

February 2024

Theoretische Physik 1: Mathematische Ergänzungen zur Vorlesung

Philipp Schicho^{a,*}

^a*Institute for Theoretical Physics, Goethe Universität Frankfurt, 60438 Frankfurt, Germany*

Abstrakt

Dieses Manuskript ist Teil der Vorlesung “Mathematische Ergänzungen zur Theoretischen Physik 1”. Das Ziel dieser Vorlesung sind Vertiefung und Anwendung der mathematischen Methoden, die in der Vorlesung “Theoretische Physik” benötigt werden.



*schicho@itp.uni-frankfurt.de

Contents

1	Einführung	3
1.1	Elementare Funktionen	3
1.2	Komplexe Zahlen und Funktionen	6
2	Analysis in einer Dimension	11
2.1	Grenzwert und Stetigkeit	11
2.2	Differentialrechnung für Funktionen einer reellen Veränderlichen	15
2.2.1	Ableitungsregeln	17
2.3	Hauptsätze der Differentialrechnung	21
2.4	Begriffe der Differentialrechnung	24
2.5	Taylor Entwicklung	26
2.6	Über die Konvergenz von Reihen	29
2.7	Hauptsätze der Integralrechnung	31
2.8	Begriffe der Integralrechnung	34
2.9	Integrations Methoden	36
2.10	Gewöhnliche Differentialgleichungen	40
2.11	Allgemeine Lösung linearer Differentialgleichungen erster Ordnung	42
2.12	Lineare Differentialgleichungen zweiter Ordnung	47
2.13	Lineare Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten	52
3	Lineare Algebra	55
3.1	Vektorräume, lineare Abbildungen	55
3.2	Skalarprodukt	58
3.3	Transponierung, Spur, Determinante	61

Vorbemerkungen

Dr. Philipp Schicho, Phys 02.131

Skript website: pschicho.github.io/lectures/metp1.pdf

Das vorliegende Manuskript basiert auf mehreren Quellen und Skripten:

References

- [1] H. van Hees, *Mathematische Ergänzungen zur Theoretischen Physik 1*.
- [2] C. B. Lang and N. Pucker, *Mathematische Methoden in der Physik*. Springer Berlin Heidelberg, Berlin, Heidelberg, 3 ed., 2016.
- [3] R. Adams and C. Essex, *Calculus: A Complete Course*. Pearson, 10 ed., 2022.
- [4] K. Hefft, *Mathematischer Vorkurs zum Studium der Physik*. Springer Spektrum, Berlin, 2 ed., 2018.
- [5] S. Großmann, *Mathematischer Einführungskurs für die Physik*. Springer Verlag, Berlin, Heidelberg, 10 ed., 2012.
- [6] M. Ganster, *Differential- und Integralrechnung*. 2011.
- [7] M. Laine, *Mathematische Methoden der Physik I*. 2013.

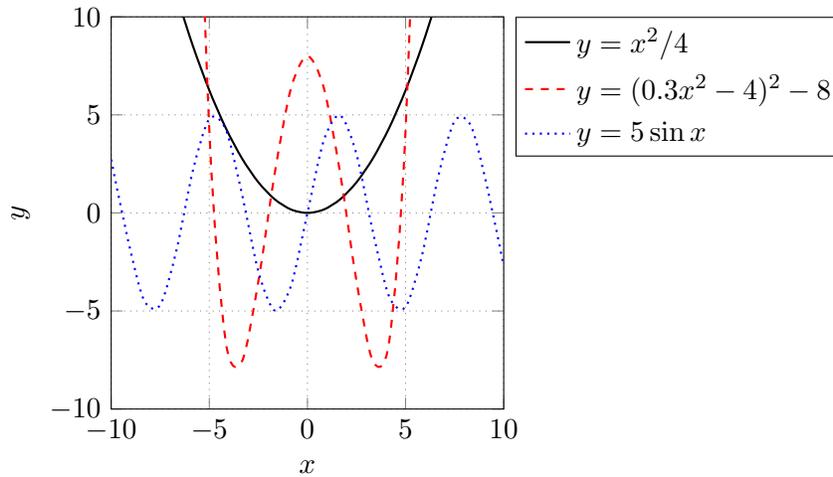


Figure 1: Elementare Beispielfunktionen.

1. Einführung

1.1. Elementare Funktionen

Referenzen: [1] Anhang B.

Um physikalische Systeme zu beschreiben benutzen und analysieren wir Funktionen. Dazu betrachten wir vorerst Funktionen einer **reellen**¹ Variablen,

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto f(x) \quad (\text{zum Beispiel } t \mapsto x(t)), \quad (1.1)$$

$$f : D \rightarrow Z, \quad x \mapsto f(x). \quad (1.2)$$

Diese sind eindeutige Abbildungen der reellen Zahlen \mathbb{R} oder einer Definitionsmenge ($D \subseteq \mathbb{R}$) in die reellen Zahlen \mathbb{R} oder einer Zielmenge ($Z \subseteq \mathbb{R}$). Jeder Zahl $x \in D$ wird eine Zahl $y \in \mathbb{R}$ zugeordnet, nämlich $y = f(x)$. Die Menge der Punkte in \mathbb{R}^2 mit den Koordinaten $(x, f(x))$ heisst **Graph der Funktion** f ; siehe Abb. 1. Folgende Definitionen sind relevant für reelle Funktionen:

Polynom. Die (un)endliche Summe von m Monomen heisst Polynom:

$$f(x) = \sum_{n=0}^m a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m, \quad (1.3)$$

¹Struktur von \mathbb{R} :

\exists zwei Verknüpfungen: $a, b \in \mathbb{R} \Rightarrow a + b \in \mathbb{R}, a, b \in \mathbb{R} \Rightarrow a \times b \in \mathbb{R}$.

\exists Nullelement: $a + 0 = a \quad \forall a \in \mathbb{R}$.

\exists Einselement: $a \times 1 = a \quad \forall a \in \mathbb{R}$

\exists Inversion: $\forall a \in \mathbb{R} \exists -a \in \mathbb{R}$ mit $a + (-a) = 0, \forall a \neq 0 \exists \frac{1}{a} \in \mathbb{R}$ mit $a \times \frac{1}{a} = 1$.

Andere Operationen: $a - b := a + (-b), \frac{a}{b} := a \times \frac{1}{b}$.

$$f(x) = \sum_{n=0}^m a_n (x - x_0)^n, \quad x_0 \in \mathbb{R}. \quad (1.4)$$

Komposition. Die Hintereinanderausführung zweier (oder mehrerer) Funktionen folgt:

$$(g \circ f)(x) := g(f(x)). \quad (1.5)$$

Beispiel 1.

$$\begin{aligned} f(x) &= x - x_0, & g(x) &= b_n x^n, \\ (g \circ f)(x) &= b_n (x - x_0)^n, & (f \circ g)(x) &= b_n x^n - x_0. \end{aligned} \quad (1.6)$$

Umkehrfunktion. Falls die Funktion $f(x)$ bijektiv ist, also die Gleichung $y = f(x)$ eindeutig als $x = g(y)$ gelöst werden kann, nennen wir g die Umkehrfunktion zu f . Bezeichnung üblicherweise: f^{-1} (n.b. nicht $\frac{1}{f}$!). Konkret bedeutet das:

$$\begin{aligned} y &= f(x) = f(f^{-1}(y)) = (f \circ f^{-1})(y) \quad \forall y \\ &\Rightarrow f \circ f^{-1} = \mathbb{1} := \text{Identität}, \end{aligned} \quad (1.7)$$

$$\begin{aligned} x &= f^{-1}(y) = f^{-1}(f(x)) = (f^{-1} \circ f)(x) \quad \forall x \\ &\Rightarrow f^{-1} \circ f = \mathbb{1}. \end{aligned} \quad (1.8)$$

Beispiel 2.

$$\begin{aligned} f : \mathbb{R}_+ &\rightarrow \mathbb{R}_+, \quad x \mapsto 2x^2, \\ f^{-1} : \mathbb{R}_+ &\rightarrow \mathbb{R}_+, \quad x \mapsto \sqrt{\frac{x}{2}}. \end{aligned} \quad (1.9)$$

Der Definitionsbereich der Funktion ist $D(f) = \mathbb{R}_+$, da diese bijektiv sein muss für die Existenz einer Umkehrfunktion.

Beispiel 3.

$$f(x) = x^m \Rightarrow \exists f^{-1}(x) =: \sqrt[m]{x} \stackrel{!}{=} x^{\frac{1}{m}} \quad \text{für } x > 0. \quad (1.10)$$

Die Definition einer allgemeinen Potenz (z.B. x^μ , $\mu \in \mathbb{R}$) folgt später.

Reihen. Weitere elementare Funktionen werden durch unendliche Reihen definiert. Eine spezielle Funktion die so definiert werden kann, ist die Exponentialfunktion:

$$f(x) = \sum_{n=0}^m a_n x^n, \quad \text{mit } a_n := \frac{1}{n!} := \frac{1}{n(n-1)\dots 1}, \quad \text{und } m \rightarrow \infty,$$

$$\exp(x) := \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \quad e := \exp(1) \approx 2.718\dots \quad (1.11)$$

Diese Darstellung konvergiert erstaunlicherweise für $\forall x$, z.B.:

$$\begin{aligned} \exp(-10) &= 1 - 10 + \frac{1}{2}100 - \frac{1}{6}1000 + \frac{1}{24}10000 - \dots \\ &= 1 - 10 + 50 - 166.7 + 416.7 - \dots \\ &\approx 0.0000453999. \end{aligned} \quad (1.12)$$

Die Umkehrfunktion von $\exp(x)$ wird mit $\ln(x)$ bezeichnet:

$$\begin{aligned} \exp(\ln(x)) &= x \quad \forall x \in \mathbb{R}_+, \\ \ln(\exp(x)) &= x \quad \forall x \in \mathbb{R} \Rightarrow \ln e = 1, \end{aligned} \quad (1.13)$$

$$\ln(x) = \ln(1 + x - 1) = \ln(1 - [1 - x]) = - \sum_{n=1}^{\infty} \frac{1}{n} (1 - x)^n, \quad (1.14)$$

wobei die Reihendarstellung für $0 < x \leq 2$ konvergiert.

Weitere elementare Funktionen sind:

- Allgemeine Potenz und Umkehrfunktion:

$$a^x := \exp(x \ln a) = \sum_{n=0}^{\infty} \frac{\ln^n a}{n!} x^n, \quad \text{für } a \in \mathbb{R}_+, \quad (1.15)$$

$$\log_a x := \frac{\ln x}{\ln a}. \quad (1.16)$$

Beweis: $y = a^x$, $\ln y = \ln \exp(x \ln a) = x \ln a$. Es gilt: $e^x = \exp(x)$, denn $\ln e = 1$.

- Hyperbolische Funktionen

$$\text{Sinus Hyperbolicus: } f : \mathbb{R} \rightarrow \mathbb{R}, \quad \sinh(x) := \frac{e^x - e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1}, \quad (1.17)$$

$$\text{Cosinus Hyperbolicus: } f : \mathbb{R} \rightarrow \mathbb{R}, \quad \cosh(x) := \frac{e^x + e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n}, \quad (1.18)$$

$$\text{Tangens Hyperbolicus: } f : \mathbb{R} \rightarrow [-1, 1], \quad \tanh(x) := \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (1.19)$$

Umkehrfunktionen (nur für bestimmte x): $\operatorname{arsinh}(x)$, $\operatorname{arcosh}(x)$, $\operatorname{artanh}(x)$.

- Trigonometrische Funktionen

$$\text{Sinus: } f : \mathbb{R} \rightarrow [-1, 1], \quad \sin(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad (1.20)$$

$$\text{Cosinus: } f : \mathbb{R} \rightarrow [-1, 1], \quad \cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad (1.21)$$

$$\text{Tangens: } f : \mathbb{R} \setminus \left\{ \left(\mathbb{Z} + \frac{1}{2} \right) \pi \right\} \rightarrow \mathbb{R}, \quad \tan(x) := \frac{\sin(x)}{\cos(x)}. \quad (1.22)$$

Umkehrfunktionen (nur für bestimmte x): $\arcsin(x)$, $\arccos(x)$, $\arctan(x)$.

1.2. Komplexe Zahlen und Funktionen

Referenzen: [1] Abschn. (2.1)–(2.2), [2] Abschn. 4.

