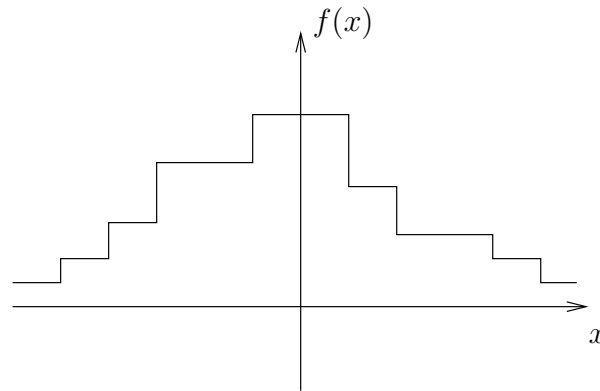
Dies gilt (falls die Behauptung richtig ist) für alle beliebig kleinen $\varepsilon > 0$ und es folgt (**). \square

Beweisidee der Behauptung. Es soll **exemplarisch** ein Ein- und Ausschaltvorgang betrachtet werden, d.h. es sei im Folgenden

$$f(x) = \begin{cases} 1, & \text{falls } -\pi \leq x \leq \pi, \\ 0, & \text{falls } |x| > \pi. \end{cases}$$

Bemerkung. Die Vorstellung ist, dass f approximativ als abzählbare Vereinigung solcher Funktion geschrieben werden kann, wie es in Abbildung 27.6 angedeutet ist. Für ein solches f modifiziere man die folgenden Argumente unter Berücksichtigung der Tatsache, dass f für große x hinreichend schnell abklingt. Wegen der Fülle der Details wird dies nicht ausgeführt.

Es sei an Abbildung 27.4 erinnert, in der die periodische Fortsetzung von $f|_{[-l\pi, l\pi]}$ angedeutet ist. Es handelt sich um eine gerade Funktion,

Abbildung 27.6: Eine Treppenfunktion f .

d.h. die Fourier-Entwicklung in (***) ist lediglich eine Entwicklung in Termen des Kosinus:

$$\int_{-l\pi}^{l\pi} f(t) \cos \left[\frac{k}{l} t \right] dt = \int_{-\pi}^{\pi} \cos \left[\frac{k}{l} t \right] dt = \frac{2l}{k} \sin \left[\frac{k}{l} \pi \right].$$

Betrachtet man schließlich zur Vereinfachung den Punkt $x = 0$, so ist (bis auf Konstanten) die Reihe

$$\sum_{k=Kl}^{\infty} \frac{1}{k} \sin \left[\frac{k}{l} \pi \right]$$

zu analysieren.

Bemerkung. Betragsmäßig kann diese Reihe lediglich durch die **divergente** harmonische Reihe abgeschätzt werden. Auch gelten beispielsweise nicht die Voraussetzungen des Riemannsches Integralkriteriums (vgl. Proposition 6, p. 326, [Hil]). Die letzte Hoffnung auf Konvergenz wird durch das alternierende Vorzeichen der trigonometrischen Funktionen genährt (man vergleiche das Konvergenzkriterium von Leibniz, Satz 4.2.4). Dies verdeutlicht nochmals die Subtilität der Konvergenzfrage.

Wegen

$$\sin \left[\frac{k}{l} \pi \right] = - \sin \left[\frac{k-l}{l} \pi \right],$$

wird wie folgt aufgespalten:

$$\begin{aligned} \sum_{k=Kl}^{\infty} \frac{1}{k} \sin \left[\frac{k}{l} \pi \right] &= \sum_{k=Kl}^{Kl+l} \frac{1}{k} \sin \left[\frac{k}{l} \pi \right] - \sum_{k=Kl+l}^{Kl+2l} \frac{1}{k} \sin \left[\frac{k-l}{l} \pi \right] \\ &+ \dots - \dots + \dots - \dots . \end{aligned}$$

Bemerkung. Streng genommen müssen hier die Partialsummen bis zu einem festen $N \in \mathbb{N}$ betrachtet werden. Da die Konvergenz der Reihe bereits verifiziert wurde, wird auch auf dieses Detail verzichtet.

O.E. sei weiter K gerade, d.h. alle vorkommenden Sinusauswertungen sind nicht-negativ. Die Summe von Kl bis $Kl+l$ ist nach oben abgeschätzt durch

$$\frac{1}{K} \sum_{k=Kl}^{Kl+l} \frac{1}{l} \sin \left[\frac{k}{l} \pi \right] .$$

Die Summe von $Kl+l$ bis $Kl+2l$ ist nach unten abgeschätzt durch (die negative Summe also nach oben durch das Negative von)

$$\frac{1}{K+2} \sum_{k=Kl+l}^{Kl+2l} \frac{1}{l} \sin \left[\frac{k-l}{l} \pi \right] = \frac{1}{K+2} \sum_{k=Kl}^{Kl+l} \frac{1}{l} \sin \left[\frac{k}{l} \pi \right] .$$

Insgesamt erhält man

$$\begin{aligned} \sum_{k=Kl}^{\infty} \frac{1}{k} \sin \left[\frac{k}{l} \pi \right] &\leq \sum_{k=Kl}^{Kl+l} \frac{1}{l} \sin \left[\frac{k}{l} \pi \right] \\ &\cdot \left[\frac{1}{K} - \frac{1}{K+2} + \frac{1}{K+2} - \frac{1}{K+4} + \frac{1}{K+4} \right. \\ &\quad \left. - \dots + \dots - \dots \right] \\ &= \frac{1}{K} \sum_{k=Kl}^{Kl+l} \frac{1}{l} \sin \left[\frac{k}{l} \pi \right] . \end{aligned}$$

Im letzten Schritt beachtet man schließlich, dass (es handelt sich wieder um eine Riemannsche Zwischensumme)

$$\sum_{k=Kl}^{Kl+l} \frac{1}{l} \sin \left[\frac{k}{l} \pi \right] \approx \int_{K\pi}^{(K+1)\pi} \sin(x) \, dx = 2 .$$

Dementsprechend ist für l hinreichend groß gezeigt (das positive Vorzeichen folgt aus analogen Argumenten)

$$0 \leq \sum_{k=Kl}^{\infty} \frac{1}{k} \sin \left[\frac{k}{l} \pi \right] \leq \frac{3}{K} .$$

Dies ist wiederum genau die Behauptung. \square

27.2 Fouriers Integralsatz (absolut integrierbar; Fourier-Integral; Diracsche Deltafunktion; weißes Rauschen)

Bevor der Kernsatz dieses Kapitels präzisiert werden kann, müssen die geeigneten Funktionenklassen eingeführt werden. Es sei zunächst an die Definition 26.2.1, *iii*), einer auf dem Intervall $[a, b]$ stückweise glatten Funktion f erinnert. Ist f nun **auf ganz \mathbb{R}** definiert, so heißt f stückweise glatt, falls die Einschränkung $f|_I$ von f auf jedes abgeschlossene Intervall $I \subset \mathbb{R}$ stückweise glatt ist.

Definition 27.2.1

*i) Man sagt, eine **stückweise glatte** Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ ist von der Klasse \mathcal{D} , falls gilt*

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty .$$

*Die Funktion heißt dann auch **absolut integrierbar**.*

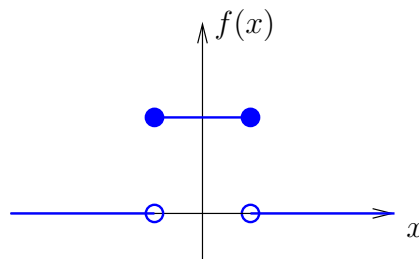
ii) Die Funktion ist von der Klasse \mathcal{D}^ , falls **zusätzlich** gilt*

$$f(x) = \frac{1}{2} \left[f(x^+) + f(x^-) \right] \quad \text{für alle } x \in \mathbb{R} .$$

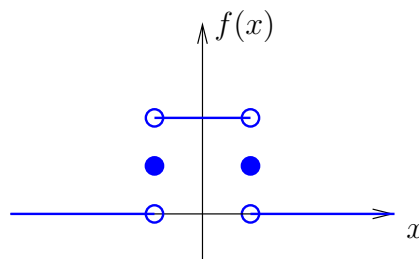
Bemerkung. Die Definition der Funktionenklasse \mathcal{D}^* ist natürlich motiviert durch Satz 26.2.1.

Beispiele.

- i)* Die Funktion $f(x) = e^{-|x|}$ ist absolut integrierbar.
- ii)* Die Funktion $f(x) = \sin(x)$ ist zwar beliebig glatt, liegt aber nicht in der Funktionenklasse \mathcal{D} , da sie nicht absolut integrierbar ist.
- iii)* Die in Abbildung 27.7 dargestellte Funktion liegt in der Klasse \mathcal{D} , aber nicht in der Klasse \mathcal{D}^* .

Abbildung 27.7: $f \in \mathcal{D}$, $f \notin \mathcal{D}^*$.

- iv)* Die in Abbildung 27.8 dargestellte Funktion liegt sowohl in der Klasse \mathcal{D} als auch in der Klasse \mathcal{D}^* .

Abbildung 27.8: $f \in \mathcal{D}$, $f \in \mathcal{D}^*$.

Bemerkung. Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ stückweise glatt und konvergiert das uneigentliche Integral $\int_{-\infty}^{\infty} |f(t)| dt$, so konvergiert auch das **Fourier-Integral**

$$\int_{-\infty}^{\infty} f(t)e^{-iut} dt$$

für **jedes** $u \in \mathbb{R}$ **absolut**. Damit kann sinnvoll definiert werden:

Definition 27.2.2

Es sei $f \in \mathcal{D}$. Dann wird der Funktion f mittels

$$\hat{f}(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-iut} dt \quad \text{für alle } u \in \mathbb{R}$$

eine Funktion $\hat{f}: \mathbb{R} \rightarrow \mathbb{C}$ zugeordnet. Die Funktion \hat{f} heißt die **Fourier-Transformierte** von f . Die Zuordnung $\mathcal{F}: f \mapsto \hat{f} = \mathcal{F}[f]$ heißt **Fourier-Transformation**.

Bemerkung. Die Wahl des Vorfaktors $1/\sqrt{2\pi}$ ist lediglich als Normierung gewählt und variiert in der Literatur.

Nun kann Fouriers Integralsatz präzise formuliert werden, der im einführenden Paragraphen dieses Kapitels bereits ausführlich plausibel gemacht wurde.

Satz 27.2.1

Es sei $f \in \mathcal{D}^*$ und \hat{f} bezeichne die Fourier-Transformierte von f . Dann gilt für alle $x \in \mathbb{R}$

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \hat{f}(u) du ,$$

wobei das uneigentliche Integral $\int_{-\infty}^{\infty} \dots$ als **Cauchyscher Hauptwert** $\lim_{R \rightarrow \infty} \int_{-R}^R \dots$ zu verstehen ist (vgl. Übungen zu Kapitel 12.4). Falls $\int_{-\infty}^{\infty} |\hat{f}(u)| du$ konvergiert, so ist das uneigentliche Integral im üblichen Sinne zu interpretieren.

Bemerkungen.

- i) Wie in der Einleitung bereits hervorgehoben, besagt Satz 27.2.1, dass eine Funktion $f \in \mathcal{D}^*$ aus ihrer kontinuierlichen Spektralzerlegung rekonstruiert werden kann, die Fourier-Transformation ist **invertierbar**.

ii) Es kann äquivalent geschrieben werden

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(t) e^{iu(x-t)} dt \right] du .$$

iii) Es gilt ein Eindeutigkeitsatz, d.h. aus $\hat{f}_1 \equiv \hat{f}_2$ folgt $f_1 = f_2$

Beispiele.

i) Es sei $f(x) = e^{-x^2/2}$. Man setzt

$$\varphi(u) := \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} e^{-iut} dt = \sqrt{2\pi} \hat{f}(u)$$

und erinnert sich an die Diskussion parameterabhängiger Integrale, insbesondere an Satz 12.5.1. Es folgt

$$\varphi'(u) = -i \int_{-\infty}^{\infty} t e^{-\frac{t^2}{2}} e^{-iut} dt .$$

Eine partielle Integration liefert für jedes fixierte $R > 0$

$$\int_{-R}^R t e^{-\frac{t^2}{2}} e^{-iut} dt = -e^{-\frac{t^2}{2}} e^{-iut} \Big|_{-R}^R - iu \int_{-R}^R e^{-\frac{t^2}{2}} e^{-iut} dt ,$$

und im Limes $R \rightarrow \infty$ folgt die lineare gewöhnliche Differentialgleichung erster Ordnung

$$\varphi'(u) = -i \left[-iu \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} e^{-iut} dt \right] = -u\varphi(u)$$

zusammen mit der Anfangsbedingung (vgl. die Diskussion der Fresnelschen Integrale in Kapitel 23.1)

$$\varphi(0) = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt = \sqrt{2\pi} .$$

Dieses Anfangswertproblem hat nach Satz 14.1.1 aber **genau eine** Lösung, und es ist leicht nachzurechnen, dass die Lösung lautet:

$$\varphi(u) = \sqrt{2\pi} e^{-\frac{u^2}{2}} ,$$

dementsprechend ist nach der Definition von φ die Fourier-Transformierte

$$\hat{f}(u) = \frac{1}{\sqrt{2\pi}} \varphi(u) = e^{-\frac{u^2}{2}} = f(u) ,$$

es ist f ein **Fixpunkt** der Fourier-Transformation.

ii) Die anschauliche Vorstellung ist, dass beispielsweise dem Kosinus eine diskrete Frequenz zugeordnet wird (Aber Vorsicht: Der Kosinus kann nicht nach Satz 27.2.1 transformiert werden, warum? Man vergleiche auch die abschließende Bemerkung dieses Kapitels.), die Funktion $e^{-x^2/2}$ wird nach dem vorherigen Beispiel auf sich selbst transformiert und die extreme Situation “in die andere Richtung” ist ein “scharfes Signal”, dem, wie nun gezeigt wird, eine Fourier-Transformierte mit **konstanter Amplitude** zugeordnet wird, man spricht von **weißem Rauschen**.