Die Erweiterung von den reellen zu den **komplexen Zahlen** ist durch die Forderung nach der Lösbarkeit von **Polynomgleichungen** motiviert. Bei der Lösung quadratischer Gleichungen der Form

$$x^2 + px + q = 0, \quad (1.23)$$

stossen wir auf das Problem, dass für $x \in \mathbb{R}$ stets $x^2 \geq 0$ gilt, d.h. im Rahmen der reellen Zahlen können wir keine Quadratwurzeln aus negativen Zahlen ziehen. Die Lösungsstrategie für Gl. (1.23) besteht darin, eine **quadratische Ergänzung** auszuführen. Offenbar gilt nämlich

$$x^2 + px + q = \left(x + \frac{p}{2} \right)^2 - \frac{p^2}{4} + q. \quad (1.24)$$

Die Gl. (1.23) ist also äquivalent zu der Gleichung

$$\left(x + \frac{p}{2} \right)^2 = \frac{p^2}{4} - q. \quad (1.25)$$

Wollen wir diese Gleichung nach x auflösen, müssen wir die Wurzel aus der rechten Seite ziehen können. Im Bereich der reellen Zahlen ist das offensichtlich nur möglich, wenn $p^2/4 - q \geq 0$ ist. Dann besitzt die Gleichung entweder eine doppelte (falls $p^2/4 - q = 0$) oder zwei unterschiedliche (falls $p^2/4 - q > 0$) Lösungen. Die Allgemeine Lösung dafür lautet

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \quad (1.26)$$

Wir erweitern nun die reellen Zahlen durch eine neue, zunächst rein symbolisch zu verstehende, “Zahl” i , die **imaginäre Einheit**, für die

$$i^2 = -1, \quad (1.27)$$

gelten soll. Dann hätte für $a > 0$ die Gleichung $x^2 = -a$ die beiden Lösungen $x = \pm i\sqrt{a}$, wobei wir voraussetzen, dass die **komplexen Zahlen**, die allgemein von der Form

$$z = x + iy, \quad x, y \in \mathbb{R}, \quad (1.28)$$

sein sollen, die gewöhnlichen Rechenregeln wie für reelle Zahlen gelten, also die sogenannten Axiome eines **Zahlenkörpers** erfüllen. Dabei soll eine komplexe Zahl definitionsgemäss durch ihren **Realteil** x und **Imaginärteil** y eindeutig bestimmt sein. Wir schreiben

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y. \quad (1.29)$$

Nehmen wir dies an, so folgt für die beiden Verknüpfungen zweier komplexer Zahlen

$$\textbf{Summe:} \quad z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2), \quad (1.30)$$

$$\textbf{Produkt:} \quad z_1 \times z_2 = (x_1 + iy_1) \times (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2), \quad (1.31)$$

wobei wir mehrfach das Assoziativ-, Kommutativ-, und Distributivgesetz verwendet haben und wir i wie eine gewöhnliche Variable behandelt haben. Das Produkt folgt ebenso durch formales Ausmultiplizieren, wobei wir im zweiten Term des Realteils die definierende Eigenschaft (1.27) der imaginären Einheit benutzt haben.

Es existieren auch “neutrale Elemente” ($z + 0 = z$, $z \times 1 = z$) und “inverse Elemente” ($z + (-z) = 0$, $z \times z^{-1} = 1$). bezüglich beider Verknüpfungen. Dies macht die Menge aller komplexen Zahlen, \mathbb{C} , zu einem **Körper**. Weiters können wir auch noch definieren:

$$\textbf{Subtraktion:} \quad z_1 - z_2 = z_1 + (-z_2) = (x_1 - x_2) + i(y_1 - y_2), \quad (1.32)$$

$$\begin{aligned} \textbf{Division:} \quad z_1/z_2 &= z_1 \times z_2^{-1} = \frac{z_1 z_2^*}{z_2 z_2^*} \\ &= \frac{(x_1 + iy_1)(x_2 - iy_2)}{x_2^2 + y_2^2} = \left(\frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} \right) + i \left(\frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} \right). \end{aligned} \quad (1.33)$$

Hier haben wir uns für die Definition des inversen z_2^{-1} , der Operation der komplexen Konjugation bedient (z^* oder \bar{z})

$$z^* = x - iy, \quad (1.34)$$

wodurch das Produkt einer komplexen Zahl mit ihrem konjugiert Komplexen

$$zz^* = (x + iy)(x - iy) = x^2 + y^2 + i(xy - yx) \in \mathbb{R}_+. \quad (1.35)$$

Daher folgt

$$z^{-1} = \frac{z^*}{zz^*} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}. \quad (1.36)$$

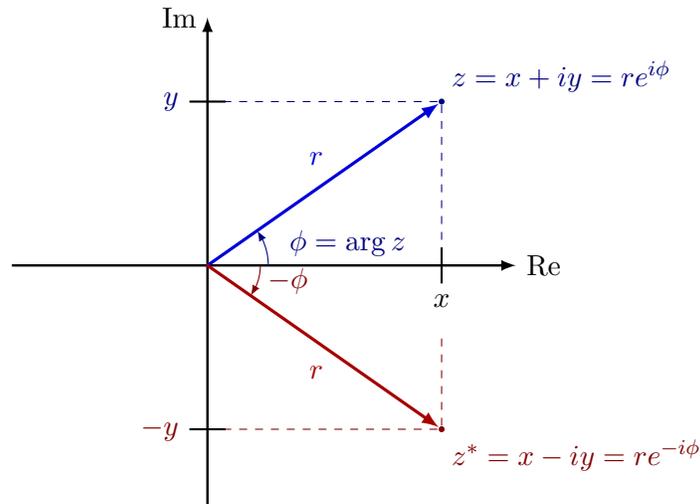


Figure 2: Die Gaußsche Zahlenebene und Polarform einer komplexen Zahl, wobei gilt $r = |z| = \sqrt{zz^*}$.

Den **Betrag** bzw. Modul, das **Argument**, und die **polare Form** einer komplexen Zahl z definieren wir als

$$|z| = \text{mod } z := \sqrt{zz^*} = \sqrt{x^2 + y^2}, \quad (1.37)$$

$$\arg z := \phi = \arctan\left(\frac{y}{x}\right), \quad (1.38)$$

$$z = x + iy = |z| \left(\frac{x}{\sqrt{x^2 + y^2}} + i \frac{y}{\sqrt{x^2 + y^2}} \right) = |z| \left(\cos \phi + i \sin \phi \right). \quad (1.39)$$

Die reellen Zahlen können wir geometrisch durch eine Zahlengerade veranschaulichen. Entsprechend kann man die komplexen Zahlen geometrisch interpretieren, wenn man das Zahlenpaar $(x, y) = (\text{Re } z, \text{Im } z)$ als Komponenten bzgl. eines kartesischen Koordinatensystems in der Euklidischen Ebene interpretiert. Dies ist die **Gaußsche Zahlenebene**. Es ist klar, dass $|z|$ geometrisch die Länge des entsprechenden z repräsentierenden Ortsvektors in der Gaußschen Zahlenebene ist (s. Abb. 2).

Offensichtlich ist $(z^*)^* = z$ für alle $z \in \mathbb{C}$. Weiter ist $\mathbb{R} \subset \mathbb{C}$, denn die komplexen Zahlen mit verschwindendem Imaginärteil sind umkehrbar eindeutig auf \mathbb{R} abbildbar. Offenbar ist $z \in \mathbb{R}$ genau dann, wenn $\text{Im } z = 0$, was zugleich $z^* = z$ impliziert. Wir haben weiter

$$\text{Re } z = \frac{z + z^*}{2}, \quad \text{Im } z = \frac{z - z^*}{2i}. \quad (1.40)$$

Die komplexe Konjugation ist auch additiv und multiplikativ

$$(z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 z_2)^* = z_1^* z_2^*. \quad (1.41)$$

Weil die beiden Verknüpfungen von Addition und Multiplikation zur Verfügung stehen, können wir die beiden kombinieren und **komplexe Polynome** definieren:

$$P_m(z) := \sum_{k=0}^m a_k z^k = a_0 + a_1 z + \dots + a_m z^m. \quad (1.42)$$

Die Koeffizienten könnten reell ($a_k \in \mathbb{R}$) oder komplex ($a_k \in \mathbb{C}$) sein. Wiederum existieren

- Komposition $[(g \circ f)(z) := g(f(z))]$
- Umkehrfunktion $[f \circ f^{-1} = f^{-1} \circ f = \mathbb{1}]$.
- Unendliche Reihen $[P_n(z)$ mit $n \rightarrow \infty]$, die zu weiteren Funktionen führen.

Wir beginnen mit der Exponentialfunktion und übernehmen die entsprechende Potenzreihe einfach von der entsprechenden reellen Funktion als Definition für die Exponentialfunktion (1.11) im Komplexen

$$\exp(z) := \sum_{k=0}^{\infty} \frac{1}{k!} z^k, \quad e := \exp(1). \quad (1.43)$$

Auch sie konvergiert für alle $z \in \mathbb{C}$.

Weiter benötigen wir noch die trigonometrischen Funktionen. Auch ihre Potenzreihen übernehmen wir aus dem Reellen, d.h. mit (1.20) bzw. (1.21) folgt (*nachrechnen!*)

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \quad (1.44)$$

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}. \quad (1.45)$$

Berechnen wir nun

$$\begin{aligned} \exp(iz) &= 1 + iz + \frac{(iz)^2}{2!} + \frac{(iz)^3}{3!} + \dots \\ &= \left(1 - \frac{z^2}{2!} + \dots\right) + i \left(z - \frac{z^3}{3!} + \dots\right) \\ &= \left(\sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}\right) + i \left(\sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}\right). \end{aligned} \quad (1.46)$$

Dabei haben wir die Reihe so umgeordnet, dass wir in einem Term den Faktor i ausklammern konnten. Das ist bei Potenzreihen erlaubt, da sie in jedem kompakten Bereich der komplexen Ebene absolut konvergiert.

Vergleichen wir nun die Reihen in den Klammern der Gl. (1.46) mit 1.44 und (1.45), erhält man die **Eulersche Formel** (1707–1783)

$$\exp(iz) = \cos z + i \sin z. \quad (1.47)$$

Für die Polardarstellung der komplexen Zahl (1.39) folgt damit

$$z = |z| \exp(i\varphi). \quad (1.48)$$

Für die Exponentialfunktion gilt auch im Komplexen die Formel

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2). \quad (1.49)$$

Beweis. Dies kann man mit Hilfe der Reihe (1.43), dem Binomischen Lehrsatz und dem Cauchy Produkt beweisen,

$$\begin{aligned} \exp(z_1 + z_2) &= \sum_{n=0}^{\infty} \frac{(z_1 + z_2)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^n \binom{n}{k} \frac{z_1^k z_2^{n-k}}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{z_1^k z_2^{n-k}}{k!(n-k)!} = \sum_{k=0}^{\infty} \frac{(z_1)^k}{k!} \sum_{n=0}^{\infty} \frac{(z_2)^n}{n!} = \exp(z_1) \exp(z_2). \end{aligned}$$

□

Diese Eigenschaft erleichtert die Rechnung mit trigonometrischen Funktionen erheblich. Es folgt z.B. genau wie (1.47) auch die Gleichung

$$\exp(-iz) = \cos z - i \sin z. \quad (1.50)$$

Wir haben damit

$$\cos z = \frac{1}{2}[\exp(iz) + \exp(-iz)], \quad \sin z = \frac{1}{2i}[\exp(iz) - \exp(-iz)]. \quad (1.51)$$

Dies erinnert an die Definition der Hyperbelfunktionen

$$\cosh z = \frac{1}{2}[\exp(z) + \exp(-z)], \quad \sinh z = \frac{1}{2}[\exp(z) - \exp(-z)]. \quad (1.52)$$

Vergleicht man (1.51) mit diesen Definitionen folgt sofort, dass

$$\cosh(iz) = \cos z, \quad \sinh(iz) = i \sin z \quad (1.53)$$

gilt. Die trigonometrischen und Hyperbelfunktionen sind im Komplexen also bis auf Konstanten im wesentlichen die gleichen Funktionen, und beide sind durch die Exponentialfunktion definiert.

Genauso folgt aus (1.51)

$$\cos(iz) = \cosh z, \quad \sin(iz) = i \sinh z. \quad (1.54)$$

Als Anwendungsbeispiel leiten wir noch die Additionstheoreme für die trigonometrischen Funktionen für reelle Argumente aus (1.49) ab. Es gilt nämlich einerseits wegen der Eulerschen Formel (1.47) und (1.48)

$$\exp[i(\varphi_1 + \varphi_2)] = \cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2), \quad (1.55)$$

$$\begin{aligned} \exp[i(\varphi_1 + \varphi_2)] &= \exp(i\varphi_1) \exp(i\varphi_2) = (\cos \varphi_1 + i \sin \varphi_1)(\cos \varphi_2 + i \sin \varphi_2) \\ &= (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i(\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2). \end{aligned} \quad (1.56)$$

Vergleicht man nun Real- und Imaginärteil von (1.55) und (1.56), folgen die bekannten Additionstheoreme

$$\cos(\varphi_1 + \varphi_2) = \cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2, \quad (1.57)$$

$$\sin(\varphi_1 + \varphi_2) = \sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2. \quad (1.58)$$

Es ist leicht zu zeigen, dass diese Additionstheoreme auch allgemein für beliebige komplexe Argumente gelten (*Übung*).

2. Analysis in einer Dimension

2.1. Grenzwert und Stetigkeit

Referenzen: [2] A.5, [3] Abschn. 1.5.

Definition 2.1. $f(x)$ hat bei $x = x_0$ den Grenzwert f_0 wenn gilt, dass

$$\forall \epsilon > 0 \exists \delta_\epsilon > 0 \quad : \quad |f(x) - f_0| < \epsilon \quad \text{falls} \quad |x - x_0| < \delta_\epsilon. \quad (2.1)$$

Bezeichnung: $f(x) \rightarrow f_0$ für $x \rightarrow x_0$, oder $\lim_{x \rightarrow x_0} f(x) = f_0$.

Der Grenzwert kann auch von links oder rechts genommen werden, z.B.: $f(x) \rightarrow f_0$ für $x \rightarrow x_0^+$, oder $\lim_{x \rightarrow x_0^+} f(x) = f_0$ bedeutet, dass nur $0 < x - x_0 < \delta_\epsilon$ betrachtet werden. Für **unendliche** Grenzwerte folgt dann $\lim_{x \rightarrow \infty} f(x) = f_0$, so dass $\forall \epsilon > 0 \exists \delta_\epsilon > 0 : |f(x) - f_0| < \epsilon$ falls $x > \delta_\epsilon$.

Beispiel 4. Die Funktion $f(x) = \frac{\sin x}{x}$ scheint an $x_0 = 0$ nicht definiert und $f(0) = \frac{0}{0}$. Aus der Definition aus Gl. (1.20) wissen wir aber

$$\sin x = x - \frac{x^3}{6} + \dots, \quad (2.2)$$

$$\frac{\sin x}{x} = 1 - \underbrace{\frac{x^2}{6} + \dots}_{\rightarrow 0 \text{ für } x \rightarrow 0^+}. \quad (2.3)$$

Daher gilt $\lim_{x \rightarrow 0^+} \frac{\sin x}{x} = 1$.

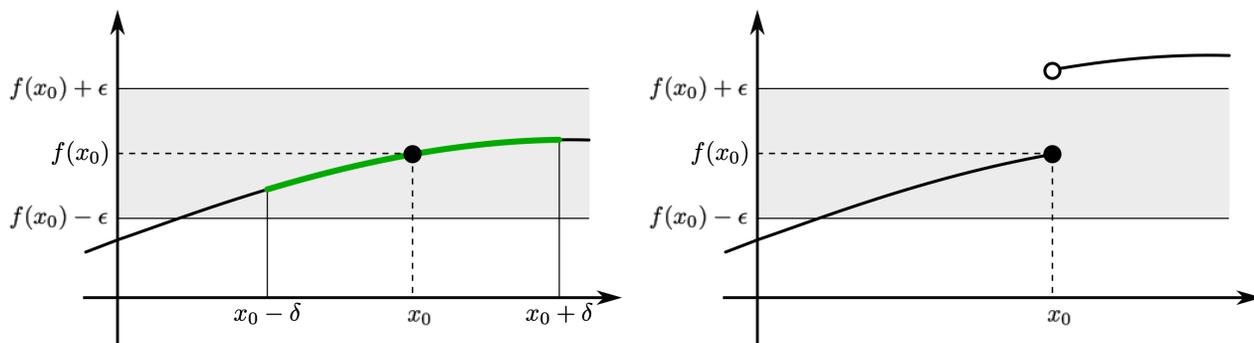


Figure 3: Beispiel einer stetigen (links) und unstetigen Funktionen (rechts), nach Definition 2.2 (cf. [4]).

Weiters können wir folgende Grenzwertsätze (ohne Beweis) benutzen:

$$\lim_{x \rightarrow x_0} cf(x) = c \lim_{x \rightarrow x_0} f(x), \quad (2.4)$$

$$\lim_{x \rightarrow x_0} [f(x) + g(x)] = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x), \quad (2.5)$$

$$\lim_{x \rightarrow x_0} [f(x) \times g(x)] = \left[\lim_{x \rightarrow x_0} f(x) \right] \times \left[\lim_{x \rightarrow x_0} g(x) \right], \quad (2.6)$$

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}, \quad \text{falls } \lim_{x \rightarrow x_0} g(x) \neq 0. \quad (2.7)$$

In den Beispielen in Abb. 1 sind alle Funktionen auf ganz \mathbb{R} definiert. Ausserdem weist ihr Graph keinerlei Unterbrechungen oder ähnliche Anomalien auf. Die mathematische Voraussetzung dafür ist, dass eine Funktion stetig ist. Anschaulich bedeutet die **Stetigkeit einer Funktion** an der Stelle x_0 , dass sich der Funktionswert um x_0 nur wenigändert, wenn man sie in einer kleinen Umgebung um x_0 betrachtet. Formal lässt sich dies so definieren:

Definition 2.2. Eine Funktion $f : D \rightarrow \mathbb{R}$ heisst stetig an der Stelle $x_0 \in D$, wenn es zu jedem $\epsilon > 0$ ein zugehöriges $\delta_\epsilon > 0$ gibt (δ_ϵ hängt von ϵ und x_0 ab!), sodass für alle $x \in D$ mit $|x - x_0| < \delta_\epsilon$ stets $|f(x) - f(x_0)| < \epsilon$ gilt.

Mit Def. 2.1 ist eine Funktion dann stetig an x_0 , wenn ihr Grenzwert dort existiert

$$\lim_{x \rightarrow x_0} f(x) = f(x_0). \quad (2.8)$$

Wenn der Grenzwert nicht existiert, oder existiert aber nicht gegen $f(x_0)$ geht, dann ist die Funktion **unstetig** an x_0 .

Informelle Definition: Wird ein beliebiges ϵ -Intervall um $f(x_0)$ vorgegeben, dann existiert dazu ein δ_ϵ -Intervall um x_0 , welches zur Gänze in das ϵ -Intervall um $f(x_0)$ abgebildet wird.

In anderen Worten: Die Funktionswerte für hinreichend kleine Umgebungen um x_0 weichen beliebig

wenig von dem Funktionswert $f(x_0)$ bei x_0 ab. Oder man die Funktion ohne Absetzen des Stifts skizzieren kann.

Beispiel 5.

$$f(x) = \begin{cases} -\frac{1}{2}, & x \leq 1, \\ +\frac{1}{2}, & x > 1. \end{cases} \quad (2.9)$$

Die Funktion ist stetig von links aber nicht von rechts: $\lim_{x \rightarrow 1^+} f(x) = +\frac{1}{2}$, $\lim_{x \rightarrow 1^-} f(x) = -\frac{1}{2}$, $f(1) = -\frac{1}{2}$.

Beispiel 6. Die Funktion $f(x) = \frac{x^2-4}{x-2}$ ist an der Stelle $x_0 = 2$ nicht definiert. Offenbar gilt aber $\lim_{x \rightarrow 2^-} f(x) = \lim_{x \rightarrow 2^+} f(x) = 4$. Dies ist auch ersichtlich da $f(x) = \frac{(x-2)(x+2)}{(x-2)} = (x+2)$. Obwohl die Funktion nicht definiert ist an x_0 , können wir damit an der Stelle $x_0 = 2$ den Funktionswert $f(2) = 4$ setzen und erhalten eine in $x_0 = 2$ stetige Funktion.

Satz 2.3. Sind zwei Funktionen f und g stetig in x_0 , dann sind auch deren folgende Verknüpfungen stetig:

(i) cf , $f + g$, $f \times g$, $\frac{f}{g}$ vorausgesetzt $g(x_0) \neq 0$,

(ii) die Komposition $g \circ f$,

(iii) die Umkehrfunktion f^{-1} .

Beweis. Für zwei in x_0 stetige Funktionen f und g ist auch die Funktion $f + g$ stetig. Voraussetzungs-gemäss gibt es zu jedem $\epsilon > 0$ ein

$$\begin{aligned} \delta_1 > 0, & \quad \text{so dass } |f(x) - f(x_0)| < \epsilon/2 & \quad \forall x \in D \quad \text{mit} \quad |x - x_0| < \delta_1, \\ \delta_2 > 0, & \quad \text{so dass } |g(x) - g(x_0)| < \epsilon/2 & \quad \forall x \in D \quad \text{mit} \quad |x - x_0| < \delta_2. \end{aligned}$$

Setzen wir nun $\delta = \min(\delta_1, \delta_2)$, dann gilt wegen der Dreiecksungleichung² für alle $|x - x_0| < \delta$

$$\begin{aligned} |[f(x) + g(x)] - [f(x_0) + g(x_0)]| &= |[f(x) - f(x_0)] + [g(x) - g(x_0)]| \\ &\leq |f(x) - f(x_0)| + |g(x) - g(x_0)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned} \quad (2.10)$$

d.h. zu jedem $\epsilon > 0$ können wir ein $\delta > 0$ finden, so dass für alle $x \in D$ mit $|x - x_0| < \delta$ stets

$$|[f(x) + g(x)] - [f(x_0) + g(x_0)]| < \epsilon \quad (2.11)$$

gilt. Das bedeutet aber gemäss der Definition, dass die Funktion $f + g$ stetig ist. \square

² Dreiecksungleichung (ohne Beweis): $|a + b| \leq |a| + |b|$ mit $a, b \in \mathbb{R}$.

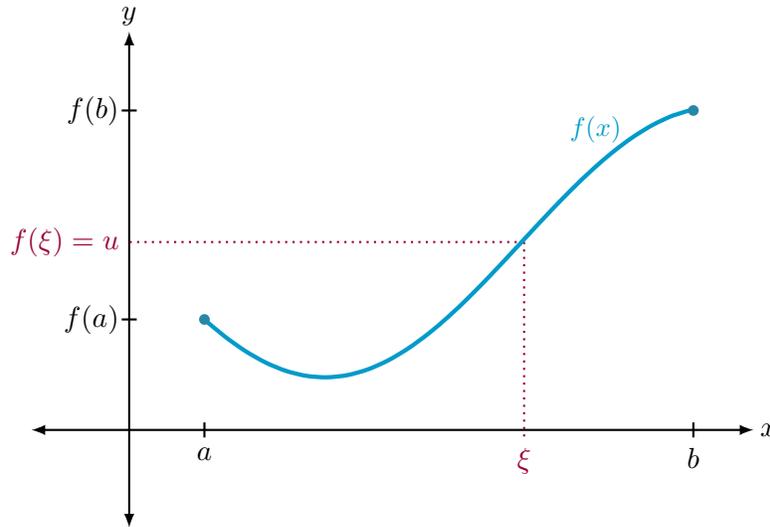


Figure 4: Zwischenwertsatz 2.5 am Beispiel einer Funktion [4].

Offensichtlich sind eine konstante Funktion $f(x) = A = \text{const}$ sowie die Funktion $f(x) = x$ in allen $x \in \mathbb{R}$ definiert und stetig. Aus den oben angegebenen Sätzen folgt dann, dass auch jedes Polynom

$$f(x) = P_m(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m = \sum_{n=0}^m a_nx^n, \quad (2.12)$$

stetig ist.

Satz 2.4 (Nullstellensatz von Bolzano (1781–1848)). *Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei stetig auf dem Intervall $[a, b]$ und es gelte $f(a) \cdot f(b) < 0$ (d.h. $f(a)$ und $f(b)$ haben unterschiedliches Vorzeichen). Dann existiert ein $\xi \in (a, b)$ mit $f(\xi) = 0$, i.e. f hat in (a, b) mindestens eine Nullstelle.*

Beweis. oBdA sei $f(a) < 0$, $f(b) > 0$. Wir setzen $I_0 = [a, b]$. Durch Halbierung von I_0 erhalten wir zwei Teilintervalle $[a, \frac{a+b}{2}]$, $[\frac{a+b}{2}, b]$, von denen bei einem, etwa $I_1 = [a_1, b_1]$, gilt, dass $f(a_1) < 0$, $f(b_1) > 0$. Durch fortwährende Intervallhalbierung erhalten wir eine Folge $I_n = [a_n, b_n]$ von Intervallen mit $I_{n+1} \subseteq I_n$, $f(a_n) < 0$, $f(b_n) > 0$ und $|b_n - a_n| = \frac{b-a}{2^n}$. Dies ist eine Nullfolge: $\lim_{n \rightarrow \infty} |b_n - a_n| = 0$.

Die Folge (a_n) ist monoton wachsend und nach oben beschränkt, (durch b), die Folge (b_n) ist monoton fallend und nach unten beschränkt, (durch a). Also sind beide Folgen konvergent. Da aber $\lim_{n \rightarrow \infty} |b_n - a_n| = 0$, müssen beide Grenzwerte gleich sein, i.e. $\exists \xi \in \mathbb{R}$ mit $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \xi$. Da beide Folgen gegen denselben Grenzwert ξ konvergieren und da weiter f voraussetzungsgemäss im ganzen Intervall $[a, b]$ stetig ist, muss folglich $f(\xi) = 0$ sein. \square

Satz 2.5 (Zwischenwertsatz). *Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Weiter sei $f(a) < f(b)$. Dann gibt es zu jedem $u \in [f(a), f(b)]$ ein $\xi \in [a, b]$, so dass $f(\xi) = u$ i.e. $f(a) < u < f(b)$ ist (entsprechendes gilt auch falls $f(a) > f(b)$ ist). Das heisst jeder Wert zwischen $f(a)$ und $f(b)$ wird angenommen.*

Beweis. Dies ist ein Sonderfall des Nullstellensatzes. Sei $g(x) = f(x) - u$. Dann ist g stetig auf $[a, b]$ mit $g(a) < 0$ und $g(b) > 0$. Nach dem Nullstellensatz existiert ein $\xi \in (a, b)$ mit $g(\xi) = 0$, i.e. $f(\xi) = u$. \square

Satz 2.6 (Satz vom Maximum und Minimum). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann nimmt f an wenigstens einer Stelle $\xi \in [a, b]$ ein Maximum (Minimum) an. Dabei ist das Maximum durch $\sup_{x \in [a, b]} f(x)$ (das Minimum durch $\inf_{x \in [a, b]} f(x)$) definiert.

Beweis. Wir beweisen den Satz für das Maximum. Der Beweis für das Minimum folgt analog. Da die Funktion im gesamten (abgeschlossenen!) Intervall definiert ist, ist ihr Bildbereich $f([a, b])$ eine beschränkte Menge und besitzt demnach ein Supremum, d.h. es gilt $f(x) \leq M$ für alle $x \in [a, b]$ mit $M = \sup_{x \in [a, b]} f(x)$. Dadurch existiert dann eine Zahlenfolge $x_n \in [a, b]$, so dass $\lim_{n \rightarrow \infty} f(x_n) = M$.