Ein “scharfes Signal” wird repräsentiert durch die **Diracsche Deltafunktion**

$$\delta(x) := \begin{cases} \infty, & \text{falls } x = 0, \\ 0, & \text{falls } x \neq 0, \end{cases}$$

$$\int_{-\infty}^{\infty} \delta(t) dt = 1.$$

Diese Schreibweise ist in der physikalischen Literatur üblich, natürlich handelt es sich hier nicht um eine Funktion, man spricht von einer **Distribution**. Ohne auf weitere Details einzugehen wird hier für $0 < \varepsilon \ll 1$ die **Approximation**

$$\delta_\varepsilon := \frac{\varepsilon}{\pi(\varepsilon^2 + t^2)}$$

betrachtet, wobei zu beachten ist, dass für alle $\varepsilon > 0$ gilt

$$\int_{-\infty}^{\infty} \delta_\varepsilon(t) dt = \frac{\varepsilon}{\pi} \lim_{R \rightarrow \infty} \frac{1}{\varepsilon} \arctan\left(\frac{t}{\varepsilon}\right) \Big|_{-R}^R = 1.$$

Die Behauptung ist nun, dass die Fourier-Transformierte lautet:

$$\hat{\delta}_\varepsilon(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\varepsilon e^{-iut}}{\pi(\varepsilon^2 + t^2)} dt = \frac{e^{-\varepsilon|u|}}{\sqrt{2\pi}}. \quad (*)$$

Nach Satz 27.2.1 ist

$$\delta_\varepsilon(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iux} \hat{\delta}_\varepsilon(u) du,$$

es ist also zu verifizieren (die Fourier-Transformierte ist eindeutig)

$$\frac{\varepsilon}{\pi(\varepsilon^2 + x^2)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iux} e^{-\varepsilon|u|} du. \quad (**)$$

Dazu wird das Integral als ein Integral über die positive reelle Achse geschrieben, nach der Definition des Betrags ist (Transformation von $\int_{-\infty}^0 \dots$ auf $\int_0^{\infty} \dots$)

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iux} e^{-\varepsilon|u|} du &= \frac{1}{2\pi} \int_0^{\infty} e^{-\varepsilon u} [e^{-iux} + e^{iux}] du \\ &= \frac{1}{2\pi} \int_0^{\infty} [e^{-u(\varepsilon+ix)} + e^{-u(\varepsilon-ix)}] du \end{aligned}$$

und die Definition des unbestimmten Integrals liefert

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iux} e^{-\varepsilon|u|} du &= \frac{1}{2\pi} \lim_{R \rightarrow \infty} \left[\frac{e^{-u(\varepsilon+ix)}}{-(\varepsilon+ix)} + \frac{e^{-u(\varepsilon-ix)}}{-(\varepsilon-ix)} \right]_{u=0}^{u=R} \\ &= \frac{1}{2\pi} \left[\frac{1}{\varepsilon+ix} + \frac{1}{\varepsilon-ix} \right] = \frac{\varepsilon}{\pi(\varepsilon^2+x^2)}, \end{aligned}$$

es folgt (**) und damit (*). Anhand von (*) sieht man schließlich, dass im Grenzwert $\varepsilon \rightarrow 0$ eine konstante Amplitude entsteht:

$$\lim_{\varepsilon \rightarrow 0} \hat{\delta}_{\varepsilon}(u) = \frac{1}{\sqrt{2\pi}} \quad \text{für alle } u \in \mathbb{R}.$$

Es kann als Faustregel festgehalten werden: **Einem scharf lokalisierten Signal entspricht ein breites Frequenzspektrum**, im obigen Extremfall ein weißes Rauschen.

iii) Formal kann geschrieben werden

$$F[\cos(\omega x)] = \sqrt{\frac{\pi}{2}} (\delta(x-\omega) + \delta(x+\omega)).$$

iv) Für weitere Beispiele sei auf die Übungen verwiesen.

Kapitel 28

Bemerkungen zur Laplace-Transformation

28.1 Die Transformation (Heaviside-Funktion; konvergenzerzeugender Faktor; exponentielle Ordnung)

Eng verwandt mit der Fourier-Transformation ist die [Laplace-Transformation](#), die hier nur kurz und exemplarisch vorgestellt werden soll. Für detailliertere Ausführungen sei auf die Literatur (z.B. [Bä] oder weiterführende Spezialliteratur verwiesen).

Ausgangspunkt ist die Beobachtung, dass die Funktionenklasse \mathcal{D}^* , in der eine Fourier-Transformierte nach Satz 27.2.1 definiert ist, für viele Anwendungen zu eng ist. Beispielsweise ist ein Einschaltvorgang zum Zeitpunkt $x = 0$ charakterisiert durch die [Heaviside-Funktion](#)

$$\theta(x) := \begin{cases} 0 & \text{für } x < 0; \\ 1 & \text{für } x \geq 0. \end{cases}$$

Diese Funktion ist **nicht absolut integrierbar** und liegt damit nicht in der Funktionenklasse \mathcal{D}^* . Allgemein werden in diesem Kapitel **stets** Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ mit der Eigenschaft

$$f(x) = 0 \quad \text{für } x < 0$$

betrachtet, m.a.W. Funktionen $f: [0, \infty) \rightarrow \mathbb{R}$, die für $x < 0$ durch die Nullfunktion fortgesetzt werden.

Die Idee zur Umgehung des obigen Problems ist es nun, einen sogenannten **konvergenzerzeugenden Faktor**

$$e^{-\alpha x}, \quad \alpha > 0 \text{ fixiert,}$$

einzuführen und das Verhalten der Funktion

$$f^*(x) = \begin{cases} 0 & \text{für } x < 0, \\ e^{-\alpha x} f(x) & \text{für } x \geq 0 \end{cases}$$

zu studieren. Die Hoffnung ist, dass f^* aufgrund der schnell fallenden Exponentialfunktion ($\alpha > 0, x \geq 0$) absolut integrierbar ist. Ist dies der Fall, so lautet die Fourier-Transformierte

$$\begin{aligned} \hat{f}^*(u) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f^*(t) e^{-iut} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\alpha t} f(t) e^{-iut} dt \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-(\alpha+iu)t} f(t) dt. \end{aligned} \quad (1)$$

Mit der komplexen Notation $z = \alpha + iu$ ergibt sich

$$\hat{f}^*(u) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-zt} f(t) dt.$$

Die rechte Seite wird jetzt nicht mehr als Funktion in $u \in \mathbb{R}$ aufgefasst, sondern als Funktion in der komplexen Variablen z , die Schreibweise ist unter Vernachlässigung des Vorfaktors $1/\sqrt{2\pi}$

$$F(z) := \mathcal{L}[f] := \int_0^{\infty} e^{-zt} f(t) dt. \quad (2)$$

Die Funktionenklasse, auf der eine solche Transformation sinnvoll definiert werden kann ist (vgl. den nachfolgenden Satz 28.1.1)

Definition 28.1.1

Es sei $f: [0, \infty) \rightarrow \mathbb{R}$ stückweise stetig.

- i) Die Funktion f heißt von **exponentieller Ordnung** γ für ein fixiertes $\gamma \in \mathbb{R}$, falls es eine Konstante $M > 0$ gibt, sodass für alle $x \in \mathbb{R}$

$$|f(x)| \leq Me^{\gamma x}.$$

- ii) Die Funktion ist heißt der **Klasse \mathcal{E}_γ** , falls zusätzlich für alle $x \in [0, \infty)$ (mit der Notation $f(0^-) = 0$) gilt

$$f(x) = \frac{1}{2} [f(x^+) + f(x^-)].$$

Beispiele.

- i) Offensichtlich sind alle Polynome für alle $\gamma > 0$ von exponentieller Ordnung γ .
- ii) Die Funktion $f: [0, \infty) \rightarrow \mathbb{R}$, $f(x) = e^{\beta x}$ für ein fixiertes $\beta \in \mathbb{R}$, ist von exponentieller Ordnung γ , falls gilt $\beta \leq \gamma$.
- iii) Ersetzt man in den obigen Beispielen den Funktionswert im Nullpunkt durch $\frac{1}{2}f(0)$, so ist f jeweils von der Klasse \mathcal{E}_γ .

Es gilt:

Satz 28.1.1

Ist f von der Klasse \mathcal{E}_γ , $\gamma \in \mathbb{R}$, so wird dieser Funktion mittels (2) die **Laplace-Transformierte**

$$F(z) = \mathcal{L}[f] = \int_0^\infty e^{-zt} f(t) dt$$

zugeordnet. Die Laplace-Transformierte existiert für alle $z \in \mathbb{C}$ mit **$\operatorname{Re} z > \gamma$** .

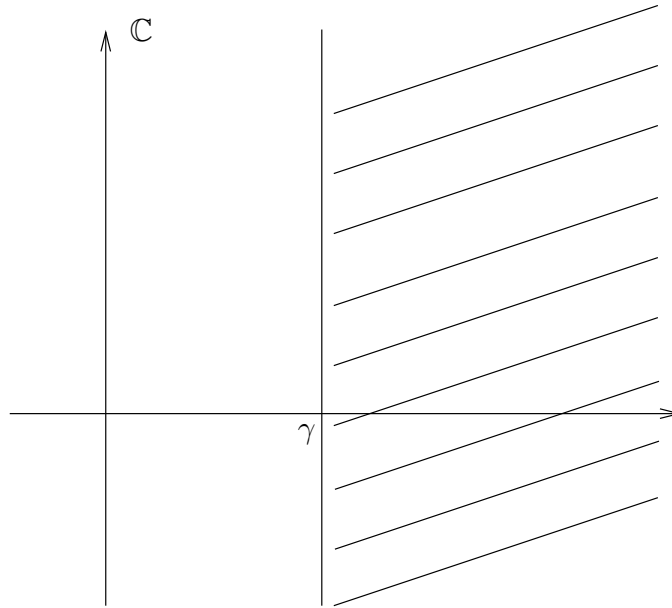


Abbildung 28.1: Das Konvergenzgebiet der Laplace-Transformation.

Bemerkungen.

- i)* An dieser Stelle ist nur wesentlich, dass f von exponentieller Ordnung ist, die Klasse \mathcal{E}_γ wird erst zur Diskussion der inversen Laplace-Transformation benötigt.
- ii)* Das Konvergenzgebiet der Laplace-Transformation einer Funktion von exponentieller Ordnung γ ist in Abbildung 28.1 angedeutet.

Beispiel. Es sei $f(x) = e^{\beta x}$, $\beta \in \mathbb{R}$ fixiert. Dann ist für $\operatorname{Re} z > \beta$

$$\begin{aligned}
 F(z) &= \int_0^\infty e^{-zt} e^{\beta t} dt \\
 &= \frac{1}{\beta - z} e^{(\beta - z)t} \Big|_0^\infty \\
 &= \frac{1}{z - \beta},
 \end{aligned}$$

man erkennt deutlich die Bedeutung des Konvergenzgebietes.

Ebenso wie bei der Fourier-Transformation ist es natürlich essentiell, dass die Funktion f aus ihrer Laplace-Transformierten zurückgewonnen

werden kann. Um das einzusehen sei f von der Klasse \mathcal{E}_γ , $\gamma \in \mathbb{R}$, und wie oben

$$f^*(x) = e^{-\alpha x} f(x),$$

$\alpha \in \mathbb{R}$, $\alpha > \gamma$ fixiert. Wegen $\alpha > \gamma$ und $f \in \mathcal{E}_\gamma$ ist leicht nachzurechnen, dass f^* absolut integrierbar ist (Übung) und die Fourier-Transformierte von f^* wurde bereits in (1) hergeleitet:

$$\hat{f}^*(u) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-(\alpha+iu)t} f(t) dt = \frac{1}{\sqrt{2\pi}} F(\alpha + iu).$$

Satz 27.2.1 liefert

$$f^*(x) = \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R e^{iux} F(\alpha + iu) du,$$

d.h.

$$\begin{aligned} f(x) &= \frac{e^{\alpha x}}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R e^{iux} F(\alpha + iu) du \\ &= \frac{1}{2\pi} \lim_{R \rightarrow \infty} \int_{-R}^R e^{(\alpha+iu)x} F(\alpha + iu) du. \end{aligned}$$

Vorsicht. An dieser Stelle wird bzgl. der **reellen** Variablen u integriert, es handelt sich noch nicht um ein komplexes Wegintegral im Sinne von Definition 22.3.1. Hier ist die Literatur oft missverständlich bzw. die Notation unsauber.

Genauer sei $\varphi_R: [-R, R] \rightarrow \mathbb{C}$ der Weg in der komplexen Ebene, der definiert ist durch die Abbildungsvorschrift

$$u \mapsto \alpha + iu,$$

α wie oben fixiert. Dann verbindet φ_R die Punkte $\alpha - iR$ und $\alpha + iR$ (parallel zur imaginären Achse) und es gilt per definitionem

$$\int_{\varphi_R} e^{zx} F(z) dz = \int_{-R}^R e^{(\alpha+iu)x} F(\alpha + iu) i du,$$

und zusammenfassend ist gezeigt

$$f(x) = \frac{1}{2\pi i} \lim_{R \rightarrow \infty} \int_{\varphi_R} e^{zx} F(z) dz. \quad (3)$$

Bemerkungen.

- i*) Mit (3) ist die **Umkehrung der Laplace-Transformation** bzw. die **inverse Laplace-Transformation** gefunden, es handelt sich um das Gegenstück zu Satz 27.2.1.
- ii*) Wie im Fall der Fouriertransformation gilt ein Eindeutigkeitsatz.
- iii*) Zur Auswertung und Interpretation von (3) sei an die ausführliche Diskussion von komplexen Kurvenintegralen in Teil IX, erinnert. Hier ist insbesondere der Residuensatz sehr hilfreich.
- iv*) Die inverse Laplace-Transformation ist in der Regel recht kompliziert (und wie gesagt oft mit der Hilfe des Residuensatzes) auszurechnen. In der Praxis entnimmt man sie oft aus Tabellen (siehe z.B. [HSZ]).
- v*) Wie eingangs erwähnt, soll hier nicht weiter auf die Laplace-Transformation eingegangen werden, für Rechenregeln und weitere Eigenschaften sei auf die Literatur verwiesen. Gleiches gilt für Anwendungen etwa bei der Diskussion von Anfangswertproblemen. Statt dessen soll in nächsten Teil das systematische Studium von Anfangswertproblemen begonnen werden.