2.2. Differentialrechnung für Funktionen einer reellen Veränderlichen

In der Physik wollen wir untersuchen, wie sich Funktionen in kleinen Umgebungen eines beliebigen Punktes im Definitionsbereich der Funktion ändern. Die gesuchte Funktion $f'(x)$ beschreibt dann kleine Änderungen um $\Delta x = x - x_0$ für $f(x)$, sodass der mögliche Fehler von zumindest quadratischer Ordnung in Δx ist

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \mathcal{O}((\Delta x)^2). \quad (2.13)$$

Das ist eine lineare Näherung. Indem man $\Delta f(x) = f(x + \Delta x) - f(x)$ einführt, bedeutet dies

$$\Delta f(x) = f'(x)\Delta x + \mathcal{O}((\Delta x)^2), \quad (2.14)$$

$$\frac{\Delta f(x)}{\Delta x} - f'(x) = \mathcal{O}((\Delta x)). \quad (2.15)$$

Der Bruch $\Delta f/\Delta x$ ist der **Differenzenquotient**. Im Limes $\Delta x \rightarrow 0$ verschwindet die rechte Seite von Gl. (2.15), da $\lim_{\Delta x \rightarrow 0} \Delta x = 0$.

Definition 2.7. Die **Ableitung** der Funktion f an der Stelle x_0 ist dann definiert als

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{(x + \Delta x) - x} = f'(x_0) = \frac{d}{dx} f(x_0), \quad (2.16)$$

vorausgesetzt, dieser Grenzwert existiert. Der Bruch df/dx ist der **Differentialquotient**. In dem Fall heisst die Funktion f **differenzierbar** an der Stelle x_0 . Sie heisst differenzierbar in einem Bereich $D' \subseteq D$ falls $f'(x)$ in jedem Punkt $x \in D'$ differenzierbar ist.

Wichtige Beispiele sind:

- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = c = \text{const}$. Dann gilt für jedes $x_0 \in \mathbb{R}$

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \equiv 0 \quad (2.17)$$

Die konstante Funktion ist also in ganz \mathbb{R} differenzierbar und ihre Ableitung verschwindet.

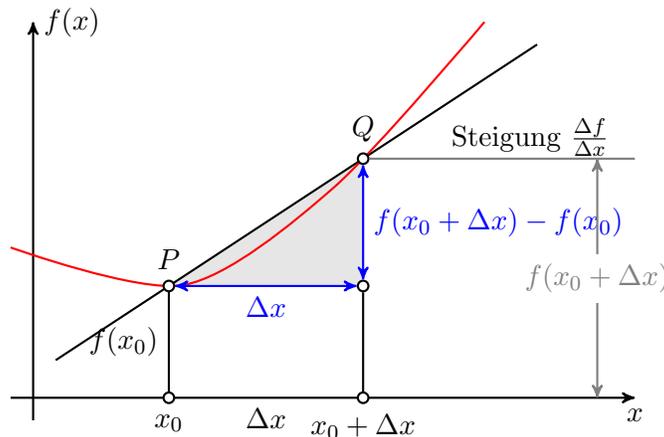


Figure 5: Graphische Darstellung der Ableitung.

- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x$. Dann folgt

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = \lim_{x \rightarrow x_0} 1 = 1. \quad (2.18)$$

- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \rightarrow x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} = \lim_{x \rightarrow x_0} (x + x_0) = 2x_0. \quad (2.19)$$

- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = ax^n$

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{ax^n - ax_0^n}{x - x_0} = a \lim_{x \rightarrow x_0} \frac{(x - x_0) \overbrace{(x^{n-1} + x^{n-2}x_0 + \dots + x_0^{n-1})}^{n\text{-Terme}}}{x - x_0} = anx_0^{n-1}. \quad (2.20)$$

- $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = \sum_{n=0}^m a_n x^n$. Die Ableitung eines Polynoms ist wieder ein Polynom und somit ist ein Polynom beliebig oft differenzierbar (mindestens zur Ordnung m):

$$f'(x_0) = \sum_{n=0}^m na_n x_0^{n-1}. \quad (2.21)$$

Die Ableitung besitzt auch eine **geometrische Bedeutung**. Betrachten wir dazu den Graphen einer Funktion $y = f(x)$. Dann bedeutet $[f(x) - f(x_0)]/(x - x_0)$ die Steigung der Sekante durch die Punkte $[x, f(x)]$ und $[x_0, f(x_0)]$. Die Ableitung ergibt, falls sie existiert, entsprechend die Steigung der **Tangente** an die durch den Funktionsgraphen gegebenen Kurve am Punkt $[x_0, f(x_0)]$. In der Physik wird oft auch $df/dt =: \dot{f}$ für Zeitkoordinaten verwendet.

2.2.1. Ableitungsregeln

Die Ableitung kann immer folgend der Definition bestimmt werden, aber in der Praxis werden Regeln bestimmt, die einmal bewiesen werden. Dabei nehmen wir an, dass alle beteiligten Funktionen im Punkt x_0 ihres Definitionsbereichs differenzierbar sind.

Summenregel. Die Ableitung ist eine **lineare Operation** ist, d.h. für beliebigen Funktionen f, g , die bei x_0 differenzierbar sind und $a, b \in \mathbb{R}$ gilt

$$\frac{d}{dx}(af)(x_0) = af'(x_0), \quad (2.22)$$

$$\frac{d}{dx}(f + g)(x_0) = f'(x_0) + g'(x_0), \quad (2.23)$$

$$\frac{d}{dx}(af + bg)(x_0) = af'(x_0) + bg'(x_0). \quad (2.24)$$

Beweis. Zum Beweis der allgemeinsten Form, nämlich Gl. (2.24), müssen wir wir den entsprechenden Limes betrachten,

$$\begin{aligned} \frac{d}{dx}(af + bg)(x_0) &= \lim_{x \rightarrow x_0} \frac{af(x) + bg(x) - [af(x_0) + bg(x_0)]}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \left[\frac{af(x) - af(x_0)}{x - x_0} + \frac{bg(x) - bg(x_0)}{x - x_0} \right] \\ &= af'(x_0) + bg'(x_0). \end{aligned} \quad (2.25)$$

Produktregel.

$$\frac{d}{dx}(fg)(x_0) = f'(x_0) \times g(x_0) + f(x_0) \times g'(x_0). \quad (2.26)$$

Beweis. Wir verwenden die Definition der Ableitung (2.16) und addieren im Nenner $0 = g(x_0 + \Delta x)f(x_0) - g(x_0 + \Delta x)f(x_0)$

$$\begin{aligned} \frac{d}{dx}(fg)(x_0) &= \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x)g(x_0 + \Delta x) - f(x_0)g(x_0)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \left[\frac{[f(x_0 + \Delta x) - f(x_0)]}{\Delta x} g(x_0 + \Delta x) + f(x_0) \frac{[g(x_0 + \Delta x) - g(x_0)]}{\Delta x} \right]. \end{aligned} \quad (2.27)$$

Nun ist eine Funktion, die in x_0 differenzierbar ist, dort auch stetig, und damit ergibt sich aus (2.27) im Limes $\Delta x \rightarrow 0$ die Produktregel (2.26).

Um zu zeigen, dass eine in x_0 **differenzierbare Funktion** f auch **stetig** ist, gehen wir auf die Definition des Limes 2.1 zurück. Da f in x_0 differenzierbar ist muss der Grenzwert $\lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x}$

existieren. Demnach gibt es zu jedem $\epsilon > 0$ ein $\delta_\epsilon > 0$, so dass

$$\left| \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} - f'(x_0) \right| < \epsilon \quad \text{falls} \quad |\Delta x| < \delta_\epsilon. \quad (2.28)$$

Wir können nun diese Ungleichung mit $|\Delta x|$ multiplizieren:

$$|f(x_0 + \Delta x) - f(x_0) - \Delta x f'(x_0)| < \epsilon |\Delta x|. \quad (2.29)$$

Für $\Delta x \rightarrow 0$ strebt die rechte Seite der Ungleichung sowie $\Delta x f'(x_0)$ unter dem Betrag auf der linken Seite gegen 0. Damit ist aber

$$\lim_{\Delta x \rightarrow 0} |f(x_0 + \Delta x) - f(x_0)| = 0 \Leftrightarrow \lim_{\Delta x \rightarrow x_0} f(x_0 + \Delta x) = f(x_0), \quad (2.30)$$

d.h. f ist an der Stelle x_0 stetig (der Grenzwert existiert). \square

Kettenregel. Seien f und g Funktionen, wobei g bei x_0 und f bei $g(x_0)$ differenzierbar seien, dann gilt

$$\frac{d}{dx}(g \circ f)(x_0) = \frac{d}{dx}g[f(x_0)] = g'[f(x_0)]f'(x_0). \quad (2.31)$$

Beweis. Wir verwenden die Definition der Ableitung (2.16) und erweitern mit der Einheit $1 = (f(x_0 + \Delta x) - f(x_0))/(f(x_0 + \Delta x) - f(x_0))$:

$$\begin{aligned} \frac{d}{dx}(g \circ f)(x_0) &= \lim_{\Delta x \rightarrow 0} \frac{g[f(x_0 + \Delta x)] - g[f(x_0)]}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{g[f(x_0 + \Delta x)] - g[f(x_0)]}{f(x_0 + \Delta x) - f(x_0)} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \\ &= \lim_{\Delta x, y \rightarrow 0} \frac{g(y_0 + \Delta y) - g(y_0)}{\Delta y} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}, \end{aligned} \quad (2.32)$$

wo wir $f(x_0) =: y_0$ und $f(x_0 + \Delta x) =: y_0 + \Delta y$ definiert haben. Da nach der gerade durchgeführten Überlegung g an der Stelle x_0 stetig ist, folgt daraus für $\Delta x \rightarrow 0$ die Kettenregel (2.31).

Quotientenregel. Wenden wir weiter die Produkt- und die Kettenregel auf die Funktion $f/g = f \cdot 1/g$ an, wobei f und g in x_0 differenzierbar und $g(x_0) \neq 0$ sei, erhalten wir die **Quotientenregel**

$$\frac{d}{dx} \left(\frac{f(x_0)}{g(x_0)} \right) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}. \quad (2.33)$$

Beweis. Wir verwenden hintereinander die Definitionen der Produkt- (2.26) und Kettenregel (2.31)

$$\begin{aligned} \frac{d}{dx} \left(\frac{f(x_0)}{g(x_0)} \right) &= \frac{d}{dx} \left(f(x_0) \frac{1}{g(x_0)} \right) = \frac{f'(x_0)}{g(x_0)} + f(x_0) \frac{d}{dx} g^{-1}(x_0) \\ &= \frac{f'(x_0)}{g(x_0)} - f(x_0) \frac{g'(x_0)}{g^2(x_0)}. \end{aligned} \quad (2.34)$$

wobei mit $g^{-1} = 1/g$ und nicht die Umkehrfunktion gemeint ist.

Differentiation der Umkehrfunktion.

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'[f^{-1}(y)]}. \quad (2.35)$$

Dazu betrachten wir noch eine Funktion $f : I \rightarrow \mathbb{R}$, die in einem Intervall I strikt monoton ist, d.h. es gilt für alle $x_1 < x_2 \in I$ stets $f(x_1) < f(x_2)$ (**strikt monoton wachsende**) oder $f(x_1) > f(x_2)$ (**strikt monoton fallende**) Funktion. Es ist dann klar, dass in diesem Intervall die Zuordnung des Funktionswertes $y = f(x)$ zum Bildpunkt x umkehrbar eindeutig ist, d.h. zu jedem Wert y im Wertebereich der Funktion existiert genau ein $x \in I$ mit $f(x) = y$, d.h. diese Gleichung ist eindeutig nach x auflösbar. Dies ermöglicht die Definition der **Umkehrfunktion** der Funktion f : $f^{-1} : f(I) \rightarrow I$. Dabei ist $f(I) = \{y \mid \text{es gibt ein } x \in I \text{ mit } f(x) = y\}$ die **Wertemenge** der Funktion f . Wir nehmen nun an, die Funktion f sei differenzierbar im Intervall I .

Beweis. Aufgrund der Definition der Umkehrfunktion gilt für alle $x \in I$

$$(f^{-1} \circ f)(x) = f^{-1}[f(x)] = x. \quad (2.36)$$

Da die rechte Seite differenzierbar und auch f differenzierbar ist, ist demnach auch f^{-1} differenzierbar und nach der Kettenregel gilt

$$\frac{d}{dx} f^{-1}[f(x)] = (f^{-1})'[f(x)]f'(x) = 1 = \frac{d}{dx} x, \quad (2.37)$$

$$(f^{-1})'[f(x)] = \frac{1}{f'(x)}. \quad (2.38)$$

Mit Hilfe dieser Formel können wir die Ableitung der Umkehrfunktionen monotoner Funktionen finden, wenn wir die Ableitung der Funktion selbst kennen. Das erkennt man einfacher, indem man $y = f(x)$ und $x = f^{-1}(y)$ schreibt. Dann ergibt Gl. (2.38) wiederum die zu zeigenden Gl. (2.35).

Beispiel 7. Betrachten wir als Beispiel $f(x) = x^2$. Setzen wir dann $y = f(x) = x^2$, folgt $x = \sqrt{y}$ (für $y > 0$). Mit Gl. (2.35) folgt daraus

$$\frac{d}{dy} \sqrt{y} = \frac{1}{f'(x)} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}. \quad (2.39)$$

Nennen wir jetzt y wieder x , folgt die Ableitungsformel für die Wurzelfunktion

$$\frac{d}{dx} \sqrt{x} = \sqrt{x}' = \frac{1}{2\sqrt{x}}. \quad (2.40)$$

Beispiel 8. *Differentiation von trigonometrischen Funktionen. Dazu verwenden wir die Definition der Ableitung (2.16)*

$$\begin{aligned}
 \frac{d}{dx} \sin x &= \lim_{\Delta x \rightarrow 0} \frac{\sin(x + \Delta x) - \sin x}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\sin x \cos \Delta x + \cos x \sin \Delta x - \sin x}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\sin x (\cos \Delta x - 1) + \cos x \sin \Delta x}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \sin x \lim_{\Delta x \rightarrow 0} \frac{\cos \Delta x - 1}{\Delta x} + \lim_{\Delta x \rightarrow 0} \cos x \lim_{\Delta x \rightarrow 0} \frac{\sin \Delta x}{\Delta x} \\
 &= (\sin x) \times (0) + (\cos x) \times (1) = \cos x,
 \end{aligned} \tag{2.41}$$

wobei wir im zweiten Schritt das trigonometrische Additionstheorem (1.58) verwenden.

Dasselbe können wir auch über die Reihendarstellung aus Gl. (1.45) und der individuellen Differentiation der Polynome zeigen,

$$\begin{aligned}
 \frac{d}{dx} \sin x &= \frac{d}{dx} \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = \sum_{k=0}^{\infty} (-1)^k (2k+1) \frac{z^{2k}}{(2k+1)!} \\
 &= \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!} = \cos x.
 \end{aligned} \tag{2.42}$$

was wiederum Gl. (1.44) entspricht.

Aus den obigen Regeln und den Definitionen der entsprechenden Funktionen, erstellen wir folgende Ableitungsliste:

$$\begin{array}{lll}
 \frac{d}{dx} c = 0 & \frac{d}{dx} \sinh x = \cosh x & \frac{d}{dx} \sin x = \cos x \\
 \frac{d}{dx} e^x = e^x & \frac{d}{dx} \cosh x = \sinh x & \frac{d}{dx} \cos x = -\sin x \\
 \frac{d}{dx} \ln |x| = \frac{1}{x} & \frac{d}{dx} \tanh x = 1 - \tanh^2 x = \frac{1}{\cosh^2 x} & \frac{d}{dx} \tan x = 1 + \tan^2 x = \frac{1}{\cos^2 x} \\
 \frac{d}{dx} x^\mu = \mu x^{\mu-1} & \frac{d}{dx} \operatorname{arsinh} x = \frac{1}{\sqrt{1+x^2}} & \frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}} \\
 \frac{d}{dx} e^{f(x)} = f'(x) e^{f(x)} & \frac{d}{dx} \operatorname{arcosh} x = \frac{1}{\sqrt{x^2-1}} & \frac{d}{dx} \arccos x = -\frac{1}{\sqrt{1-x^2}} \\
 \frac{d}{dx} \ln |f(x)| = \frac{f'(x)}{f(x)} & \frac{d}{dx} \operatorname{artanh} x = \frac{1}{1-x^2} & \frac{d}{dx} \arctan x = \frac{1}{1+x^2}
 \end{array} \tag{2.43}$$

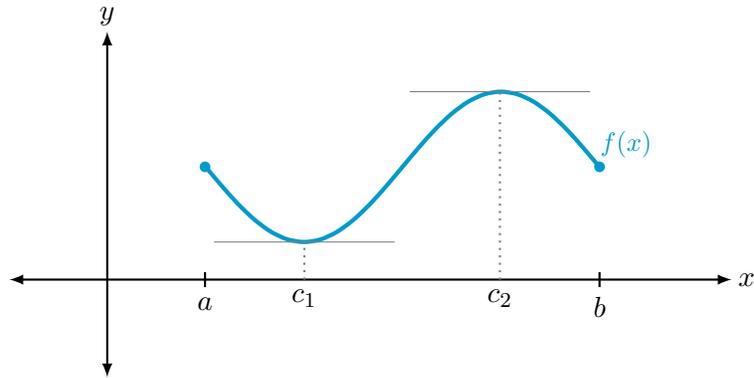


Figure 6: Graphische Darstellung des Satzes von Rolle mit zwei Extremstellen $f(c_1) = f(c_2) = 0$.

2.3. Hauptsätze der Differentialrechnung

Referenzen: [1] Abschn. (4.1), [2] Abschn. A.6.3.

Globale Eigenschaften von differenzierbaren Funktionen können oftmals durch lokale Eigenschaften, wie der Ableitung, beschränkt werden. Die folgenden Sätze sind die Grundlage zur Minimierung, Maximierung und Kurvendiskussion von Funktionen.

Satz 2.8 (Satz von Fermat). Sei $f : [a, b] \rightarrow \mathbb{R}$ an $x_0 \in (a, b)$ differenzierbar. Hat f an der Stelle x_0 ein lokales Maximum oder Minimum, gilt notwendigerweise $f'(x_0) = 0$.

Beweis. Für ein lokales Maximum beweisen wir dies, indem wir die Folgen (x_n) und (x'_n) betrachten, wobei gilt

$$x_n = x_0 + \frac{1}{n}, \quad x'_n = x_0 - \frac{1}{n}. \quad (2.44)$$

Hat f an x_0 ein lokales Maximum, dann gilt

$$\begin{aligned} f'(x_0) &= \lim_{n \rightarrow \infty} \frac{f(x_0 + \frac{1}{n}) - f(x_0)}{\frac{1}{n}} \leq 0, \\ f'(x_0) &= \lim_{n \rightarrow \infty} \frac{f(x_0 - \frac{1}{n}) - f(x_0)}{-\frac{1}{n}} \geq 0. \end{aligned} \quad (2.45)$$

Damit ist $f'(x_0) = 0$, da dies für alle $f'(x_0) \neq 0$ einen Widerspruch zur Folge hat. \square

Satz 2.9 (Satz von Rolle (Michel Rolle 1652–1719)). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $x \in (a, b)$ differenzierbar. Gilt dann $f(a) = f(b)$, existiert mindestens eine Stelle $\xi \in (a, b)$ mit $f'(\xi) = 0$; cf. Abb. 6.

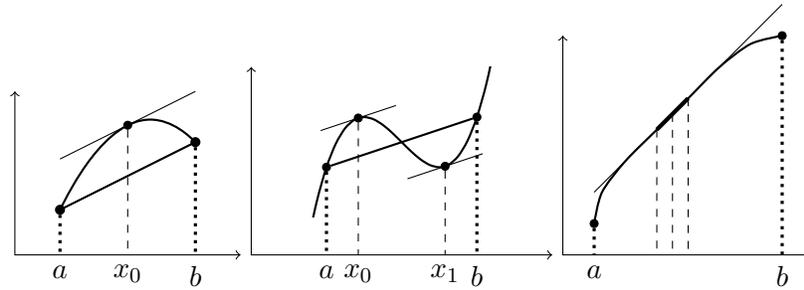


Figure 7: Beispiele für den Satz 2.10 Ein (links), zwei (mitte), unendlich viele Stellen für x_0 .

Beweis. Ist $f(x) = \text{const}$, ist die Behauptung trivial, denn dann ist $f'(x) = 0$ für alle $x \in (a, b)$. Falls dies nicht der Fall ist, gibt es wenigstens eine Stelle x_0 mit $f(x_0) > f(a) = f(b)$ oder $f(x_0) < f(a) = f(b)$. Da die Funktion f in dem abgeschlossenen Intervall $[a, b]$ voraussetzungsgemäss stetig ist, nimmt sie nach dem Satz vom Maximum und Minimum (cf. Satz 2.6) irgendwo in diesem Intervall das Infimum $f(\xi) = \inf_{x \in [a, b]} f(x)$ oder das Supremum $f(\xi) = \sup_{x \in [a, b]} f(x)$ ihres Bildbereiches an. Im ersten Fall muss offenbar das Supremum und im zweiten Fall das Infimum an wenigstens einer Stelle ξ im offenen Intervall (a, b) angenommen werden, d.h. es liegt zumindest ein Extremum in (a, b) vor, und nach dem eben bewiesenen Satz 2.8 ist $f'(\xi) = 0$.

Satz 2.10 (1. Mittelwertsatz der Differentialrechnung). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gibt es eine Stelle $\xi \in (a, b)$ mit, cf. Abb. 7

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \quad (2.46)$$

Dies ist eine Verallgemeinerung des Satzes von Rolle.

Beweis. Wir wenden den Satz von Rolle 2.9 auf folgende Funktion an:

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \quad (2.47)$$

Diese Funktion erfüllt die Voraussetzungen des Satzes von Rolle. Da f auf $[a, b]$ stetig und auf (a, b) differenzierbar ist, gilt dies sicher auch für g . Ausserdem ist $g(a) = g(b) = f(a)$. Folglich gibt es eine Stelle $\xi \in (a, b)$ mit $g'(\xi) = 0$. Nun ist aber

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \Rightarrow 0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}. \quad (2.48)$$

□

Folge 2.10.1. Man kann folgende Überlegungen anstellen:

(i) f sei differenzierbar und monoton wachsend auf $]a, b[$. dann ist $f'(x) \geq 0 \forall x \in]a, b[$. Aus dem ersten Mittelwertsatzes (2.46)

$$\underbrace{f(b) - f(a)}_{>0} = f'(\xi) \underbrace{(b - a)}_{>0}, \quad (2.49)$$

folgt daher, dass $f'(\xi) > 0$ Gleichsam für monoton fallende Funktionen.

(ii) Es gebe eine Unsicherheit in der Lagebestimmung des Messgeräts, i.e. in der Koordinate Δx . Dies führt zu einem systematischen Fehler im Messwert: $\Delta f = f'(x_0)\Delta x$.

Satz 2.11 (2. Mittelwertsatz der Differentialrechnung (Cauchy 1789–1857)). Seien die Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann existiert mindestens eine Stelle $\xi \in (a, b)$ mit

$$g'(\xi)[f(b) - f(a)] = f'(\xi)[g(b) - g(a)]. \quad (2.50)$$

Ist weiterhin $g'(x) \neq 0$ auf (a, b) , so ist $g'(\xi) \neq 0$ und $g(b) \neq g(a)$, und es folgt

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \quad (2.51)$$

Beweis. Wir wenden den Satz von Rolle 2.9 auf folgende Funktion an:

$$\begin{aligned} \zeta(x) &= g(x)[f(b) - f(a)] - f(x)[g(b) - g(a)], \\ \zeta(b) &= g(b)[f(b) - f(a)] - f(b)[g(b) - g(a)] = -g(b)f(a) + g(a)f(b), \\ \zeta(a) &= g(a)[f(b) - f(a)] - f(a)[g(b) - g(a)] = +g(a)f(b) - g(b)f(a) = \zeta(b), \\ \implies \zeta'(\xi) &= g'(x)[f(b) - f(a)] - f'(x)[g(b) - g(a)] = 0. \end{aligned} \quad (2.52)$$

□

Satz 2.12 (Regel von de l'Hospital (Marquis de l'Hospital 1661–1704)). Seien die Funktionen f und g (geeignet) differenzierbar auf und gelte $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = l$ mit $\lim_{x \rightarrow a} g'(x) \neq 0$. Dann folgt aus $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ oder ∞ , die Aussage

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = l. \quad (2.53)$$

Beweis. Laut erweitertem Mittelwertsatz 2.11 mit $f(a) = g(a) = 0$ gilt:

$$\frac{f(b)}{g(b)} = \frac{f'(\xi)}{g'(\xi)}, \quad \text{mit } a < \xi < b. \quad (2.54)$$

Jetzt betrachtet man $b \rightarrow a^+$, was wiederum $\xi \rightarrow a^+$ inkludiert.

Die Regel von de l'Hospital erleichtert in vielen Fällen die Bestimmung eines Grenzwertes. Hat man nach einmaliger Anwendung wiederum einen unbestimmten Ausdruck, kann der Satz wiederum auf die Funktionen f' und g' angewendet werden, sofern die entsprechenden Voraussetzungen vorliegen.

- Grenzwerte der Form $(\frac{\infty}{\infty})$ oder $(0 \times \infty)$ können genommen werden, weil sie in die Form $(\frac{0}{0})$ übergeführt werden können:

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(x)}{g(x)} &= \lim_{x \rightarrow a} \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}} \stackrel{(2.53)}{=} \lim_{x \rightarrow a} \frac{\frac{-g'(x)}{g^2(x)}}{\frac{-f'(x)}{f^2(x)}} = \lim_{x \rightarrow a} \frac{g'(x) f^2(x)}{f'(x) g^2(x)}, \\ 1 &= \lim_{x \rightarrow a} \frac{g'(x) f(x)}{f'(x) g(x)} \Leftrightarrow (2.53). \end{aligned} \tag{2.55}$$

- Die Aussage gilt auch, falls $a = \pm\infty$ ist. Wir setzen $x = \frac{1}{y}$ und erhalten

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{y \rightarrow 0^+} \frac{f(\frac{1}{y})}{g(\frac{1}{y})} \stackrel{(2.53)}{=} \lim_{y \rightarrow 0^+} \frac{f'(\frac{1}{y})(-y^2)}{g'(\frac{1}{y})(-y^2)} = \lim_{y \rightarrow 0^+} \frac{f'(\frac{1}{y})}{g'(\frac{1}{y})} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}. \tag{2.56}$$

- Ausdrücke der Form $f(x)^{g(x)} = (1^\infty), (0^0), (\infty^0)$ können durch $\lim f(x)^{g(x)} = \lim e^{g(x) \ln f(x)} = e^{\lim g(x) \ln f(x)}$ auf zuvor diskutierte Fälle zurückgeführt werden.
- Ausdrücke der Form $f(x) - g(x) = (\infty - \infty)$ müssen zuerst in eine der zuvor erwähnten Formen gebracht werden, um damit weiter arbeiten zu können.