Teil XI

Gewöhnliche Differentialgleichungen (Teil 3): Anfangswertprobleme

Kapitel 29

Der Banachsche Fixpunktsatz, Nullstellenbestimmung

29.1 Der Banachsche Fixpunktsatz (Newton-Verfahren; Fixpunkt; Fixpunktiteration; Verfahrensfunktion; vollständig; Ba- nachraum; Kontraktion)

Eine typische Anwendung des Banachschen¹ Fixpunktsatzes ist die iterative Bestimmung von Nullstellen einer gegebenen Funktion, das sogenannte [Newton-Verfahren](#). Die wesentliche Folgerung des Banachschen Fixpunktsatzes wird aber der [Satz von Picard²-Lindelöf³](#) zur Existenz und Eindeutigkeit von Lösungen von Anfangswertproblemen sein.

Doch zunächst zur Idee des Newton-Verfahrens im einfachsten Fall: Es sei $f: \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion mit einer oder mehreren Nullstellen, die in der Regel nicht explizit berechnet werden können (bekanntlich schon bei Polynomen höheren Grades) und statt dessen mit Hilfe eines iterativen Verfahrens approximativ bestimmt werden sollen. Die Idee dazu ist in Abbildung 29.1 skizziert.

Gestartet wird mit einem beliebigen Punkt $x_0 \in \mathbb{R}$ (nahe einer Nullstelle). An den Graphen von f (hier stets als differenzierbar angenommen) legt man im Punkt $\begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix}$ die Tangente (die ja die

¹S. Banach, 1892–1945; Lwów.

²E. Picard, 1856–1941; Toulouse, Paris.

³E. Lindelöf, 1870–1946; Helsingfors.

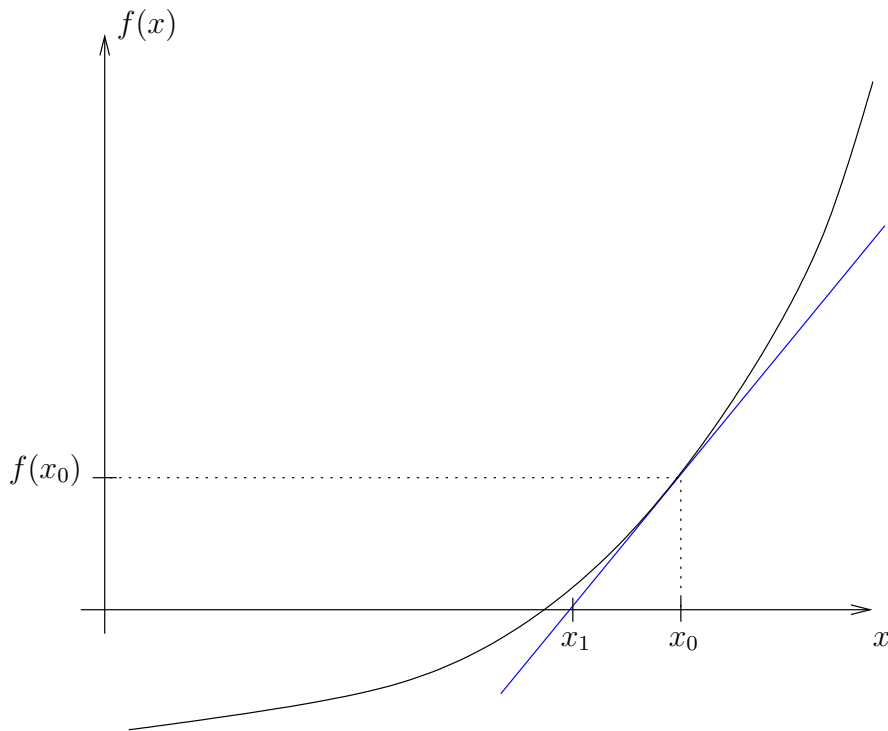


Abbildung 29.1: Zum Newton Verfahren.

Funktion in erster Ordnung approximiert) und hofft, dass die Nullstelle der Tangente näher bei der Nullstelle von f liegt als x_0 . Die Nullstelle der Tangente sei x_1 . Mit dem Punkt x_1 wiederholt man anschließend das Verfahren

Die Tangente ist gegeben durch

$$t(x) = f(x_0) + f'(x_0)(x - x_0) ,$$

die Nullstelle berechnet sich zu

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} ,$$

vorausgesetzt dass $f'(x_0) \neq 0$. Allgemein lautet die Iteration für $k \in \mathbb{N}_0$ (falls f' keine Nullstellen hat)

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)} =: \Phi(x_k) .$$

Falls die so gewonnene Folge $\{x_k\}$ konvergiert mit Grenzwert x , so folgt (vgl. die Diskussion rekursiver Folgen in Kapitel 4)

$$x = \Phi(x) ,$$

es ist x ein **Fixpunkt** der Funktion Φ .

Auch allgemein sind **Fixpunktiterationen** von der Form

$$x_{k+1} = \Phi(x_k), \quad k \in \mathbb{N}_0,$$

mit einer geeigneten **Verfahrensfunktion** Φ . Entscheidend ist dabei die Frage nach der **Konvergenz** der Iteration.

Vor einer formalen Behandlung des Problems sei die grundsätzliche Idee anhand der Abbildungen 29.2 und 29.3 verdeutlicht. **Rot** eingetragen ist jeweils die Winkelhalbierende, die Funktion Φ sei jeweils die **blau** eingetragene Gerade (Φ ist in der Regel natürlich nicht linear).

Gestartet wird mit einem beliebigen Punkt x_0 . Dazu wird der Wert $\Phi(x_0)$ ermittelt. Lläuft man auf der Höhe $\Phi(x_0)$ (also parallel zur x -Achse) bis zur Winkelhalbierenden, so findet man einen Punkt $\begin{pmatrix} x_1 \\ \Phi(x_0) \end{pmatrix}$. Da dieser Punkt aber auf der Winkelhalbierenden liegt, folgt $x_1 = \Phi(x_0)$, der erste Iterationsschritt ist auf diese Weise graphisch ausgeführt.

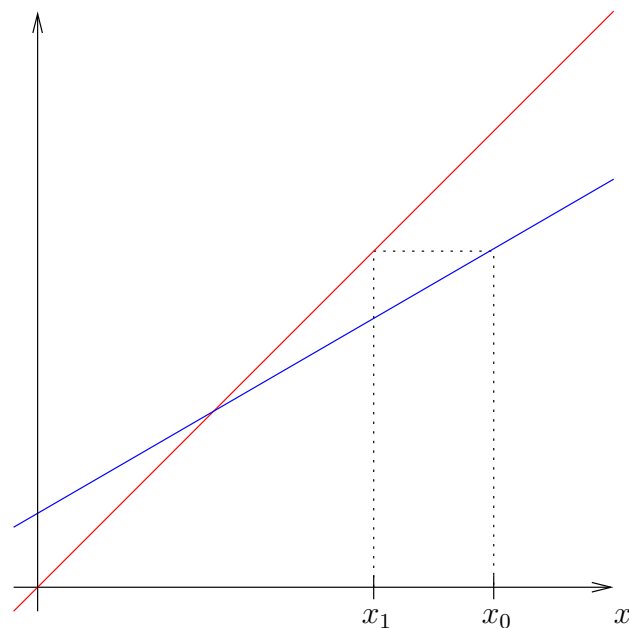


Abbildung 29.2: Konvergenz des Verfahrens.

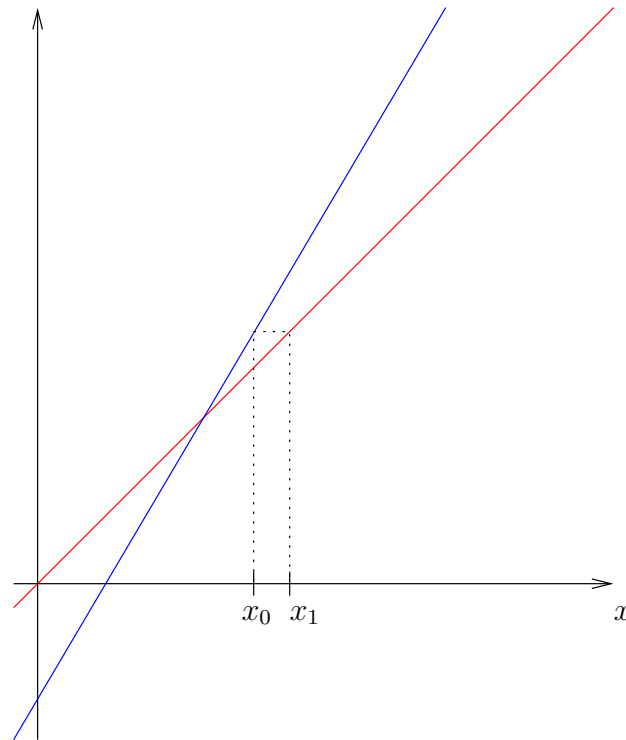


Abbildung 29.3: Divergenz des Verfahrens.

Im ersten Beispiel ist die Steigung von Φ kleiner als die der Winkelhalbierenden (kleiner als 1) es gilt für alle $y, z \in \mathbb{R}$

$$|\Phi(y) - \Phi(z)| \leq c|y - z|$$

mit einer positiven Konstanten $c < 1$. Man erkennt, dass x_1 näher am Schnittpunkt x der Winkelhalbierenden und Φ liegt als x_0 , wobei für den Schnittpunkt gilt $x = \Phi(x)$, der Schnittpunkt ist der gesuchte Fixpunkt.

Im zweiten Beispiel hingegen gilt für alle $y, z \in \mathbb{R}$

$$|\Phi(y) - \Phi(z)| > |y - z|$$

und iterativ entfernt man sich immer mehr vom gesuchten Schnittpunkt, es gibt keine Hoffnung auf die Konvergenz des Verfahrens.

Nachdem in Kapitel 6.1 der allgemeine Begriff eines Vektorraumes eingeführt wurde (vgl. Definition 6.1.1), wurden alle anschließenden Betrachtungen auf den \mathbb{R}^n reduziert, wobei oft auch allgemeinere Situationen hätten zugelassen werden können. Für die Hauptanwendung

des Banachschen Fixpunktsatzes im nächsten Kapitel ist es an dieser Stelle jedoch notwendig, allgemeine Vektorräume V zu betrachten (insbesondere **Funktionenräume**), die mit einer Norm $\|\cdot\|_V$ versehen sind. Der Begriff “Norm” ist in Definition 6.2.1 bereits für beliebige Vektorräume eingeführt, ein Konvergenzbegriff in V kann wörtlich aus Kapitel 4 übertragen werden:

Definition 29.1.1

*Es sei V ein Vektorraum versehen mit einer Norm $\|\cdot\|_V$ ($(V, \|\cdot\|_V)$ ist ein **normierter Raum**).*

*i) Eine Folge $\{x_n\}$ von Elementen aus V heißt **konvergent**, wenn es ein $x \in V$ gibt, sodass gilt: Für jedes $\varepsilon > 0$ existiert eine (**von ε abhängige**) Zahl $N(\varepsilon) \in \mathbb{N}$ mit*

$$\|x_n - x\|_V < \varepsilon \quad \text{für alle } n \geq N(\varepsilon).$$

*ii) Die Folge heißt **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $N(\varepsilon) \in \mathbb{N}$ gibt, sodass gilt:*

$$\|x_n - x_m\|_V < \varepsilon \quad \text{für alle } n, m \geq N(\varepsilon).$$

Bemerkungen.

- i) Die Notationen und Bezeichnungen aus Kapitel 4 werden unverändert übernommen.*
- ii) Für reelle Zahlenfolgen fallen die Begriffe “konvergente Folge” und “Cauchy-Folge” zusammen (Cauchys Konvergenzkriterium, 4.1.7). Ein Analogon ist in vielen normierten Räumen falsch (vgl. nachfolgendes Beispiel und Definition 29.1.2).*

Um überhaupt die Konvergenz einer Fixpunktiteration in einem normierten Raum erwarten zu können, muss dieser einer zusätzlichen Bedingung genügen:

Definition 29.1.2

*Es sei $(V, \|\cdot\|_V)$ ein normierter Raum. Dann heißt V **vollständig**, falls **jede** Cauchy-Folge in V konvergiert, d.h. jede Cauchy-Folge aus V konvergiert gegen einen Grenzwert, der **in V liegt**. Ein vollständiger normierter Raum heißt **Banachraum**.*

Beispiel. Man betrachte den Vektorraum $C^0([a, b])$ der stetigen Funktionen $f: [a, b] \rightarrow \mathbb{R}$ (Übung: Ist dies ein Vektorraum?) versehen mit der Norm

$$\|f\|_{C^0} = \left[\int_a^b |f(x)|^2 dx \right]^{\frac{1}{2}}$$

(Übung: Ist dies eine Norm auf C^0 ?) Anhand des einführenden Beispiels *ii*) aus Kapitel 5.1 wird unmittelbar deutlich, dass dieser normierte Raum nicht vollständig sein kann (Übung!). Es sei ausdrücklich betont, dass es sich hier nicht um ein “künstlich konstruiertes” Beispiel handelt: Das Beispiel gehört zu dem Problemkreis, der eine der wichtigsten und schwierigsten Disziplinen in der Mathematik überhaupt erst notwendig macht, die **Regularitätstheorie für Lösungen partieller Differentialgleichungen**. Die Existenz von Lösungen kann nämlich nicht etwa im Raum der entsprechend oft differenzierbaren Funktionen gezeigt werden, zunächst müssen sogenannte **schwache Lösungen** (verallgemeinerte Lösungen) studiert werden (diese sind a priori in der Regel nicht einmal stetig, obwohl sie in einem gewissen Sinne eine Differentialgleichung lösen). Die Kunst ist es dann, die Eigenschaft “Lösung einer bestimmten Differentialgleichung” geschickt auszunutzen, um evtl. doch Differenzierbarkeitseigenschaften der Lösung beweisen zu können. Dies geht jedoch weit über den Rahmen der Vorlesung hinaus.

Als weiteres Hilfsmittel benötigt man den Begriff der “Lipschitz-Stetigkeit”, der bereits in Kapitel 10.1 kurz diskutiert wurde. Wegen seiner fundamentalen Bedeutung in diesem Kapitel soll er hier nochmals präzisiert werden.