Beispiel 9. Grenzwerte der Form $(\frac{0}{0})$; cf. Bsp. 4:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \lim_{x \rightarrow 0} \frac{\cos x}{1} = 1, \tag{2.57}$$

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \lim_{x \rightarrow 0} \frac{\frac{1}{1+x}}{1} = 1. \tag{2.58}$$

Beispiel 10. Der Logarithmus wächst bei $x \rightarrow 0$ langsamer als jede Potenz für $n > 0$:

$$\underbrace{\lim_{x \rightarrow 0^+} x^n \ln x}_{(0 \times \infty)} = \lim_{x \rightarrow 0^+} \underbrace{\frac{\ln x}{x^{-n}}}_{(\frac{\infty}{\infty})} \stackrel{(2.53)}{=} \lim_{x \rightarrow 0^+} \frac{\frac{1}{x}}{-nx^{-n-1}} = \lim_{x \rightarrow 0^+} -\frac{x^n}{n} = 0. \tag{2.59}$$

Beispiel 11. Die Exponentialfunktion wächst bei $x \rightarrow \infty$ schneller als jede Potenz:

$$\lim_{x \rightarrow \infty} \frac{x^n}{e^x} = \lim_{x \rightarrow \infty} \frac{nx^{n-1}}{e^x} = \dots = \lim_{x \rightarrow \infty} \frac{n!}{e^x} = 0. \tag{2.60}$$

2.4. Begriffe der Differentialrechnung

Referenzen: [1] Abschn. (4.2–3),

Falls eine Funktion, $y(x)$, als Lösung der Gleichung $F(x, y) = \text{const.}$ definiert wird, nennt man sie eine *implizite Funktion*. Die explizite Lösung $y(x)$ ist eine *explizite Funktion*.

Partielle Ableitung. Funktioniert gleich der “Ableitung nach einem Parameter”, ist aber symmetrisch mit Bezug auf alle Variablen, die als Koordinaten betrachtet werden:

$$\frac{\partial f(x, a)}{\partial x} := \left. \frac{df}{dx} \right|_{a=\text{const.}}, \quad (2.61)$$

$$\frac{\partial f(x, a)}{\partial a} := \left. \frac{df}{da} \right|_{x=\text{const.}}. \quad (2.62)$$

Beispiel 12.

$$f(x) = ax^2, \quad \frac{\partial f}{\partial x} = 2ax, \quad \frac{\partial f}{\partial a} = x^2. \quad (2.63)$$

Implizite Differentiation.

$$\begin{aligned} 0 &\stackrel{!}{=} \Delta F = F(x_0 + \Delta x, y(x_0 + \Delta x)) - F(x_0, y(x_0)) \\ 0 &= \frac{F(x_0 + \Delta x, y(x_0 + \Delta x)) - F(x_0, y(x_0 + \Delta x))}{\Delta x} + \frac{F(x_0, y(x_0 + \Delta x)) - F(x_0, y(x_0))}{\Delta x} \\ &\stackrel{\Delta x \rightarrow 0}{=} \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{\partial y}{\partial x}, \end{aligned} \quad (2.64)$$

wo wir im letzten Schritt die Kettenregel 2.31 verwendet haben. Daraus folgt, dass

$$\frac{\partial y}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}, \quad \text{falls } \frac{\partial F}{\partial y} \neq 0. \quad (2.65)$$

Beispiel 13.

$$F(x, y) = x^{\frac{2}{3}} + y^{\frac{2}{3}} = a^{\frac{2}{3}}, \quad y'(x) = -\frac{\frac{2}{3}x^{-\frac{1}{3}}}{\frac{2}{3}y^{-\frac{1}{3}}} = -\left(\frac{y}{x}\right)^{\frac{1}{3}}. \quad (2.66)$$

Beispiel 14.

$$c = \frac{1}{2}mv^2, \quad v = \sqrt{\frac{2c}{m}}, \quad \frac{dv}{dm} = -\frac{1}{2} \frac{\sqrt{2c}}{m^{\frac{3}{2}}} = -\frac{1}{2} \frac{v}{m}, \quad (2.67)$$

$$F(v, m) = \frac{1}{2}mv^2, \quad \frac{dv}{dm} = -\frac{\frac{\partial F}{\partial m}}{\frac{\partial F}{\partial v}} = -\frac{\frac{1}{2}v^2}{mv} = -\frac{1}{2} \frac{v}{m}. \quad (2.68)$$

Höhere Ableitungen.

$$(f')' =: f'' =: f^{(2)}, \quad f^{(n)} := (f^{(n-1)})', \quad \frac{\partial^2 f}{\partial x \partial a}, \quad \frac{\partial^{m+n} f}{\partial x^m \partial a^n}, \quad \frac{\partial^2 f}{\partial a \partial x}. \quad (2.69)$$

2.5. Taylor Entwicklung

Referenzen: [1] Abschn. (1.2), [2] Abschn. A.7.6.

Die Kenntnis von f' liefert gewisse Rückschlüsse auf die Funktion f selbst, z.B. Monotonie, mögliche lokale Extrema. Die Kenntnis von f'' liefert darüberhinaus eine Information, ob dieses Wachsen bzw. Fallen von f' zunimmt oder abnimmt. Dies führt zur Überlegung, ob bei Kenntnis aller Ableitungen an einer Stelle x_0 die Funktion global oder zumindest auf einem Intervall rekonstruierbar ist.

Sei $f : D \rightarrow \mathbb{R}$ mit offenem Definitionsbereich $D \subseteq \mathbb{R}$ eine Funktion, für die $f^{(n)}$ in einem abgeschlossenen Intervall $I \subset D$ stetig und daher mindestens $(n + 1)$ -mal ($n \in \{0, \dots, n\}$) stetig differenzierbar sei. Sei $a \in D$. Wir suchen dann eine Näherung von f durch ein **Polynom** n -ten Grades für Argumente “in der Nähe” von a . Sei $x \in I$. Dann gibt es Koeffizienten $c_k \in \mathbb{R}$, so dass

$$f(x) = \sum_{k=0}^n c_k (x - a)^k + R_n(x, a). \quad (2.70)$$

Der erste Term ist das “Taylor Polynom” und der zweite Term das “Lagrange-Restglied”, für das gilt $R_n(x, a) = \mathcal{O}[(x - a)^{n+1}]$ und daher

$$\lim_{n \rightarrow \infty} \frac{R_n(x, a)}{(x - a)^{n+1}} = \text{const.} \quad (2.71)$$

Um die Koeffizienten c_k in (2.70) zu bestimmen, leiten wir diese Gleichung k -mal ($k \in \{0, 1, \dots, n\}$) ab und setzen danach $x = a$:

$$\left. \frac{d^k}{dx^k} f(x) \right|_{x=a} =: f^{(k)}(a) = k! c_k, \quad (2.72)$$

denn wegen (2.71) verschwindet die k -te Ableitung des Restglieds für $x = a$. Damit wird

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k + R_n(x, a). \quad (2.73)$$

Dies ist die **Taylorsche Formel**:

$$T_n(x, a) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k \quad \text{heisst Taylorpolynom } n\text{-ter Ordnung,} \quad (2.74)$$

$$R_n(x, a) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - a)^{n+1} \quad \text{heisst Restglied nach Lagrange.} \quad (2.75)$$

Beweis. ObdA sei $x > a$, und x fest. Betrachte die Hilfsfunktion und deren Ableitung an ξ

$$g(t) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x-t)^k - m \frac{(x-t)^{n+1}}{(n+1)!}, \quad (2.76)$$

$$\begin{aligned} g'(\xi) &= 0 - \sum_{k=0}^n \frac{f^{(k+1)}(\xi)}{k!} (x-\xi)^k + \sum_{k=1}^n \frac{f^{(k)}(\xi)}{(k-1)!} (x-\xi)^{k-1} + m \frac{(x-\xi)^n}{n!} \\ &= - \sum_{k=0}^n \frac{f^{(k+1)}(\xi)}{k!} (x-\xi)^k + \sum_{k=0}^{n-1} \frac{f^{(k+1)}(\xi)}{k!} (x-\xi)^k + m \frac{(x-\xi)^n}{n!} \\ &= - \frac{f^{(n+1)}(\xi)}{n!} (x-\xi)^n + m \frac{(x-\xi)^n}{n!}, \end{aligned} \quad (2.77)$$

mit $t \in [a, x]$ wobei $m = m(x, a)$ so gewählt wird, dass $g(a) = 0$. Nachdem auch $g(x) = 0$ ist, sind die Voraussetzung des Satzes von Rolle 2.9 erfüllt, daher $\exists \xi \in (a, x)$ mit $g'(\xi) = 0$. Sodass aus (2.77) folgt $m = f^{(n+1)}(\xi)$, das nun dem Lagrangeschen Restglied entspricht. Setzt man nun in $g(t)$ für $t = a$, so erhalten wir die Taylor-Formel. \square

Falls nun f im Intervall (a, x) bzw. (x, a) sogar beliebig oft stetig differenzierbar ist und ein $S > 0$ existiert, so dass $|f^{(n+1)}(\xi)| < S$ für alle $\xi \in (a, x)$ bzw. $\xi \in (x, a)$ ist, folgt aus (2.75)³

$$|R_n(x, a)| < S \frac{|x-a|^{n+1}}{(n+1)!} \rightarrow 0 \quad \text{für } n \rightarrow \infty, \quad (2.78)$$

dass das Restglied verschwindet. In diesem Fall besitzt f die **Taylor-Reihe**

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k = \sum_{k=0}^{\infty} \frac{(x-a)^k}{k!} \left(\frac{d}{dx} \right)^k f(a) = \exp \left[(x-a) \frac{d}{dx} \right] f(a), \quad (2.79)$$

wobei wir im letzten Schritt die Exponentialfunktion des Differentialoperators durch eine formale Reihe schreiben konnten.

Beispiel 15 (Geometrische-Reihe).

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 \dots \quad \text{für } |x| < 1. \quad (2.80)$$

Beispiel 16 (Exponentialfunktion). Wir betrachten die Taylor-Entwicklung von $f(x) = \exp(x)$ um $a = 0$. Da

$$f^{(k+1)}(x) = \exp(x), \quad (2.81)$$

diese Funktion die Bedingungen für die Taylor-Entwicklung für jedes $x \in \mathbb{R}$ erfüllt, gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \quad (2.82)$$

³Um zu zeigen, dass für beliebige $y > 0$ $\lim_{n \rightarrow \infty} y^n/n! = 0$ gilt, bemerken wir, dass für $n > 2y$ und $N \in \mathbb{N}$ die Abschätzung $0 \leq y^{n+N}/(n+N)! \leq y^n y^N / ((2y)^N n!) = y^n/n! (1/2)^N$ gilt. Der letzte Ausdruck strebt aber für $N \rightarrow \infty$ gegen 0 für fixes n .

Beispiel 17 (Sinus). Wir betrachten die Taylor-Entwicklung von $f(x) = \sin(x)$ um $a = 0$. Da

$$f(x) = \sin x, \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad \dots, \quad (2.83)$$

gilt also

$$f^{(2k)}(0) = 0, \quad f^{(2k+1)}(0) = (-1)^k, \quad k \in \{0, 1, 2, \dots\}. \quad (2.84)$$

Die Taylor-Reihe für den Sinus ist also

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}. \quad (2.85)$$

Es ist klar, dass wegen $-1 \leq \sin x \leq 1$ und $-1 \leq \cos x \leq 1$ aufgrund der Restgliedformel die Taylorreihen des Sinus (2.85) als auch des Cosinus sicher überall in \mathbb{R} und tatsächlich gegen $\sin x$ bzw. $\cos x$ konvergieren.

Gleichsam findet man die Reihendarstellungen von $\exp(x)$, $\sinh(x)$, $\cosh(x)$, $\sin(x)$, $\cos(x)$, ..