Definition 29.1.3

Es sei $(V, \|\cdot\|_V)$ ein normierter Raum und $D \subset V$. Eine Abbildung $\Phi: D \rightarrow V$ heißt **Lipschitz-stetig** auf D , falls es eine Konstante $L > 0$ gibt, sodass für alle $\underline{\mathbf{x}}, \underline{\mathbf{y}} \in D$:

$$\|\Phi(\underline{\mathbf{x}}) - \Phi(\underline{\mathbf{y}})\|_V \leq L \|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|_V .$$

Eine Konstante L mit dieser Eigenschaft heißt **Lipschitz-Konstante**. Ist Φ Lipschitz-stetig mit einer Konstanten $L < 1$, so heißt Φ **kontrahierend**, man spricht auch von einer **Kontraktion**.

Beispiele.

- i) Eine typische Lipschitz-stetige Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ ist die Betragsfunktion. Diese ist jedoch keine Kontraktion.
- ii) Weitere Beispiele wurden bereits in den Übungen diskutiert.

Als Hauptsatz dieses Paragraphen kann nun formuliert werden:

Satz 29.1.1 (*Banachscher Fixpunktsatz*)

Es sei $(V, \|\cdot\|_V)$ ein Banachraum und $M \subset V$ **abgeschlossen**. Weiterhin sei $\Phi: M \rightarrow M$ eine **Kontraktion** mit einer Lipschitz-Konstanten $L < 1$. Dann existiert **genau ein Fixpunkt** $z \in M$ von Φ in M , d.h. es gilt

$$\Phi(z) = z .$$

Beweis. Der Beweis ist **konstruktiv**, d.h. er zeigt gleichzeitig einen Weg auf, wie die Lösung (zumindest approximativ) gefunden werden kann. Zunächst wird die Existenz eines Fixpunktes gezeigt, dann die Eindeutigkeit.

Existenz. Es sei $x_0 \in M$ beliebig und für $k \in \mathbb{N}_0$ sei

$$x_{k+1} := \Phi(x_k) .$$

Dann gilt nach Voraussetzung für alle $k \geq 1$

$$\|x_{k+1} - x_k\|_V = \|\Phi(x_k) - \Phi(x_{k-1})\|_V \leq L\|x_k - x_{k-1}\|_V .$$

Ist nun $n \in \mathbb{N}$ und $k \geq n$, so folgt iterativ

$$\begin{aligned} \|x_{k+1} - x_k\|_V &\leq L\|x_k - x_{k-1}\|_V \\ &\leq L^2\|x_{k-1} - x_{k-2}\|_V \\ &\leq L^{k-n+1}\|x_{k-(k-n)} - x_{k-(k-n+1)}\|_V \\ &= L^{k-n+1}\|x_n - x_{n-1}\|_V . \end{aligned} \quad (*)$$

Dies zeigt für alle $m \geq n \in \mathbb{N}$ (zu beachten ist die Voraussetzung $L < 1$)

$$\begin{aligned} \|x_m - x_n\|_V &= \|(x_m - x_{m-1}) + (x_{m-1} - x_{m-2}) + \cdots + (x_{n+1} - x_n)\|_V \\ &\leq \sum_{k=n}^{m-1} \|x_{k+1} - x_k\|_V \\ &\leq \|x_n - x_{n-1}\|_V \sum_{k=n}^{m-1} L^{k-n+1} \\ &\leq \|x_n - x_{n-1}\|_V \sum_{l=1}^{\infty} L^l \\ &= \frac{L}{1-L} \|x_n - x_{n-1}\|_V . \end{aligned}$$

Es wurde aber bereits gezeigt (siehe $(*)$ mit den entsprechenden Indizes)

$$\|x_n - x_{n-1}\|_V \leq L^{n-1} \|x_1 - x_0\|_V ,$$

woraus folgt

$$\|x_m - x_n\|_V \leq \frac{L^n}{1-L} \|x_1 - x_0\|_V . \quad (**)$$

Ist schließlich $\varepsilon > 0$ fixiert und $N = N(\varepsilon)$ hinreichend groß gewählt mit

$$\frac{L^N}{1-L} \|x_1 - x_0\|_V < \varepsilon ,$$

so folgt für alle $n, m \geq N(\varepsilon)$

$$\|x_m - x_n\|_V < \varepsilon,$$

die Folge ist eine Cauchy-Folge und konvergiert gegen einen Grenzwert $z \in V$, da V nach Voraussetzung vollständig ist. Da M abgeschlossen ist und $x_k \in M$ für alle $k \in \mathbb{N}_0$, gilt auch $z \in M$.

Nachdem die Konvergenz der Folge verifiziert ist, kann jetzt in der Gleichung

$$x_{k+1} = \Phi(x_k)$$

zur Grenze übergegangen werden (vgl. die Diskussion rekursiver Folgen in Kapitel 4.1) und mit der Stetigkeit von Φ folgt wie behauptet

$$z = \lim_{k \rightarrow \infty} x_{k+1} = \lim_{k \rightarrow \infty} \Phi(x_k) = \Phi(z).$$

Eindeutigkeit. Gäbe es einen weiteren Fixpunkt $\tilde{z} \in M$, so würde gelten

$$\|\tilde{z} - z\|_V = \|\Phi(\tilde{z}) - \Phi(z)\|_V \leq L\|\tilde{z} - z\|_V < \|\tilde{z} - z\|_V,$$

es ergibt sich ein Widerspruch und Satz 29.1.1 ist vollständig bewiesen. \square

Bemerkungen.

- i)* Es sei betont, dass die Iteration für jeden beliebigen Startwert konvergiert. Natürlich hängt aber die Konvergenzgeschwindigkeit davon ab, wie nahe beim Fixpunkt gestartet wird.
- ii)* Im Grenzwert $m \rightarrow \infty$ wird aus (**) für alle $n \in \mathbb{N}$

$$\|z - x_n\|_V \leq \frac{L^n}{1 - L} \|x_1 - x_0\|_V.$$

Der Fehler (d.h. der Abstand von x_n zum gesuchten Fixpunkt) im n -ten Iterationsschritt kann somit explizit abgeschätzt werden.

29.2 Nullstellenbestimmung (Konvergenz des Newton-Verfahrens; modifiziertes Newton-Verfahren)

Wie zu Beginn dieses Kapitels veranschaulicht, ist die näherungsweise Berechnung von Nullstellen eine der Anwendungen des Banachschen Fixpunktsatzes. Dies ist nicht auf den Fall einer Dimension beschränkt,

dennoch sei hier der Einfachheit wegen lediglich der Fall einer Funktion $f: I \rightarrow \mathbb{R}$ betrachtet, wobei $I \subset \mathbb{R}$ ein (verallgemeinertes) Intervall bezeichne. Gesucht sind die Punkte $x \in I$, für die

$$f(x) = 0$$

richtig ist. Ist eine weitere (wie f als differenzierbar angenommene) Funktion $a(x) \neq 0$ für jedes $x \in I$, so ist die Bedingung “ x ist Nullstelle von f ” äquivalent dazu, dass x ein **Fixpunkt** der Abbildung

$$\Phi(x) = x - a(x)f(x)$$

ist. Das bedeutet nämlich

$$a(x)f(x) = 0$$

und da $a(x)$ nicht gleich Null sein soll, ist dies äquivalent zu $f(x) = 0$. (beim Newton-Verfahren wird $a = 1/f'(x)$ gewählt). Da f differenzierbar ist auf I , ist auch die Funktion Φ differenzierbar auf I , und nach dem Mittelwertsatz der Differentialrechnung (11.2.3) ist Φ eine Kontraktion, falls eine Konstante $0 < L < 1$ existiert, sodass für alle $x \in I$ gilt

$$|\Phi'(x)| = |1 - (af)'(x)| \leq L < 1. \quad (*)$$

Auch wenn diese Bedingung auf ganz I erfüllt ist, muss zur Konvergenz einer Iteration **hinreichend nahe bei einer Nullstelle gestartet werden**, d.h. der Funktionswert im Startpunkt sei bereits “recht klein”. Zudem muss sichergestellt sein, dass die Abbildung Φ ein Teilintervall in der Nähe der gesuchten Nullstelle **in sich selbst abbildet**, genauer gesagt: Es sei $x_0 \in I$, $I \supset \tilde{I} = [x_0 - r, x_0 + r]$, $r > 0$ fixiert, es gelte $(*)$ für alle $x \in \tilde{I}$ und es sei weiter

$$|f(x_0)| \leq \frac{r(1-L)}{|a(x_0)|} \quad (**)$$

Dann gilt nach der Definition von Φ

$$|\Phi(x_0) - x_0| \leq |a(x_0)||f(x_0)|$$

und aus $(**)$ folgt für alle $x \in \tilde{I}$

$$\begin{aligned} |\Phi(x) - x_0| &\leq |\Phi(x) - \Phi(x_0)| + |\Phi(x_0) - x_0| \\ &\leq L|x - x_0| + |a(x_0)||f(x_0)| \\ &\leq L|x - x_0| + (1-L)r \leq r, \end{aligned}$$

es ist also $x \in \tilde{I}$ und $\Phi|_{\tilde{I}}$ ist eine Abbildung von \tilde{I} in sich selbst. Damit kann Satz 29.1.1 angewandt werden und zusammenfassend ergibt sich

Satz 29.2.1

Es sei wie oben $f: \mathbb{R} \supset I \rightarrow \mathbb{R}$ eine differenzierbare Funktion und für die (ebenfalls differenzierbare) Funktion $a: I \rightarrow \mathbb{R}$ gelte $a(x) \neq 0$ für alle $x \in I$. Weiterhin sei $\tilde{I} = [x_0 - r, x_0 + r] \subset I$ und für alle $x \in \tilde{I}$ gelte mit einer Konstanten L

$$|1 - (af)'(x)| \leq L < 1 .$$

*Zudem gelte (**). Dann hat f genau eine Nullstelle $z \in \tilde{I}$. Für beliebiges $x_0 \in \tilde{I}$ wie oben ist diese Nullstelle der Grenzwert der Iteration*

$$x_{k+1} = x_k - a(x_k)f(x_k) , \quad k \in \mathbb{N} ,$$

wobei die Fehlerabschätzung

$$|z - x_k| \leq \frac{L^k}{1 - L} |a(x_0)| |f(x_0)|$$

richtig ist.

Bemerkungen.

- i)* Ist f definiert auf ganz \mathbb{R} , so wird die Bedingung (*) aber evtl. nur auf einem kleinen Teilintervall richtig sein.
- ii)* Die Bedingung (**) stellt sicher, dass die Verfahrensfunktion eine Abbildung des Intervalls \tilde{I} in sich selbst ist. Sie ist je leichter zu verifizieren, je näher bei der Nullstelle gestartet wird.
- iii)* Eine ungefähre Kenntnis der Nullstelle und damit eine geeignete Wahl des Startwertes erhält man, indem man zwei Punkte sucht, in denen f unterschiedliches Vorzeichen hat. Dazwischen muss eine Nullstelle liegen (warum?). Durch Intervallhalbierung nähert man sich weiter der Nullstelle.

iv) Hat f eine Nullstelle, so konvergiert das Newton-Verfahren ($a(x) = f'(x)^{-1}$) stets, vorausgesetzt der Startwert liegt hinreichend nahe bei der Nullstelle (Übung!).

Beispiele. Beispielrechnungen zum Newton-Verfahren sind als Übung auf einer Rechenmaschine zu implementieren. Auch bei einfachen Beispielen wird man feststellen, dass die rigorose Verifikation der Voraussetzungen von Satz 29.2.1 in der Regel nicht einfach ist. Ein weiterer Nachteil des Newton-Verfahrens ist es, dass in jedem Schritt $f'(x)^{-1}$ bestimmt werden muss, was mit einem großen Rechenaufwand verbunden ist.

Als Alternative zum Newton-Verfahren kann man etwa $a(x) = \textit{konstant} = (f'(x_0))^{-1}$ wählen, man spricht vom sogenannten **modifizierten Newton-Verfahren**. Dieses konvergiert zwar nicht so schnell, erfordert aber nicht so viel Rechenaufwand und die Voraussetzungen des Satzes können leichter verifiziert werden.

Beispiel. Es sei $f(x) = x^2 + x - 2$, $x_0 = 9/10$ und $\tilde{I} = [x_0 - r, x_0 + r]$. Dabei sollen r und die Kontraktionskonstante L so bestimmt werden, dass das **modifizierte** Newton-Verfahren nach Satz 29.2.1 konvergiert. Es ist

$$a = \textit{konstant} = \frac{1}{f'(x_0)} = \frac{1}{2x_0 + 1} = \frac{5}{14}.$$

Die Kontraktionsbedingung ist

$$\begin{aligned} & \left| 1 - \frac{5}{14}(2x + 1) \right| \leq L \\ \Leftrightarrow & \left| \frac{14}{5} - 2x - 1 \right| \leq \frac{14}{5}L \\ \Leftrightarrow & \left| 2\left(x - \frac{9}{10}\right) \right| \leq \frac{14}{5}L. \end{aligned}$$

Auf \tilde{I} muss demnach mit einer Konstanten $L < 1$ gelten

$$|x - x_0| \leq \frac{14}{10}L,$$

d.h. die erste Bedingung lautet

$$r \leq \frac{14}{10}L .$$

Die Bedingung (**) aus Satz 29.2.1 wird zu

$$\frac{29}{100} \frac{5}{14} = \frac{29}{280} \leq r(1 - L) .$$

Alle Voraussetzungen des Satzes sind also erfüllt, falls

$$\frac{29}{280} \leq r(1 - L) \leq \frac{14}{10}L(1 - L)$$

und falls $L < 1$. Dies bedeutet, dass $L < 1$ so gewählt werden können muss, dass

$$\frac{29}{280} \frac{10}{14} = \frac{29}{392} < L(1 - L) .$$

Dies ist in der Tat aber möglich (z.B. ist die Wahl $L = 0.925$ zulässig, dann wird $r = L14/10$ gewählt), und die Iteration konvergiert nach Satz 29.2.1 (mit $z = 1$ als Grenzwert).

Kapitel 30

Der Satz von Picard-Lindelöf

30.1 Zwei Beispiele (Nicht-Eindeutigkeit; Nicht-Existenz)

Zur Motivation des Satzes sollen hier zunächst zwei Beispiele diskutiert werden.

Beispiel. Betrachtet sei das (skalare) Anfangswertproblem

$$y' = \sqrt{|y(x)|}, \quad y(2) = 1.$$

Die Behauptung ist, dass die Funktionen ($a \leq 0$ fixiert)

$$\phi_a(x) = \begin{cases} x^2/4, & \text{falls } x > 0, \\ 0, & \text{falls } a \leq x \leq 0, \\ -(x-a)^2/4, & \text{falls } x < a, \end{cases}$$

sowie

$$\psi(x) = \begin{cases} x^2/4, & \text{falls } x > 0, \\ 0, & \text{falls } x \leq 0, \end{cases}$$

das Anfangswertproblem lösen. Mittels einer Probe wird die Behauptung leicht als Übung verifiziert, wobei zu beachten ist, dass die stückweise definierten Funktionen tatsächlich differenzierbar sind. Die Funktionsgraphen sind in den Abbildungen 30.1 und 30.2 skizziert.

Beobachtung.

- i)* Es gibt **unendlich viele Lösungen des Anfangswertproblems**.
M.a.W.: **Anfangswertprobleme sind i.A. nicht eindeutig lösbar.**

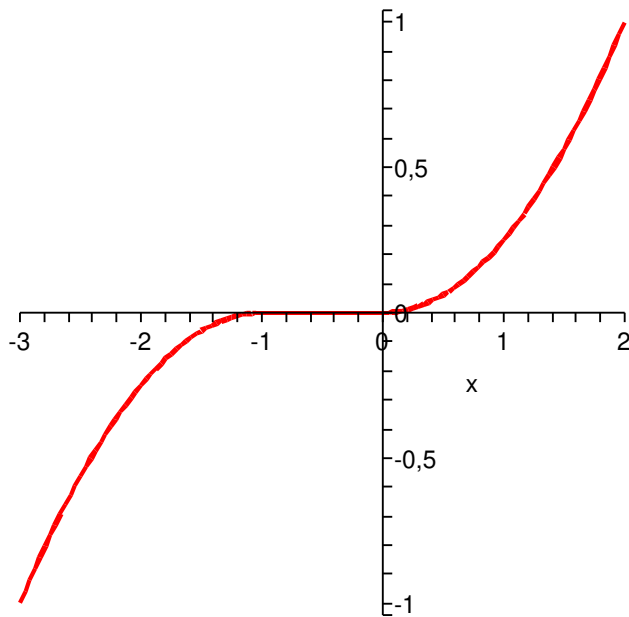


Abbildung 30.1: Eine mögliche Lösung des Anfangswertproblems.

- ii) Die Funktion $\sqrt{|y|}$ (als Funktion in der Variablen y) ist **nicht Lipschitz-stetig**.
- iii) Die Lösungen **existieren für alle Zeiten**.

Beispiel. Betrachtet sei die Differentialgleichung

$$y' = e^{y(x)} \sin(x)$$

mit der Lösung (Probe!)

$$y_c(x) = -\ln(\cos(x) + c),$$

wobei die Integrationskonstante c durch eine Anfangsbedingung festgelegt ist.

Beobachtung.

- i) Ist $y(0) = y_0 < -\ln(2)$, so ist $c > 1$, die Lösung existiert für alle Zeiten.

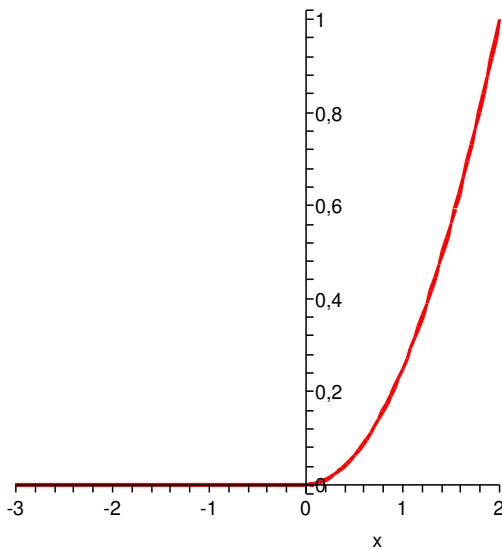


Abbildung 30.2: Eine weitere mögliche Lösung des Anfangswertproblems.

ii) Ist $y(0) = -\ln(2)$, so lautet die Lösung

$$y(x) = -\ln(1 + \cos(x)) ,$$

die Lösung existiert nur im Intervall $(-\pi, \pi)$, sie strebt gegen unendlich, falls $x \rightarrow \pm\pi$, sie kann nicht fortgesetzt werden. M.a.W.: Die Existenz von Lösungen zu Anfangswertproblemen kann nicht ohne Zusatzvoraussetzungen angenommen werden.

30.2 Der Satz (Lipschitz-Bedingung; Picard-Iteration)

In diesem Kapitel wird das folgende Anfangswertproblem für ein System gewöhnlicher Differentialgleichungen erster Ordnung betrachtet:

$$\underline{y}' = F(x, \underline{y}(x)) , \quad \underline{y}(x_0) = \underline{y}^{(0)} .$$

Dabei ist die gesuchte Funktion y eine Funktion eines reellen Zeitintervalls in den \mathbb{R}^N , es handelt sich also um N Gleichungen.

Erinnerung. Gleichungen höherer Ordnung können auf ein System erster Ordnung zurückgeführt werden (vgl. Kapitel 16). Deshalb sind Systeme erster Ordnung von besonderem Interesse.

Satz 30.2.1

Es sei $I \subset \mathbb{R}$ ein Intervall und $x_0 \in I$. Die Funktion

$$F : I \times \mathbb{R}^N \rightarrow \mathbb{R}^N$$

sei stetig und *erfülle eine Lipschitz-Bedingung bzgl. der zweiten Variablen*. D.h.: Es existiert eine Konstante $L > 0$, sodass für alle $x \in I$ und für alle $\underline{\mathbf{y}}^{(1)}, \underline{\mathbf{y}}^{(2)} \in \mathbb{R}^N$ gilt

$$\|F(x, \underline{\mathbf{y}}^{(1)}) - F(x, \underline{\mathbf{y}}^{(2)})\| \leq L \|\underline{\mathbf{y}}^{(1)} - \underline{\mathbf{y}}^{(2)}\| .$$

Dann gibt es zu jedem Anfangswert $\underline{\mathbf{y}}^{(0)} \in \mathbb{R}^N$ *genau eine* stetig differenzierbare Abbildung $\underline{\mathbf{y}} : I \rightarrow \mathbb{R}^N$ mit

$$\begin{aligned} \underline{\mathbf{y}}'(x) &= F(x, \underline{\mathbf{y}}(x)) \quad \text{für alle } x \in I , \\ \underline{\mathbf{y}}(x_0) &= \underline{\mathbf{y}}^{(0)} . \end{aligned}$$

Bemerkung. Es gibt allgemeinere Varianten des Satzes, in denen die Lipschitzbedingung nicht auf einem kompletten Streifen der Form $S = I \times \mathbb{R}^n$ gefordert ist. Darauf soll an dieser Stelle jedoch nicht näher eingegangen werden.

Beweis. Der Beweis beginnt mit einer

Beobachtung.

- i) Falls eine Lösung $\underline{\mathbf{y}}$ existiert, so gilt nach dem Hauptsatz der Differential- und Integralrechnung

$$\underline{\mathbf{y}}(x) = \underline{\mathbf{y}}^{(0)} + \int_{x_0}^x F(t, \underline{\mathbf{y}}(t)) dt .$$

- ii) Betrachtet man die Abbildung $\Phi : C^0(I; \mathbb{R}^N) \rightarrow C^0(I; \mathbb{R}^N)$ (Φ ordnet also einer stetigen Funktion eine stetige Funktion zu), die definiert ist durch

$$(\Phi(\underline{\mathbf{f}}))(x) := \underline{\mathbf{y}}^{(0)} + \int_{x_0}^x F(t, \underline{\mathbf{f}}(t)) dt \quad \text{für alle } x \in I$$

(die Bildfunktion ist festgelegt, wenn ihr Wert für jedes $x \in I$ bekannt ist), so ist nach *i*) ein Fixpunkt dieser Abbildung zu suchen.

Es sei nun $\delta > 0$ fixiert und J das Zeitintervall $J := [x_0 - \delta, x_0 + \delta]$. Gezeigt werden soll, dass Φ für hinreichend kleines δ **eine Kontraktion auf dem Banachraum der auf J stetigen Funktionen ist**. Die Norm ist dabei die Supremumsnorm

$$\|\underline{\mathbf{g}}\|_\infty = \max_{x \in [x_0 - \delta, x_0 + \delta]} \|\underline{\mathbf{g}}(x)\| .$$

(Die Vollständigkeit dieses Raumes ist als Übung zu beweisen.)

Zum Beweis obiger Behauptung werden beliebige Funktionen $\underline{\mathbf{f}}, \underline{\mathbf{g}} \in C^0(J; \mathbb{R}^N)$ fixiert. In J gilt für alle $x > x_0$ ($x < x_0$ analog, Vorzeichen beachten)

$$(\Phi(\underline{\mathbf{g}}))(x) - (\Phi(\underline{\mathbf{f}}))(x) = \int_{x_0}^x (F(t, \underline{\mathbf{g}}(t)) - F(t, \underline{\mathbf{f}}(t))) dt ,$$

woraus unmittelbar die Abschätzung

$$\|(\Phi(\underline{\mathbf{g}}))(x) - (\Phi(\underline{\mathbf{f}}))(x)\| \leq \int_{x_0}^x \|(F(t, \underline{\mathbf{g}}(t)) - F(t, \underline{\mathbf{f}}(t))) dt\|$$

folgt. Da F nach Voraussetzung eine Lipschitz-Bedingung bzgl. der zweiten Variable erfüllt, ist gezeigt

$$\begin{aligned} \|(\Phi(\underline{\mathbf{g}}))(x) - (\Phi(\underline{\mathbf{f}}))(x)\| &\leq \int_{x_0 - \delta}^{x_0 + \delta} L \|\underline{\mathbf{g}} - \underline{\mathbf{f}}\|_\infty dt \\ &= 2\delta L \|\underline{\mathbf{g}} - \underline{\mathbf{f}}\|_\infty . \end{aligned}$$

Dies ist für alle $x \in J$ richtig und auf der linken Seite kann zum Supremum übergegangen werden mit dem Resultat

$$\|\Phi(\underline{\mathbf{g}}) - \Phi(\underline{\mathbf{f}})\|_\infty \leq 2\delta L \|\underline{\mathbf{g}} - \underline{\mathbf{f}}\|_\infty .$$

Ist δ hinreichend klein, so liefert der Banachsche Fixpunktsatz die eindeutige Lösung des Anfangswertproblems auf J . Stückweise wird der Satz anschließend auf ganz I bewiesen. \square

Bemerkungen.

- i) Die Beweiskonstruktion nennt man auch die **Methode der sukzessiven Approximation** oder **Picard-Iteration**. Die Methode führt **konstruktiv** auf Näherungslösungen.
- ii) Die Iteration wird in der Regel mit der Funktion $\underline{y}_0(x) \equiv \underline{y}^{(0)}$ gestartet.

Beispiel. Es werde das Anfangswertproblem

$$y' = x + y, \quad y(0) = 1,$$

auf dem Streifen

$$S = \{(x, y) : -1 \leq x \leq 1, y \in \mathbb{R}\}$$

betrachtet. Klarerweise erfüllt die rechte Seite der Differentialgleichung

$$F(x, y) = x + y$$

die im Satz verlangte Lipschitz-Bedingung. Die Picard-Iteration sieht wie folgt aus:

$$y_0(x) = 1,$$

$$y_1(x) = 1 + \int_0^x (t + 1) dt = 1 + x + \frac{x^2}{2},$$

$$y_2(x) = 1 + \int_0^x (t + 1 + t + t^2/2) dt = 1 + x + x^2 + \frac{x^3}{3!},$$

$$y_3(x) = 1 + \int_0^x (t + 1 + t + t^2 + t^3/3!) dt = 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{4!}$$

$$\begin{aligned} y_4(x) &= 1 + \int_0^x (t + 1 + t + t^2 + t^3/3 + t^4/4!) dt \\ &= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{3 \cdot 4} + \frac{x^5}{5!}. \end{aligned}$$

Induktiv kann leicht gezeigt werden:

$$y_n(x) \xrightarrow{n \rightarrow \infty} y(x) = 2e^x - x - 1$$

und mit Hilfe einer Probe verifiziert man, dass es sich tatsächlich um eine Lösung des Anfangswertproblems handelt.

Kapitel 31

Einige einfache Differentialgleichungen

31.1 Geometrische Interpretation (Richtungsfeld; Linienelement; Isokline)

In diesem Abschnitt soll anhand eines einfachen Beispiels auf die geometrische Deutung von gewöhnlichen Differentialgleichungen bzw. Anfangswertproblemen eingegangen werden. Es sei also

$$y' = F(x, y(x))$$

und $F: \mathbb{R}^2 \supset U \rightarrow \mathbb{R}$. Die angesprochene geometrische Deutung erfolgt über das **Richtungsfeld** einer Differentialgleichung: Jedem Punkt $(x, z) \in U$ wird durch die rechte Seite F der Differentialgleichung eine Tangentenrichtung zugeordnet. Ist $y(x)$ eine Lösung der Differentialgleichung, so muss die Kurve $(x, y(x))$ in jedem Punkt in die zugehörige Tangentenrichtung zeigen, Lösungen der Differentialgleichung folgen in diesem Sinne dem Richtungsfeld.

Als Skizze trägt man in jedem Punkt $(x, z) \in U$ eine kleine Strecke mit der Steigung $F(x, z)$ ein (ein **Linienelement**). Zu diesem Zweck werden in der Regel zunächst die **Isoklinen** bestimmt, das sind Kurven auf denen die Linienelemente gleiche Steigung haben.

Beispiel. Betrachtet werde die Differentialgleichung

$$y' = y + x ,$$

die rechte Seite der Differentialgleichung ist hier $F(x, z) = z + x$. Die Isoklinen sind gegeben durch ($k \in \mathbb{R}$)

$$F(x, z) = z + x = k, \quad \text{d.h.} \quad z = k - x.$$

Das Richtungsfeld ist in Abbildung 31.1 angedeutet, ebenso die Lösungen, die sich dem Richtungsfeld anpassen. Die Isokline $z = -1 - x$ ist gleichzeitig Lösung. Weitere Details werden in den Übungen diskutiert.