Beispiel 18 (Logarithmus). Wir betrachten die Taylor-Entwicklung von $f(x) = \ln(1+x)$ um $a = 0$. Da $f(x)$ beliebig oft differenzierbar auf $(-1, \infty)$ ist für alle $k \in \mathbb{N}$, folgt

$$f(x) = \ln(1+x), \quad f'(x) = \frac{1}{1+x}, \quad f^{(k)}(x) = \frac{(k-1)!(-1)^{k+1}}{(1+x)^k}, \quad \Rightarrow \frac{f^{(k)}(0)}{k!} = \frac{(-1)^{k+1}}{k}. \quad (2.86)$$

Mit $f(0) = 0$ gilt dann für $x \in (-1, \infty)$ die Taylor-Reihe für den Logarithmus

$$T_n(x, 0) = \sum_{k=1}^n \frac{(-1)^{k+1}}{k} x^k \quad \text{für } n \rightarrow \infty. \quad (2.87)$$

Beispiel 19 (Binomische Reihe). Wir betrachten die Taylor-Entwicklung von $f(x) = (1+x)^\mu$ um $a = 0$ für $\mu \in \mathbb{R}$ und $x \in (-1, 1)$. Da $f(x)$ beliebig oft differenzierbar auf $(-1, \infty)$ ist für alle $k \in \mathbb{N}$, folgt

$$f(x) = (1+x)^\mu, \quad f^{(k)}(x) = \mu(\mu-1)\dots(\mu-k+1)(1+x)^{\mu-k}, \quad \Rightarrow \frac{f^{(k)}(0)}{k!} = \binom{\mu}{k}, \quad (2.88)$$

wobei der Binomialkoeffizient $\binom{\mu}{k} = \frac{\mu(\mu-1)\dots(\mu-k+1)}{k!}$. Mit $f(0) = 1$ folgt dann für $|x| < 1$ die Binomische Reihe

$$T_n(x, 0) = \sum_{k=0}^n \binom{\mu}{k} x^k \quad \text{für } n \rightarrow \infty. \quad (2.89)$$

Beispiel 20. Wir betrachten die Taylor-Entwicklung von $f(x) = \arctan(x)$ um $a = 0$. Die Reihendarstellung folgt aus Gl. (2.43)

$$\frac{d \arctan x}{dx} = \frac{1}{1+x^2} = 1 - x^2 + x^4 + \dots = \sum_{k=0}^{\infty} (-1)^k x^{2k}, \quad (2.90)$$

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}, \quad (2.91)$$

wobei wir im letzten Schritt die Terme individuell integriert haben.

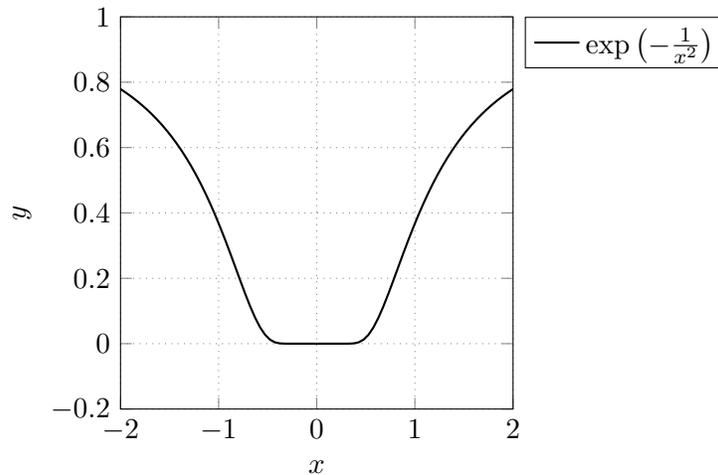


Figure 8: Gegenbeispiel für eine Taylor-Reihe.

Beispiel 21. Es gibt auch Gegenbeispiele zu konvergierenden Taylor-Reihen. Dazu betrachten wir $f(x) = \exp(-\frac{1}{x^2})$ um $a = 0$ wie in Abb. 8. Die Ableitungen an $a = 0$ sind

$$\begin{aligned}
 f(0) &= \exp(-\infty) = 0, \\
 f'(0) &= \lim_{x \rightarrow 0} \frac{2}{x^3} f(x) = \lim_{y \rightarrow \infty} \frac{2y^3}{e^{y^2}} \stackrel{(2.53)}{=} 0, \\
 f^{(n)}(0) &= 0 \quad \forall n.
 \end{aligned} \tag{2.92}$$

Daher konvergiert die Taylor-Reihe zwar formal mit $\sum_{k=0}^{\infty} \frac{0}{k!} x^k = 0$, aber das Ergebnis ist nicht richtig, weil das Restglied nach Lagrange R_n bei $n \rightarrow \infty$ nicht unterdrückt wird:

$$\begin{aligned}
 f^{(2)}(x) &= \left(\frac{4}{x^6} - \frac{6}{x^4} \right) f(x), \\
 f^{(3)}(x) &= \left(\dots + \frac{4!}{x^5} \right) f(x), \\
 f^{(n)}(x) &= \left(\dots + \frac{(n+1)!}{x^{n+2}} \right) f(x),
 \end{aligned} \tag{2.93}$$

$$R_n = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} = \left(\dots + \frac{(n+2)!}{\xi^{n+3}} \right) f(\xi) \frac{x^{n+1}}{(n+1)!} = \left(\dots + (n+2) \left(\frac{x}{\xi} \right)^{n+1} \right) \frac{1}{\xi^2} f(\xi), \tag{2.94}$$

wobei der letzte Ausdruck für $n > 0$ wächst da $\frac{x}{\xi} > 1$ weil $a < \xi < x$.

2.6. Über die Konvergenz von Reihen

Satz 2.13. Falls eine Funktion durch eine Reihe definiert ist, i.e. $f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$, dann stellt diese Reihe auch die Taylor-Reihe um den Punkt $x = a$ dar.

Beweis. Die Koeffizienten der Taylor-Reihe sind $\frac{f^{(k)}(a)}{k!}$. Daher ist

$$\begin{aligned} f(x) &= c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3 + \dots & f(a) &= c_0, \\ f'(x) &= c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \dots & f'(a) &= c_1, \\ f^{(2)}(x) &= 2c_2 + 6c_3(x-a) + \dots & f^{(2)}(a) &= 2!c_2, \\ f^{(3)}(x) &= 6c_3 + \dots & f^{(3)}(a) &= 3!c_3, \end{aligned} \quad (2.95)$$

wodurch $c_k = \frac{f^{(k)}(a)}{k!}$. Es spielt daher keine Rolle wie die Reihendarstellung konstruiert wird.

Zur praktischen Analyse der Konvergenz von Reihen kann man sehr oft das Quotientenkriterium heranziehen. Dazu seien

$$f_n(x) := \sum_{k=0}^n c_k(x-x_0)^k, \quad (2.96)$$

$$\rho_k(x) := \left| \frac{c_{k+1}(x-x_0)}{c_k} \right| = \left| \frac{c_{k+1}(x-x_0)^{k+1}}{c_k(x-x_0)^k} \right|. \quad (2.97)$$

wobei ρ auch **Konvergenzradius** genannt wird. Demnach ist die Reihe konvergent, falls der Grenzwert $\lim_{n \rightarrow \infty} f_n(x)$ existiert und sei $\rho(x) = \lim_{k \rightarrow \infty} \rho_k(x)$, falls existent. Dann gilt

$$\begin{aligned} \rho(x) < 1 & \quad \text{die Reihe ist absolut konvergent,} \\ \rho(x) = 1 & \quad \text{keine Aussage möglich (kann konvergieren falls das Vorzeichen alterniert),} \\ \rho(x) > 1 & \quad \text{die Reihe ist divergent.} \end{aligned} \quad (2.98)$$

Die Werte von x mit $\rho(x) < 1$ bilden den **Konvergenzbereich**. Ist der Limes $\rho(x) = 0$, bedeutet dies, dass die Potenzreihe für alle $x \in \mathbb{R}$ konvergiert.

Beweis. Beweis der Konvergenz. Falls $\rho(x) < 1$ dann $\exists k_0$ sodass $\rho_k(x) < \delta < 1 \forall k > k_0$. Dann gilt:

$$f_n(x) = \underbrace{\sum_{k=0}^{k_0} c_k(x-x_0)^k}_{\text{endlich}} + \underbrace{\sum_{k=k_0+1}^n c_k(x-x_0)^k}_{|\sum_{k=k_0+1}^n c_k(x-x_0)^k| \leq \sum_{k=k_0+1}^n |c_k(x-x_0)^k|}, \quad (2.99)$$

wobei wir den zweiten Term mittels der Dreiecksungleichung aus Fn. 2 nach oben hin abschätzen konnten. Da aber laut Gl. (2.97) jeder Summand auf den k_0 -Term zurückgeführt werden kann folgt, dass

$$\sum_{k=k_0+1}^n |c_k(x-x_0)^k| < |c_{k_0}(x-x_0)^{k_0}|(\delta + \delta^2 + \delta^3 + \dots) \quad (2.100)$$

Dies ist aber wiederum die Geometrische Reihe die für $\delta < 1$ konvergiert und somit auch einen endlichen Beitrag zu $f_n(x)$ ergibt.

Mit absolut konvergenten Reihen kann man in der Regel genau so umgehen wie mit Polynomen. Reihen können umgeordnet, dividiert, multipliziert und die Ordnung von Ableitungen (Integration) und Summierung vertauscht werden i.e. $\frac{d}{dx} \sum_k c_k(x-x_0)^k = \sum_k c_k k(x-x_0)^{k-1}$.

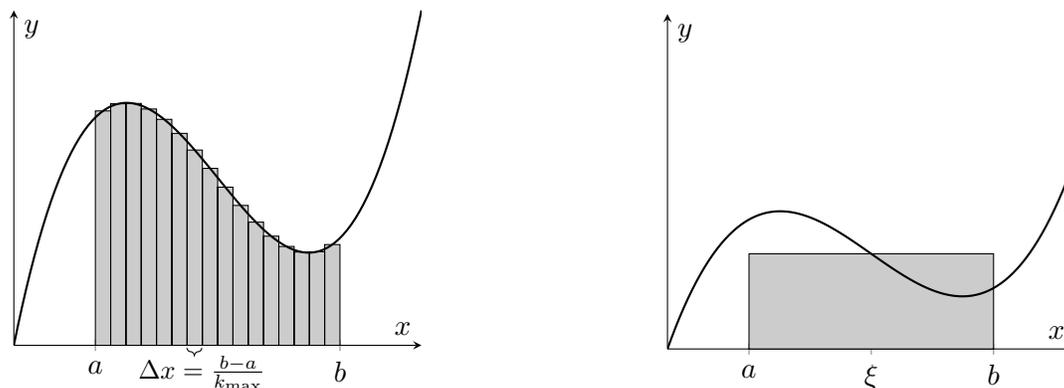


Figure 9: Links: Integration im Intervall (a, b) mit endlicher Intervallsegmentierung. Rechts: Beispiel zum Mittelwertsatz der Integralrechnung 2.18.

2.7. Hauptsätze der Integralrechnung

Referenzen: [1] Abschn. (5.1.1), [2] Abschn. A.7.1–2.

Die Integralrechnung ist, in gewissem Sinne die Umkehrung der Differentialrechnung. Sie beschäftigt sich generell mit dem geometrischen Inhaltsbegriff (Bogenlänge einer Kurve, Flächeninhalte von Flächen, Volumen von Körpern) und ist daher eng mit der **Maßtheorie** verknüpft. In dieser Vorlesung behandeln wir nur das **Riemann-Integral**. Für an der modernen Lebesgue-Integrationstheorie Interessierte sei auf [5] verwiesen, wo man eine Darstellung findet, die nicht die allgemeine Maßtheorie voraussetzt.

Definition 2.14. Sei $f(x)$ stetig $\forall x \in]a, b[$. Die **Stammfunktion** bzw. das **unbestimmte Integral** von $f(x)$ ist eine stetige und differenzierbare Funktion $F(x)$, die die Differentialgleichung

$$F'(x) = f(x) \quad \forall x \in]a, b[, \quad (2.101)$$

erfüllt.

Das Ziel der Integralrechnung für reellwertige Funktionen einer reellen Variablen ist, anschaulich formuliert, die Berechnung des Flächeninhaltes unter der Kurve der Funktion $f : [a, b] \rightarrow \mathbb{R}$, wobei $[a, b] \subset \mathbb{R}$ ein endliches Intervall ist. Zunächst setzen wir nur voraus, dass die Funktion f beschränkt sei, d.h. es gibt Zahlen m und M , so dass $m \leq f(x) \leq M$ für alle $x \in [a, b]$. Um den Flächeninhalt zu bestimmen, wird das Intervall $[a, b]$ in n Teilintervalle $[x_{j-1}, x_j]$ mit $j \in \{0, 1, \dots, n\}$ zerlegt, wobei $a = x_0 < x_1 < \dots < x_n = b$ ist und die Funktion durch eine entsprechende **Treppenfunktion** zu approximieren. Dabei ist eine Treppenfunktion eine Funktion, die auf jedem Teilintervall $[x_{j-1}, x_j]$ konstant ist.