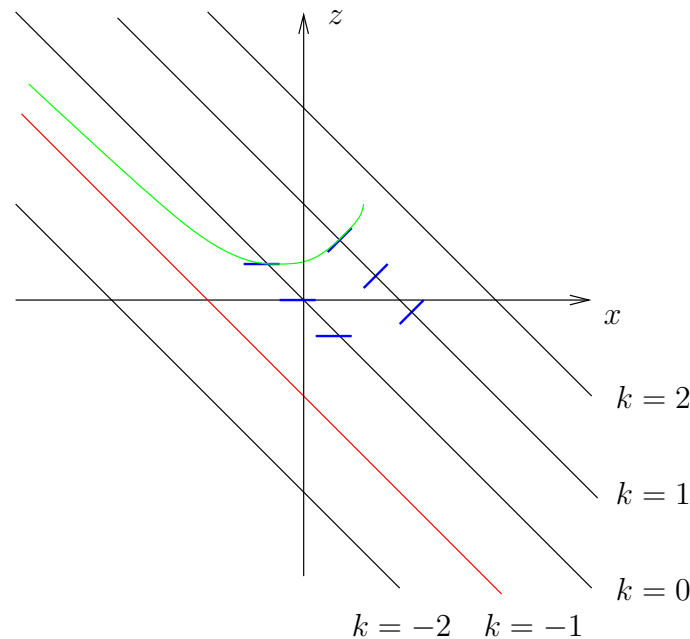


Abbildung 31.1: Das Richtungsfeld der Differentialgleichung.

31.2 Spezielle Typen von Differentialgleichungen

(Trennung der Variablen)

Explizite Lösungen von gewöhnlichen Differentialgleichungen können gefunden werden, falls die Gleichung (die rechte Seite) eine geeignete Struktur hat. In diesem Paragraphen werden vier einfache Typen von Differentialgleichungen vorgestellt, wobei stets der skalare Fall ($N = 1$) betrachtet wird. Alle auftretenden Funktionen seien im Folgenden zumindest stetig und, falls erforderlich, stetig differenzierbar.

Typ 1. Die einfachsten Differentialgleichungen sind vom Typ

$$y' = F(x),$$

d.h. die rechte Seite hängt nicht von der gesuchten Lösung ab. Nach dem Hauptsatz der Differential- und Integralrechnung sind die Lösungen von der Form

$$y(x) = \int_{x_0}^x F(t) dt + konst. ,$$

eine Anfangsbedingung $y(x_0) = y^{(0)}$ liefert

$$y = y^{(0)} + \int_{x_0}^x f(t) dt .$$

Typ 2. Sehr hilfreich ist eine Produktstruktur der Differentialgleichung:

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

wobei die Funktion f auf einem Intervall $I = (a, b)$ gegeben sei, g auf $J = (c, d)$. Es gelte $g(y) \neq 0$ für alle $y \in J$. Dann ist eine **Trennung der Variablen** möglich: Aus

$$\frac{y'}{g(y)} = f(x)$$

wird **formal** abgeleitet

$$\frac{dy}{g(y)} = f(x) dx .$$

Integration ergibt

$$\int \frac{1}{g(y)} dy = \int f(x) dx .$$

Ist Ψ eine Stammfunktion von $\frac{1}{g}$ und ist Φ eine Stammfunktion von f , so sind Lösungen in **impliziter Form** gegeben durch

$$\Psi(y) = \Phi(x) + c , \quad c \in \mathbb{R} . \quad (*)$$

Es ist weiter nach Voraussetzung

$$\frac{d}{dy} \Psi(y) = \frac{1}{g(y)} \neq 0 ,$$

woraus die Existenz einer Umkehrfunktion Ψ^{-1} von Ψ folgt, d.h.

$$y(x) = \Psi^{-1}(\Phi(x) + c)$$

ist Lösung.

Frage: Ist obiges **heuristisches Vorgehen** gerechtfertigt, d.h. löst eine Funktion $y(x)$ mit (*) tatsächlich die Differentialgleichung?

Es gilt:

$$\begin{aligned} \underbrace{\frac{d}{dx}\Psi(y(x))}_{= \frac{d\Psi}{dy}(y(x))y'(x)} &= \underbrace{\frac{d}{dx}[\Phi(x) + c]}_{= f(x)}, \\ &= \frac{d\Psi}{dy}(y(x))y'(x) \end{aligned}$$

also

$$\frac{1}{g(y(x))}y'(x) = f(x)$$

und damit

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

die heuristische Vorgehensweise war somit berechtigt.

Beispiel. Es sei $I = \mathbb{R}$ und

$$y' = xy,$$

in obiger Notation ist $f(x) = x$ und $g(y) = y$.

i) Die Funktion $y \equiv 0$ löst klarerweise die Gleichung.

ii) Ist $y \neq 0$, d.h. $0 \notin J$, so folgt

$$\int \frac{1}{y} dy = \int x dx,$$

implizit ist y gegeben durch

$$\ln(|y|) = \frac{x^2}{2} + c,$$

in diesem Fall kann explizit aufgelöst werden (hier für $y > 0$):

$$y(x) = e^{x^2/2}e^c \quad \text{bzw.} \quad y(x) = Ke^{x^2/2}, \quad K \in \mathbb{R}.$$

Typ 3. Ist mit reellen Konstanten $a, b \neq 0, c$

$$y' = h(ax + by + c),$$

so betrachtet man die Transformation

$$u(x) = ax + by(x) + c .$$

Es folgt (die zweite Gleichheit ergibt sich aus der Differentialgleichung und der Definition von u)

$$u'(x) = a + by'(x) = a + bh(u) .$$

Diese Gleichung ist aber vom Typ 2 ($g(u) = a + bh(u)$, $f(x) \equiv 1$, $u'(x) = f(x)g(u)$) und kann im Prinzip gelöst werden, eine Rücktransformation liefert y .

Beispiel. Es sei

$$y' = e^{2x+3y} , \quad u = 2x + 3y .$$

Zu lösen ist zunächst die Gleichung

$$u' = 2 + 3e^u ,$$

eine Trennung der Variablen liefert

$$\int \frac{1}{2 + 3e^u} du = \int dx .$$

Anhand von

$$\frac{d}{du} \left[\frac{1}{2}u - \frac{1}{2} \ln(2 + 3e^u) \right] = \frac{1}{2 + 3e^u}$$

erkennt man

$$\frac{1}{2}u - \frac{1}{2} \ln(2 + 3e^u) = x + K , \quad K \in \mathbb{R} .$$

Die Definition von u liefert

$$2x + 3y - \ln(2 + 3e^{2x+3y}) = 2x + 2K ,$$

die Lösungen sind in impliziter Form gegeben durch

$$3y - \ln(2 + 3e^{2x+3y}) = c , \quad c \in \mathbb{R} .$$

Typ 5. Für $x \neq 0$ gelte

$$y' = h\left(\frac{y}{x}\right) .$$

Hier ist die Transformation

$$u(x) = \frac{y(x)}{x}$$

zu betrachten, sie führt auf

$$u'(x) = \frac{y'(x)}{x} - \frac{y(x)}{x^2} = \frac{h(u) - u}{x}.$$

Diese Differentialgleichung ist aber wieder vom Typ 2, ist u Lösung, so ist $y(x) = xu(x)$ Lösung der ursprünglichen Gleichung.

Beispiel. Für $x > 0$ sei

$$y' = \frac{x^2 + xy + y^2}{x^2} = 1 + \frac{y}{x} + \frac{y^2}{x^2}.$$

In obiger Notation ist $h(z) = 1 + z + z^2$ und die Differentialgleichung in u lautet

$$u'(x) = \frac{1 + u^2}{x},$$

eine Trennung der Variablen liefert

$$\int \frac{1}{1 + u^2} du = \int \frac{1}{x} dx,$$

die anschließende Integration führt auf

$$\arctan(u) = \ln(|x|) + c, \quad c \in \mathbb{R},$$

und schließlich:

$$y(x) = t \tan(\ln(|x|) + c), \quad c \in \mathbb{R}.$$

Kapitel 32

Numerische Methoden

32.1 Einführung (stetige Abhängigkeit der Lösung von den Daten; Einschrittverfahren; Mehrschrittverfahren)

Die meisten Anfangswertprobleme lassen sich nicht explizit lösen und man bedient sich numerischer Methoden, um zumindest eine Näherungslösung zu finden. Damit diese überhaupt funktionieren können, muss eine Lösung existieren und diese muss eindeutig sein. Dies kann etwa mit Hilfe des Satzes von Picard-Lindelöf überprüft werden. Ebenso muss aber auch garantiert sein, dass “kleine” Änderungen in den Eingabedaten (die auf einer Rechenmaschine nicht exakt dargestellt werden können) nur “kleine” Änderungen in der Lösung produzieren. Dazu sei ohne Beweis der folgende Satz vorausgeschickt:

Satz 32.1.1 (*stetige Abhängigkeit von den Daten*)

Es sei $\emptyset \neq \Omega \subset [a, b] \times \mathbb{R}^N$ offen, $F, G \in C^0(\Omega; \mathbb{R}^N)$, $(a, \underline{\mathbf{u}}_0), (a, \underline{\mathbf{v}}_0) \in \Omega$. Die Funktion F erfülle auf ganz Ω eine (globale) Lipschitz-Bedingung (bzgl. der zweiten Variablen) mit einer Konstanten L . Weiter gelte für ein $\delta > 0$

$$\sup_{(x, \underline{\mathbf{y}}) \in \Omega} \|F(x, \underline{\mathbf{y}}) - G(x, \underline{\mathbf{y}})\| \leq \delta$$

und $\underline{\mathbf{u}} \in C^1([a, b]; \mathbb{R}^N)$ sei (in Ω) eine Lösung des Anfangswertproblems

$$\underline{\mathbf{y}}' = F(x, \underline{\mathbf{y}}(x)) , \quad \underline{\mathbf{y}}(a) = \underline{\mathbf{u}}_0 .$$

Das Anfangswertproblem

$$\underline{\mathbf{y}}' = G(x, \underline{\mathbf{y}}(\mathbf{x})) , \quad \underline{\mathbf{y}}(a) = \underline{\mathbf{v}}_0$$

habe eine Lösung $\underline{\mathbf{v}} \in C^1([a, b]; \mathbb{R}^n)$. Dann gilt für alle $x \in [a, b]$

$$\|\underline{\mathbf{u}}(x) - \underline{\mathbf{v}}(x)\| \leq \|\underline{\mathbf{u}}_0 - \underline{\mathbf{v}}_0\| e^{L(x-a)} + \frac{\delta}{L} \left(e^{L(x-a)} - 1 \right) .$$

Nachdem die Voraussetzungen für die Existenz, die Eindeutigkeit und die stetige Abhängigkeit von den Daten überprüft sind, kann nun das Anfangswertproblem

$$\underline{\mathbf{y}}' = F(x, \underline{\mathbf{y}}(x)) , \quad \underline{\mathbf{y}}(x_0) = \underline{\mathbf{y}}_0 ,$$

numerisch untersucht werden. Es sei daran erinnert, dass das Anfangswertproblem äquivalent ist zur Integralgleichung

$$\underline{\mathbf{y}}(x) = \underline{\mathbf{y}}^{(0)} + \int_{x_0}^x F(t, \underline{\mathbf{y}}(t)) dt$$

ist. Es sei im Folgenden $I = [x_0, b] \subset \mathbb{R}$ ein Intervall, auf dem eine Lösung (wie oben) existiert. In diesem Intervall werden **äquidistante Stützstellen**

$$x_i = x_0 + ih , \quad h = \frac{b - x_0}{n} , \quad i = 1, \dots, n .$$

betrachtet. Bezeichnet φ die exakte Lösung des Anfangswertproblems, so wird versucht, mit Hilfe der Integralgleichung Näherungen $\underline{\mathbf{y}}_i = \underline{\mathbf{y}}(x_i)$ für $\varphi(x_i)$ zu finden, beispielsweise:

$$\underline{\mathbf{y}}(x_1) = \underline{\mathbf{y}}_0 + \int_{x_0}^{x_0+h} F(t, \varphi(t)) dt .$$

Dabei steht natürlich noch die unbekannte exakte Lösung im Integranden auf der rechten Seite, diese wird deshalb numerisch behandelt.

Die numerischen Verfahren unterscheiden sich dadurch, wie dies geschieht:

- i*) **Einschrittverfahren**: Der approximative Wert \underline{y}_{i+1} berechnet sich nur aus dem vorausgehenden Wert \underline{y}_i .
- ii*) **Mehrschrittverfahren**: Zur Berechnung von \underline{y}_{i+1} werden weitere Werte herangezogen (z.B. $\underline{y}_{i-k}, \dots, \underline{y}_i, 1 \leq k \leq i$).

32.2 Das Polygonzugverfahren von Euler (konsistentes Verfahren; Ordnung eines Einschrittverfahrens)

Zur Vereinfachung sei in den folgenden Paragraphen stets o.E. $N = 1$, d.h. es wird jeweils nur eine Gleichung erster Ordnung betrachtet (das macht insbesondere die Taylor-Entwicklungen übersichtlicher). Es bezeichne $\varphi: [x_0, b] \rightarrow \mathbb{R}^n$ die Lösung des Anfangswertproblems

$$\varphi' = F(x, \varphi(x)), \quad \varphi(x_0) = \varphi_0.$$

Idee. Für kleine Schrittweiten h gilt näherungsweise

$$\frac{1}{h}(\varphi(x+h) - \varphi(x)) \approx \varphi'(x) = F(x, \varphi(x)),$$

d.h.

$$\varphi(x+h) \approx \varphi(x) + hF(x, \varphi(x)).$$

Für fixiertes $n \in \mathbb{N}$ setzt man $x_i := x_0 + ih$, $h = (b - x_0)/n$, $i = 1, \dots, n$, und die obige Idee motiviert die Rekursionsformel

$$y_0 := \varphi(x_0), \quad y_{i+1} := y_i + hF(x_i, y_i), \quad i = 0, \dots, n-1. \quad (1)$$

Weitere Idee. In Verallgemeinerung von (1) kann die Iterationsvorschrift

$$y_0 := \varphi(x_0), \quad y_{i+1} := y_i + h\Phi(x_i, y_i, h), \quad i = 0, \dots, n-1. \quad (2)$$

betrachtet werden, wobei Φ so gewählt werden kann soll, dass ein "besseres" Verfahren entsteht.

Zur weiteren Analyse dieser Idee sei nun φ die Lösung des Anfangswertproblems

$$z' = F(t, z(t)) , \quad z(x) = y .$$

Bemerkung. Die Notation ist motiviert durch die folgende Situation: Ist im k^{ten} Schritt zum Zeitpunkt x eine Näherung y für $\varphi(x)$ gefunden, so soll untersucht werden, wie gut lokal zu diesen Daten die Rekursion mit der Differentialgleichung übereinstimmt.

Es sei weiter

$$\Delta(x, y, h) := \begin{cases} \frac{1}{h}(\varphi(x+h) - \varphi(x)) , & \text{falls } h \neq 0 , \\ F(x, y) , & \text{falls } h = 0 , \end{cases}$$

$\Delta(x, y, h)$ ist somit der **Differenzenquotient der exakten Lösung**. Für die Funktion Φ gilt nach (2)

$$\Phi(x_i, y_i, h) = \frac{y_{i+1} - y_i}{h} ,$$

es handelt sich um den **Differenzenquotienten der Näherungslösung zur Schrittweite h** . Ein Maß für den lokalen Fehler ist die Größe

$$\tau(x, y, h) := \Delta(x, y, h) - \Phi(x, y, h) ,$$

τ heißt **lokaler Diskretisierungsfehler** des durch Φ gemäß (2) definierten Verfahrens. Die Erwartung an ein “vernünftiges” Verfahren ist natürlich

$$\lim_{h \rightarrow 0} \tau(x, y, h) = 0 \tag{3}$$

und wegen

$$\lim_{h \rightarrow 0} \Delta(x, y, h) = F(x, y)$$

ist (3) äquivalent zu

$$\lim_{h \rightarrow 0} \Phi(x, y, h) = F(x, y) . \tag{4}$$

Definition 32.2.1

Die Funktion Φ bzw. das gemäß (2) durch Φ definierte Verfahren heißt *konsistent*, falls (4) gilt. Das Verfahren ist *von der Ordnung p* , falls

$$\tau(x, y, h) = O(h^p) .$$

Bemerkungen.

- i) Es sei an die Definition 13.2.1 der Landauschen Symbole erinnert.
- ii) Der Begriff “konsistent” sagt noch nichts über die Konvergenz des Verfahrens aus.
- iii) Das Euler-Verfahren ist trivialerweise konsistent, da in diesem Fall

$$\Phi(x, y, h) \equiv F(x, y) .$$

Zur Berechnung der Ordnung des Euler-Verfahrens sei wieder φ die exakte Lösung des Anfangswertproblems

$$z' = F(t, z(t)) , \quad z(x) = y .$$

Ist die rechte Seite F der Differentialgleichung hinreichend glatt, so gilt das auch für die Lösung φ und eine Taylor-Entwicklung ergibt für ein $\theta \in (0, 1)$

$$\varphi(x + h) = \sum_{k=0}^{p-1} \varphi^{(k)}(x) \frac{1}{k!} h^k + \frac{1}{p!} h^p \varphi^{(p)}(x + \theta h) .$$

An dieser Stelle wird die Information ausgenutzt, dass φ Lösung der Differentialgleichung ist (die Indizes bezeichnen die partiellen Ableitungen nach der entsprechenden Variablen):

$$\begin{aligned} \varphi''(x) &= \frac{d}{dx} \underbrace{F(x, \varphi(x))}_{=\varphi'(x)} \\ &= F_x(x, \varphi(x)) + F_y(x, \varphi(x))\varphi'(x) \\ &= F_x(x, \varphi(x)) + F_y(x, \varphi(x))F(x, \varphi(x)) , \end{aligned}$$

und analog können höhere Ableitungen ausgedrückt werden, etwa

$$\begin{aligned}\varphi'''(x) &= F_{xx}(x, \varphi(x)) + 2F_{xy}(x, \varphi(x))F'(x, \varphi(x)) \\ &\quad + F_{yy}(x, \varphi(x))F'^2(x, \varphi(x)) + F_y(x, \varphi(x))F'_x(x, \varphi(x)) \\ &\quad + F_y^2(x, \varphi(x))F'(x, \varphi(x)) .\end{aligned}$$

Mit der Notation $y = \varphi(x)$ ergibt sich

$$\begin{aligned}\Delta(x, y, h) &= \varphi'(x) + \frac{h}{2!}\varphi''(x) + \dots + \frac{h^{p-1}}{p!}\varphi^{(p)}(x + \theta h) \\ &= F(x, y) + \frac{h}{2}(F_x(x, y) + F_y(x, y)F'(x, y)) + O(h^2) .\end{aligned}\quad (5)$$

Im Spezialfall des Euler-Verfahrens ist

$$\tau(x, y, h) = \Delta(x, y, h) - F(x, y) = O(h) ,$$

das Euler-Verfahren ist von der Ordnung 1.

Bemerkung. Im vektoriellen Fall $N > 1$ gelten zu (5) analoge Formeln.

Die Gleichung (5) zeigt eine mögliche andere Wahl von Φ , nämlich

$$\Phi(x, y, h) := F(x, y) + \frac{h}{2}(F_x(x, y) + F_y(x, y)F'(x, y)) ,$$

mit dieser Wahl ist

$$\tau(x, y, h) = \Delta(x, y, h) - \Phi(x, y, h) = O(h^2) ,$$

d.h. mit diesem Φ in der Rekursion (2) erhält man ein Verfahren der Ordnung 2. Der **Nachteil** ist jedoch, dass in jedem Schritt nicht nur F sondern auch die partiellen Ableitungen von F ausgewertet werden müssen. Deshalb werden in den nächsten Paragraphen geeignetere Verfahren höherer Ordnung vorgestellt.

32.3 Das Verfahren von Heun

Beim Verfahren von Heun wird mit dem **Ansatz**

$$\Phi(x, y, h) := a_1 F(x, y) + a_2 F(x + p_1 h, y + p_2 h F'(x, y)) \quad (1)$$

gestartet. Dabei sollen die reellen Koeffizienten a_1 , a_2 , p_1 und p_2 so bestimmt werden, dass die Ordnung des Verfahrens maximal wird. Betrachtet man die Funktion $h \mapsto \Phi(x, y, h)$, so liefert eine Taylor-Entwicklung

$$\begin{aligned}\Phi(x, y, h) &= \Phi(x, y, 0) + h \frac{d}{dh} \Big|_{h=0} \Phi(x, y, h) + O(h^2) \\ &= (a_1 + a_2)F(x, y) + a_2h(p_1F_x(x, y) + p_2F_y(x, y)F(x, y)) \\ &\quad + O(h^2).\end{aligned}$$

Ein Vergleich mit (5), 32.2, zeigt:

$$\tau(x, y, h) = \Delta(x, y, h) - \Phi(x, y, h) = O(h^2),$$

falls gilt

$$a_1 + a_2 = 1, \quad a_2p_1 = \frac{1}{2}, \quad a_2p_2 = \frac{1}{2}. \quad (2)$$

M.a.W.: Gilt (2), so liefert Φ ein Verfahren zweiter Ordnung. Zu beachten ist, dass (2) aus drei nichtlinearen Gleichungen besteht, eine mögliche Lösung ist

$$a_1 = a_2 = \frac{1}{2}, \quad p_1 = p_2 = 1.$$

Damit wird der Ansatz (1) zu

$$\Phi(x, y, h) = \frac{1}{2} \left[F(x, y) + F(x + h, y + hF(x, y)) \right] \quad (3)$$

und zusammen mit (2), (32.2), liefert (3) die Rekursionsvorschrift (nach Heun \approx 1900)

$$y_{i+1} = y_i + \frac{h}{2} \left[F(x_i, y_i) + F(x_i + h, y_i + hF(x_i, y_i)) \right]. \quad (4)$$

Beobachtung. Im Hinblick auf den Rechenaufwand des Verfahrens ist es wesentlich zu beobachten, dass nach (4) in jedem Schritt die Funktion F an zwei verschiedenen Stellen auszuwerten ist. Beim Euler-Verfahren war die rechte Seite F der Differentialgleichung in jedem Schritt nur an einer Stelle auszuwerten.

32.4 Das modifizierte Euler-Verfahren

Das modifizierte Euler-Verfahren geht auf Collatz (1960) zurück und basiert auf der Beobachtung, dass die Wahl

$$a_1 = 0, \quad a_2 = 1, \quad p_1 = p_2 = \frac{1}{2}$$

eine weitere Lösung des Gleichungssystems (2), 32.3, liefert. Einsetzen in (1), 32.3 liefert

$$\Phi(x, y, h) = F\left(x + \frac{h}{2}, y + \frac{h}{2}F(x, y)\right)$$

und als Rekursionsvorschrift

$$y_{i+1} = y_i + hF\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}F(x_i, y_i)\right). \quad (1)$$

Beobachtung. Auch in (1) ist zu beobachten, dass pro Schritt zwei Auswertungen der rechten Seite F der Differentialgleichung nötig sind.

32.5 Runge-Kutta-Verfahren (mehrstufiges Einschrittverfahren)

Alle bisher vorgestellten Verfahren gehören zur Klasse der Runge-Kutta-Verfahren (1895), bei der in Verallgemeinerung von (1), 32.3, der Ansatz

$$\Phi(x, y, h) := \sum_{i=1}^m \gamma_i k_i(x, y, h) \quad (1)$$

gemacht wird. Dabei sind die γ_i reelle Koeffizienten, die $k_i(x, y, h)$ sind Funktionen mit Werten im \mathbb{R}^N (wieder o.E. $N = 1$) und $m \in \mathbb{N}$ ist fixiert. Der Ansatz für die k_i lautet

$$\begin{aligned} k_1(x, y, h) &= F(x, y), \\ k_2(x, y, h) &= F\left(x + \alpha_2 h, y + h\beta_{21}k_1(x, y, h)\right), \\ &\vdots \\ k_m(x, y, h) &= F\left(x + \alpha_m h, y + h \sum_{l=1}^{m-1} \beta_{ml} k_l(x, y, h)\right). \end{aligned} \quad (2)$$

Idee. Man versuche wie oben mit Taylor-Entwicklungen (jetzt höherer Ordnung, vgl. (5), 32.2) die Koeffizienten

$$\gamma_1, \dots, \gamma_m, \alpha_2, \dots, \alpha_m, \beta_{il}, \quad 2 \leq i \leq m, \quad 1 \leq l < i,$$

so zu bestimmen, dass das Verfahren eine möglichst hohe Konsistenzordnung p bekommt.

Mit der Stetigkeit von F kann man induktiv zeigen

$$\lim_{h \rightarrow 0} k_i(x, y, h) = F(x, y)$$

und zusammen mit (1) folgt

$$\lim_{h \rightarrow 0} \Phi(x, y, h) = F(x, y) \sum_{i=1}^m \gamma_i,$$

also als notwendige Bedingung

Satz 32.5.1

*Das nach (1) und (2) gegebene m -stufige Einschrittverfahren ist **konsistent**, falls*

$$\gamma_1 + \gamma_2 + \dots + \gamma_m = 1.$$

Die Koeffizienten von Runge-Kutta-Verfahren werden in der Regel schematisch festgehalten (vgl. Tabelle 32.1)

Spezialfälle.

i) Ist $m = 1$, so reduziert sich (1) auf

$$\Phi(x, y, h) = \gamma_1 F(x, y)$$

und nach Satz 32.5.1 ist $\gamma_1 = 1$ zu wählen, es handelt sich um das Euler-Verfahren.

ii) Ist $m = 2$, so führt die Suche nach Verfahren möglichst hoher Ordnung wieder auf die Bestimmungsgleichungen (2) mit der Notation $a_i = \gamma_i$, $i = 1, 2$, $p_1 = \alpha_2$, $p_2 = \beta_{21}$. Das Verfahren von Heun wird durch das Koeffizientenschema aus Tabelle 32.2 beschrieben, das modifizierte Euler-Verfahren durch das Schema aus Tabelle 32.3

| | | | | | |
|------------|--------------|--------------|---------|------------------|------------|
| α_2 | β_{21} | | | | |
| α_3 | β_{31} | β_{32} | | | |
| \vdots | \vdots | \vdots | | | |
| α_m | β_{m1} | β_{m2} | \dots | $\beta_{m(m-1)}$ | |
| | γ_1 | γ_2 | \dots | γ_{m-1} | γ_m |

Tabelle 32.1: Koeffizientenschema eines Runge-Kutta-Verfahrens.

| | |
|---|---------|
| 1 | 1 |
| | 1/2 1/2 |

Tabelle 32.2: Koeffizientenschema des Verfahrens von Heun.

| | |
|-----|-----|
| 1/2 | 1/2 |
| | 0 1 |

Tabelle 32.3: Koeffizientenschema des modifizierten Euler-Verfahrens.

| | |
|------------------|--|
| $\alpha_2 = 1/2$ | $\beta_{22} = 1/2$ |
| $\alpha_3 = 1$ | $\beta_{31} = -1 \quad \beta_{32} = 2$ |
| | $\gamma_1 = 1/6 \quad \gamma_2 = 2/3 \quad \gamma_3 = 1/6$ |

Tabelle 32.4: Koeffizientenschema der einfachen Kutta-Regel.

iii) $m = 3$: **Einfache Kutta-Regel**. Taylor-Entwicklungen von $\Delta(x, y, h)$ und $\Phi(x, y, h)$ liefern ein Verfahren der Ordnung 3, falls ein System von 7 **nichtlinearen** Gleichungen für die Koeffizienten erfüllt ist. Ein Lösungsansatz ist etwa

Wieder ist durch $\sum_{i=1}^3 \gamma_i = 1$ die achte Gleichung gegeben. Das Iterationsschema zur einfachen Kutta-Regel lautet:

$$y_{i+1} = y_i + \frac{h}{6} \left[f(x_i, y_i) + 4k_2(x_i, y_i, h) + k_3(x_i, y_i, h) \right],$$

$$k_2(x, y, h) = f\left(x + \frac{h}{2}, y + \frac{h}{2}f(x, y)\right),$$

$$k_3(x, y, h) = f\left(x + h, y + h\left[2f\left(x + \frac{h}{2}, y + \frac{h}{2}f(x, y)\right) - f(x, y)\right]\right).$$

Die Ordnung des Verfahrens ist 3 und pro Schritt sind 3 Auswertungen der rechten Seite der Differentialgleichung erforderlich (deshalb spricht man auch von der “(3,3)“-Regel).

iv) Für $m = 4$ existieren 4-stufige Verfahren der Ordnung 4:

- **klassisches Runge-Kutta Verfahren**;
- **Gills-Formel**;
- **3/8-Regel**;
- **Kuntzmann-Verfahren**;
- **England-Verfahren**.

Bemerkungen.

i) Verfahren fünfter Ordnung lassen sich **nicht durch $m = 5$ -stufige Verfahren realisieren**, die Ordnung 5 ist bei 6-stufigen Verfahren optimal.

| | | | | | | | | | |
|----------------------|---|---|---|---|---|---|---|---|---|
| Stufe des Verfahrens | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Bestmögliche Ordnung | 1 | 2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 |

Tabelle 32.5: Verfahren bis zur Stufe 9.

ii) Ein Verfahren hoher Ordnung erfordert eine große Stufenzahl. Eine stark anwachsende Zahl nichtlinearer Gleichungen zur Bestimmung der Koeffizienten ist zu lösen. In der Literatur sind Verfahren bis zur Stufe 9 dokumentiert (vgl. Tabelle 32.5).

32.6 Konvergenz von Einschrittverfahren (globaler Diskretisierungsfehler; Verfahrensfehler)

Schließlich ist eine **globale Fehlerabschätzung** herzuleiten, d.h. es geht um die entscheidende Frage, **mit welchem Fehler das Endresultat tatsächlich behaftet ist**. Dazu wird zur Vereinfachung der skalare Fall $N = 1$ betrachtet, das Anfangswertproblem lautet

$$y' = f(x, y), \quad y(x_0) = y_0,$$

wobei die Funktion $f: [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$ stetig und Lipschitz-stetig bzgl. der zweiten Variablen sei. Es sei weiter $x_0 \in [a, b]$, $\varphi(x)$ bezeichne die exakte Lösung des Anfangswertproblems und man setzt ($n \in \mathbb{N}$)

$$h_n := \frac{x - x_0}{n}, \quad x_i = x_0 + ih_n, \quad i = 0, 1, \dots, n.$$

Das Verfahren

$$y_{i+1} = y_i + h_n \Phi(x_i, y_i, h_n), \quad i = 0, 1, \dots, n-1, \quad (1)$$

liefert den Wert

$$y_n = y(x, h_n)$$

als Näherung für den gesuchten Wert $\varphi(x)$. Die Güte der Näherung ist durch den **globalen Diskretisierungsfehler**

$$E(x, h_n) := y(x, h_n) - \varphi(x)$$

quantifiziert. Das Einschrittverfahren heißt **konvergent**, falls

$$\lim_{n \rightarrow \infty} E(x, h_n) = 0.$$

Zur Bestimmung des globalen Diskretisierungsfehlers sei hier ohne Beweis der folgende Satz angegeben:

Satz 32.6.1

Betrachtet sei obiges Anfangswertproblem mit der exakten (eindeutigen) Lösung φ . Es sei $\Phi: G \rightarrow \mathbb{R}$ stetig,

$$G := \{(x, y, h) : a \leq x \leq b, |y - \varphi(x)| \leq \gamma, |h| \leq h_0\}$$

mit Konstanten $\gamma, h_0 > 0$. Φ sei Lipschitz-stetig auf G bzgl. y mit einer Konstanten L . Für den *lokalen Diskretisierungsfehler* $\tau(x, \varphi(x), h)$ des Verfahrens (1) gelte

$$|\tau(x, \varphi(x), h)| = |\Delta(x, \varphi(x), h) - \Phi(x, \varphi(x), h)| \leq Mh^p$$

für ein $p > 0, M > 0$. Für den *globalen Diskretisierungsfehler* folgt: Es gibt ein $\bar{h} > 0$, sodass für alle $x \in [a, b]$ und für alle $n \in \mathbb{N}$ mit

$$|h_n| = \left| \frac{x - x_0}{n} \right| \leq \bar{h}$$

gilt:

$$|E(x, h_n)| \leq |h_n|^p \frac{M}{L} \left[e^{L|x-x_0|} - 1 \right]. \quad (2)$$

Für $\gamma = \infty$ ist $\bar{h} = h_0$.

Bemerkungen.

- i) Zum Beweis benötigt man Satz ??? (stetige Abhängigkeit von den Daten).
- ii) Zu gegebenen x und $\varepsilon > 0$ kann mit (2) die Schrittweite h so bestimmt werden, dass $|E(x, h)| < \varepsilon$.
- iii) Das Problem ist: Zuvor müssen M und L bestimmt bzw. abgeschätzt werden. Das kann in konkreten Fällen schwierig sein.

Beim Euler-Verfahren ist $\Phi(x, y, h) = f(x, y)$, also (falls f partiell stetig nach y differenzierbar ist)

$$L \approx \left| \frac{\partial \Phi}{\partial y} \right| = \left| \frac{\partial f}{\partial y} \right|.$$

In Abschnitt ??? wurde weiter gezeigt: $p = 1$ und

$$M \approx \frac{1}{2} \left| f_x(x, \varphi(x)) + f_y(x, \varphi(x)) f(x, \varphi(x)) \right|,$$

d.h. die Daten können noch relativ einfach abgeschätzt werden. Beim klassischen Runge-Kutta Verfahren hingegen sind schon Ableitungen bis zur Ordnung 4 zu untersuchen.

Nach der letzten Bemerkung ist aus (2) eine “praktikalere” Abschätzung abzuleiten. Dazu werden asymptotische Entwicklungen der Form

$$y(x, h) = \varphi(x) + \text{Terme in Potenzen von } h$$

benötigt:

Satz 32.6.2

Es sei $f \in C^{k+2}([a, b] \times \mathbb{R}; \mathbb{R})$ und $y(x, h)$ die durch ein Einschrittverfahren der Ordnung $p \leq k$ gelieferte Näherung für die exakte Lösung φ des Anfangswertproblems. Dann gilt

$$y(x, h) = \varphi(x) + e_p(x)h^p + e_{p+1}(x)h^{p+1} + \dots + e_k(x)h^k + E_{k+1}(x, h)h^{k+1}$$

*mit $e_p(x_0) = 0$, $h \in H_x := \{(x - x_0)/n : n \in \mathbb{N}\}$, $x \in [a, b]$. Die Funktionen e_i sind **unabhängig von h** und für alle $x \in [a, b]$ gilt*

$$\sup_{h \in H_x} |E_{k+1}(x, h)| < \infty.$$

Als Konsequenz aus Satz 32.6.2 folgt

$$E(x, h) = y(x, h) - \varphi(x) = e_p(x)h^p + O(h^{p+1}) \quad (3)$$

und (3) soll nun weiter ausgenutzt werden.

Idee. Zusätzlich zu $y(x, h)$ berechnet man $y(x, h/2)$.

Ersetzt man in (3) h durch $h/2$, so ergibt sich

$$E(x, h/2) = e_p(x)h^p 2^{-p} + O((h/2)^{p+1}), \quad (4)$$

wird (4) von (3) abgezogen, so gelangt man zu

$$\begin{aligned} y(x, h) - y(x, h/2) &= E(x, h) - E(x, h/2) \\ &= e_p(x) \left(\frac{h}{2}\right)^p (2^p - 1) + O(h^{p+1}). \end{aligned}$$

Eingesetzt in (4) ergibt sich schließlich

$$E(x, h/2) = \frac{y(x, h) - y(x, h/2)}{2^p - 1} + O(h^{p+1}). \quad (5)$$

Beispielsweise liefert (5) für obiges Runge-Kutta Verfahren der Ordnung 4 die Fehlerabschätzung

$$E(x, h/2) = \frac{1}{15} [y(x, h) - y(x, h/2)] + O(h^5).$$

Neben der Frage zur Kontrolle des Diskretisierungsfehlers stellt sich die Frage nach der

Kontrolle des Rundungsfehlers?

Zur Beantwortung dieser Frage bezeichnen y_i die exakten Werte, \tilde{y}_i die tatsächlich berechneten, mit Rundungsfehlern behafteten Werte. es folgt

$$\begin{aligned} \tilde{y}_0 &= y_0 + \varepsilon_0 \quad (\text{Startwert}), \\ \tilde{y}_{i+1} &= \tilde{y}_i + h\Phi(x_i, \tilde{y}_i, h) + \varepsilon_{i+1} \quad (\text{Rekursion}), \end{aligned}$$

wobei stets die Annahme $|\varepsilon_i| \leq \varepsilon$, $\varepsilon > 0$ fixiert, getroffen sei, d.h. in jedem Schritt seien die Rundungsfehler durch ε kontrolliert. Setzt man

$$r_i := r(x_i, h) := \tilde{y}_i - y_i,$$

so gilt

$$\begin{aligned}
 |r_{i+1}| &= |\tilde{y}_{i+1} - y_{i+1}| \\
 &= |\tilde{y}_i + h\Phi(x_i, \tilde{y}_i, h) + \varepsilon_{i+1} - y_i - h\Phi(x_i, y_i, h)| \\
 &\leq |\varepsilon_{i+1}| + |r_i| + |h| |\Phi(x_i, \tilde{y}_i, h) - \Phi(x_i, y_i, h)| \\
 &\leq |r_i|(1 + |h|L) + |\varepsilon_{i+1}|.
 \end{aligned}$$

Dabei wurde in der letzten Ungleichung die Lipschitz-Bedingung für Φ ausgenutzt. Mit der Idealisierung $r_0 = 0$, einem Induktionsbeweis und der Ungleichung $1 + t \leq e^t$ ergibt sich

$$\begin{aligned}
 |r_{i+1}| &\leq \varepsilon \sum_{l=0}^i (1 + L|h|)^l \\
 &= \varepsilon \frac{(1 + L|h|)^{i+1} - 1}{L|h|} \\
 &\leq \frac{\varepsilon}{|h|} \frac{e^{(i+1)L|h|} - 1}{L}.
 \end{aligned}$$

Ist wie üblich $h = (x - x_0)/n$, $x \in [a, b]$, $n \in \mathbb{N}$ gegeben, so folgt für $r(x, h) := r_n$

$$\begin{aligned}
 |r(x, h)| &\leq \frac{\varepsilon}{h} \frac{1}{L} [e^{nL|h|} - 1] \\
 &= \frac{\varepsilon}{h} \frac{1}{L} [e^{L|x-x_0|} - 1].
 \end{aligned}$$

Der **Verfahrensfehler**

$$V(x, h) := E(x, h) + r(x, h),$$

also die Summe aus Diskretisierungsfehler und Rundungsfehler, ist damit (und mit Satz 32.6.1) abgeschätzt durch:

Satz 32.6.3 *Mit den obigen Bezeichnungen gilt für alle $x \in [a, b]$, für $h_n \in H_x$ mit $|h_n| \leq \bar{h}$*

$$|V(x, h)| \leq \left[M|h|^p + \frac{\varepsilon}{|h|} \right] \frac{e^{L|x-x_0|} - 1}{L}.$$

Bemerkung. Die Abschätzung ist zwar recht grob, legt aber nahe: Bei Einschrittverfahren ist die **Schrittweite nicht zu klein zu wählen**,

| h | 10^{-2} | 10^{-3} | 10^{-4} | 10^{-5} |
|-----------|------------------------|------------------------|------------------------|------------------------|
| $V(0, h)$ | $-0.276 \cdot 10^{-4}$ | $-0.229 \cdot 10^{-7}$ | $-0.478 \cdot 10^{-6}$ | $-0.227 \cdot 10^{-5}$ |

Tabelle 32.6: Zum Verfahrensfehler eines Einschrittverfahrens.

da sonst der Rundungsfehler überwiegen kann.

Beispiel. Die exakte Lösung des Anfangswertproblems

$$y' = -200xy^2, \quad y(-1) = \frac{1}{101},$$

lautet

$$\varphi(x) = \frac{1}{1 + 100x^2}.$$

Der mit dem Runge-Kutta Verfahren (12-stellig) berechnete Näherungswert $y(0, h)$ für $\varphi(0) = 1$ ist in Tabelle ??? angegeben, hier wäre also $h = 10^{-3}$ die optimale der angegebenen Schrittweiten.

Literaturverzeichnis

- [AO1] Ansorge, R., Oberle, H.J., Mathematik für Ingenieure 1. Wiley-VCH, Berlin, 2000.
- [AO2] Ansorge, R., Oberle, H.J., Mathematik für Ingenieure 2. Wiley-VCH, Weinheim, 2003.
- [Bä] Bärwolff, G., Höhere Mathematik für Naturwissenschaftler und Ingenieure. Spektrum-Elsevier, München, 2005. 2te Auflage!!!
- [BHW] Burg, K., Haf, H., Wille, F., Höhere Mathematik für Ingenieure. Teubner, Stuttgart/Leipzig/Wiesbaden, 2002.
- [Di] Dirschmid, H.J., Mathematische Grundlagen der Elektrotechnik. Vieweg, Braunschweig/Wiesbaden, 1990.
- [Fi] Fischer, G., Lineare Algebra. Vieweg, Braunschweig/Wiesbaden, 1986.
- [Fo] Forster, O., Analysis 1. Vieweg, Braunschweig, 1983.
- [HSZ] Hackbusch, W., Schwarz, H.R., Zeidler, E., Teubner-Taschenbuch der Mathematik. Teubner, Wiesbaden, 2003.
- [Hi1] Hildebrandt, S., Analysis 1. Springer, Berlin, Heidelberg, 2002.
- [Hi2] Hildebrandt, S., Analysis 2. Springer, Berlin, Heidelberg, 2003.
- [BH] Büchter, A., Henn, H.W., Elementare Stochastik.
- [He] Henze ...
- [Ke] Kemnitz, A., Mathematik zum Studienbeginn. Vieweg, Braunschweig/Wiesbaden, 2002.
- [Kr1] Krämer, W., So überzeugt man mit Statistik. Campus, Frankfurt, 1994.

- [Kr2] Krämer, W., So lügt man mit Statistik. Campus, Frankfurt, 1998.
- [Kr3] Krämer, W., Statistik verstehen. Eine Gebrauchsanweisung. Campus, Frankfurt.
- [Mü] [http ...](#)