Nun kann man für jedes Teilintervall das Infimum bzw. das Supremum der Funktionswerte in diesem Intervall verwenden. Nach dem Satz vom Infimum und Supremum existieren nämlich beide Größen,

weil voraussetzungsgemäss die Funktion f beschränkt ist. Dann bezeichnen wir als **Unter- und Obersumme** der Funktion f bzgl. der Zerlegung $[x_{j-1}, x_j]$ des Intervalls $[a, b]$ die Grössen

$$\begin{aligned} S_u &= \sum_{j=1}^n (x_j - x_{j-1}) \inf\{f([x_{j-1}, x_j])\}, \\ S_o &= \sum_{j=1}^n (x_j - x_{j-1}) \sup\{f([x_{j-1}, x_j])\}. \end{aligned} \quad (2.102)$$

Der Flächeninhalt unter der Kurve für alle Zerlegungen des Intervalls $[a, b]$ liegt stets zwischen diesen beiden Werten und wir erhalten den Flächeninhalt, indem wir die Unterteilung immer feiner machen. Die Anzahl der Intervalle $n \rightarrow \infty$, wobei zugleich $\max_{j \in \{1, \dots, n\}} (x_j - x_{j-1}) \rightarrow 0$ gehen soll. Der Flächeninhalt existiert dann sicher, wenn S_u und S_o gegen denselben Wert streben. Wir sagen dann, die Funktion f sei über das Intervall $[a, b]$ **Riemann-integrierbar**, und wir definieren das Integral als den entsprechenden Grenzwert

$$\int_a^b dx f(x) = \lim_{\Delta x \rightarrow 0} \sum_{k=1}^{n=k_{\max}} f(x_k) \Delta x = \lim_{n \rightarrow \infty} S_u = \lim_{n \rightarrow \infty} S_o. \quad (2.103)$$

Satz 2.15. *Wir bemerken sogleich, dass wenn $f, g : [a, c] \rightarrow \mathbb{R}$ über die Intervalle $[a, b]$ und $[b, c]$ Riemann-integrierbar sind, folgt dass folgende Integranden Riemann-integrierbar sind*

$$\int_a^b dx [\lambda_1 f(x) + \lambda_2 g(x)] = \lambda_1 \int_a^b dx f(x) + \lambda_2 \int_a^b dx g(x), \quad \text{für } \lambda_1, \lambda_2 \in \mathbb{R}, \quad (2.104)$$

$$\int_a^c dx f(x) = \int_a^b dx f(x) + \int_b^c dx f(x), \quad (2.105)$$

$$\int_a^b dx f(x) = - \int_b^a dx f(x), \quad \text{für } a < b, \quad (2.106)$$

$$\int_a^b dx \lambda_1 = (b - a) \lambda_1, \quad \text{für } a < b, \quad (2.107)$$

$$\int_a^a dx f(x) = 0. \quad (2.108)$$

Beweis. Die Beweise folgen aus der Additivität der Definition (2.103):

$$\int_a^b dx \lambda_1 = \lim_{\Delta x \rightarrow 0} \sum_{k=1}^{k_{\max}} f(x_k) \Delta x = \lim_{\Delta x \rightarrow 0} \{\lambda_1 k_{\max} \Delta x\} = \lim_{\Delta x \rightarrow 0} \{\lambda_1 (b - a)\}. \quad (2.109)$$

Satz 2.16 (Hauptsatz der Differential- und Integralrechnung). *Nach dem vorigen Abschnitt können wir für eine auf $[a, b]$ stetige Funktion f die **Integralfunktion***

$$F(x) = \int_a^x dx' f(x'), \quad (2.110)$$

definieren. Der **Hauptsatz der Analysis** besagt, dass dann F eine auf $[a, b]$ differenzierbare Funktion ist und folgende Differentialgleichung erfüllt (cf. (2.101))

$$F'(x) = f(x). \quad (2.111)$$

Beweis. Offenbar ist für $\Delta x > 0$

$$F(x + \Delta x) - F(x) = \left[\int_a^{x+\Delta x} - \int_a^x \right] dx' f(x') = \left[\int_a^{x+\Delta x} + \int_x^a \right] dx' f(x') = \int_x^{x+\Delta x} dx' f(x'). \quad (2.112)$$

Seien nun

$$m = \inf_{x' \in [x, x+\Delta x]} f(x'), \quad M = \sup_{x' \in [x, x+\Delta x]} f(x'). \quad (2.113)$$

Beide Werte existieren, und es gibt nach dem Satz vom Maximum und Minimum 2.6 (Bolzano-Weierstrass), Zahlen $\xi_{\Delta x}, \xi'_{\Delta x} \in [x, x + \Delta x]$, so dass $m = f(\xi_{\Delta x})$ und $M = f(\xi'_{\Delta x})$. Demnach ist

$$\int_x^{x+\Delta x} dx f(\xi_{\Delta x}) \leq F(x + \Delta x) - F(x) \leq \int_x^{x+\Delta x} dx f(\xi'_{\Delta x}). \quad (2.114)$$

Nach Ausführung der Integrale am Infimum bzw. Supremum (vgl. Gl. (2.107)) und der Division durch Δx , erhalten wir

$$\begin{aligned} \Delta x f(\xi_{\Delta x}) &\leq F(x + \Delta x) - F(x) \leq \Delta x f(\xi'_{\Delta x}) \\ f(\xi_{\Delta x}) &\leq \frac{F(x + \Delta x) - F(x)}{\Delta x} \leq f(\xi'_{\Delta x}). \end{aligned} \quad (2.115)$$

Für $\Delta x \rightarrow 0$ gehen nun $\xi_{\Delta x}, \xi'_{\Delta x} \rightarrow x$ und wegen der Stetigkeit von f entsprechen $f(\xi_{\Delta x}), f(\xi'_{\Delta x}) \rightarrow f(x)$, womit (2.111) bewiesen ist. Analog gehen wir für $\Delta x < 0$ vor indem wir Gl. (2.106) für allgemeine Integrale benutzen. Dann müssen wir nur die obige Betrachtung für das Integral

$$\int_x^{x+\Delta x} dx f(x) = - \int_{x+\Delta x}^x dx f(x) \quad (2.116)$$

anstellen. Dann ergibt sich auch für den Grenzwert $\Delta x \rightarrow 0$ für $\Delta x < 0$, dass $F'(x) = f(x)$ ist. Damit ist F differenzierbar und besitzt die Ableitung $F' = f$ für alle $x \in [a, b]$.

Definition 2.17. Wir bemerken noch, dass demnach für jede **Stammfunktion** $\tilde{F} : [a, b] \rightarrow \mathbb{R}$ von f , gilt

$$\int_a^b dx f(x) = \tilde{F}(b) - \tilde{F}(a). \quad (2.117)$$

Dabei heisst \tilde{F} Stammfunktion zu f , falls

$$\tilde{F}'(x) = f(x) \quad (2.118)$$

gilt. Man bezeichnet eine beliebige Stammfunktion von f auch als **unbestimmtes Integral** und schreibt

$$\tilde{F}(x) = \int dx f(x) + C, \quad C = \text{const.} \quad (2.119)$$

Beweis. Die Eindeutigkeit der Stammfunktion bis auf eine additive Konstante ergibt sich, indem wir annehmen, es existiert eine weitere Stammfunktion $\tilde{F}(x) = F(x) + \tilde{F}(a)$. Da aber $\tilde{F}'(x) = F'(x) = f(x)$ und also $\tilde{F}'(x) - F'(x) \equiv 0$. Nun muss eine auf $[a, b]$ differenzierbare Funktion, deren Ableitung überall im Intervall $[a, b]$ verschwindet, konstant sein (vgl. Gl. (2.46)). Da $F'(a) = 0$ ist, gilt also die Behauptung.

Satz 2.18 (Mittelwertsatz der Integralrechnung). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $\varphi : [a, b] \rightarrow \mathbb{R}$ mit $\varphi \geq 0$ integrierbar. Dann gibt es eine Zahl $\xi \in [a, b]$, so dass

$$\int_a^b dx f(x)\varphi(x) = f(\xi) \int_a^b dx \varphi(x). \quad (2.120)$$

Beweis. Nach den Sätzen zur Riemann-Integrierbarkeit ist mit den obigen Voraussetzungen $f\varphi$ Riemann-integrierbar. Da f auf dem abgeschlossenen Intervall stetig ist, existieren $M = \sup_{x \in [a, b]} f(x)$ und $m = \inf_{x \in [a, b]} f(x)$. Da weiter $\varphi \geq 0$ ist, folgt $m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x)$, und damit

$$m \int_a^b dx \varphi(x) \leq \int_a^b f(x)\varphi(x) \leq M \int_a^b dx \varphi(x). \quad (2.121)$$

Da auch $\int_a^b dx \varphi(x) \geq 0$ ist, gibt es demnach eine Zahl μ mit $m \leq \mu \leq M$, so dass

$$\int_a^b dx f(x)\varphi(x) = \mu \int_a^b dx \varphi(x). \quad (2.122)$$

Da $f : [a, b] \rightarrow \mathbb{R}$ stetig ist, nimmt f die Werte m und M für irgendwelche Werte $\xi_1, \xi_2 \in [a, b]$ tatsächlich an: $f(\xi_1) = m$, $f(\xi_2) = M$. Wegen des Zwischenwertsatzes für stetige Funktionen 2.5 gibt es eine Zahl ξ zwischen ξ_1 und ξ_2 mit $f(\xi) = \mu$. Wegen (2.122) ist damit der Zwischenwertsatz der Integralrechnung bewiesen.

Für eine konstante Funktion $\varphi(x) = \text{const.}$ vereinfacht sich der Satz zu

$$\int_a^b dx f(x) = f(\xi)(b - a), \quad (2.123)$$

und der Beweis vereinfacht sich indem man den Mittelwertsatz der Differentialrechnung 2.10 benutzt

$$\int_a^b dx f(x) = F(b) - F(a) = F'(\xi)(b - a) = f(\xi)(b - a). \quad (2.124)$$

□

2.8. Begriffe der Integralrechnung

Referenzen: [1] Abschn. (5.1.2)

Mit dem Hauptsatz der Differential- und Integralrechnung können wir nun sehr viele Integrale berechnen, indem wir eine beliebige Stammfunktion zu der zu integrierenden Funktion suchen. Dafür

gibt es einige Rechentechniken, die wir hier kurz zusammenfassen. Zunächst kann man aus jeder Ableitungsregel eine Integrationsregel herleiten. Dazu können wir zunächst die Ableitungen der elementaren Funktionen aus Gl. (2.43) invertieren

$$\begin{aligned}
 \int^x dy 0 &= c & \int^x dy \cosh y &= \sinh x & \int^x dy \cos y &= \sin x \\
 \int^x dy e^y &= e^x & \int^x dy \sinh y &= \cosh x & \int^x dy \sin y &= -\cos x \\
 \int^x dy \frac{1}{y} &= \ln|x| & \int^x dy \frac{1}{\cosh^2 y} &= \tanh x & \int^x dy \frac{1}{\cos^2 y} &= \tan x \\
 \int^x dy y^\mu &= \frac{1}{\mu+1} x^{\mu+1} & \int^x dy \frac{1}{\sqrt{1+y^2}} &= \operatorname{arsinh} x & \int^x dy \frac{1}{\sqrt{1-y^2}} &= \arcsin x \\
 \int^x dy f'(y) e^{f(y)} &= e^{f(x)} & \int^x dy \frac{1}{\sqrt{y^2-1}} &= \operatorname{arcosh} x & \int^x dy \frac{1}{\sqrt{1-y^2}} &= -\arccos x \\
 \int^x dy \frac{f'(y)}{f(y)} &= \ln|f(x)| & \int^x dy \frac{1}{1-y^2} &= \operatorname{artanh} x & \int^x dy \frac{1}{1+y^2} &= \arctan x \quad (2.125)
 \end{aligned}$$

Hier haben wir die Notation $\int^x dy f(y) = [F(y)]_x = F(x)$ benutzt. Ableitungen von elementaren Funktionen sind immer elementare Funktionen. Deren Integrale sind aber nicht immer elementare Funktionen und werden häufig als *spezielle Funktionen* bezeichnet, e.g.

$$\operatorname{Li}_2(x) = -\int_0^x \frac{dy}{y} \ln(1-y), \quad x < 1, \quad \text{Dilogarithmus,} \quad (2.126)$$

$$E_1(x) = \int_x^\infty \frac{dy}{y} e^{-y}, \quad x > 0, \quad \text{exponentielles Integral,} \quad (2.127)$$

$$\Gamma(x) = \int_0^\infty dy y^{x-1} e^{-y}, \quad x > 0, \quad \text{Gammafunktion,} \quad (2.128)$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^\infty dy e^{-y^2}, \quad x > 0, \quad \text{Errorfunktion.} \quad (2.129)$$

Uneigentliche Integrale. Bislang betrachteten wir Riemann-Integrale $\int_a^b dx f(x)$ unter den Voraussetzungen, dass a, b endlich waren und f beschränkt auf $[a, b]$. Wir suchen nun Erweiterungen des Riemann-Integrals (falls möglich) auf unbeschränkte Integrationsintervalle bzw. auf unbeschränkte Funktionen.

Konkret können Integrale auch an a, b divergieren und/oder a, b bei $\pm\infty$ liegen. Dann wird das Integral als Grenzwert definiert, falls existent. Die Funktion $f(x)$ ist dann *integrierbar* auf $]a, b[$. Integrale die durch einen Grenzwert definiert sind, heissen *uneigentliche Integrale*.

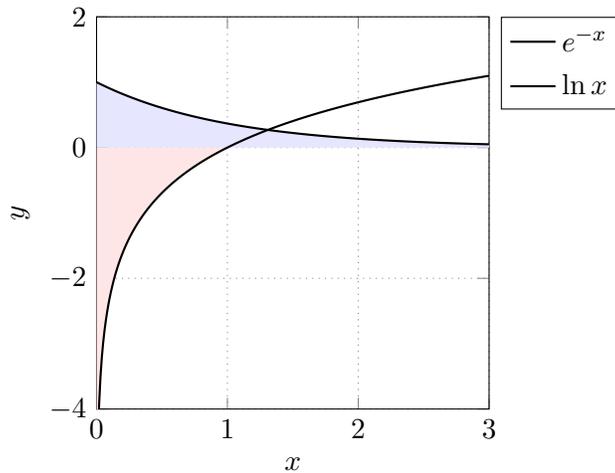


Figure 10: Beispiele für Uneigentliche Integrale aus Bsp. 22.

Beispiel 22.

$$\begin{aligned} \int_0^1 dx \ln x &:= \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 dx \ln x = \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^1 dx \frac{d}{dx} [x \ln x - x] \\ &= \lim_{\epsilon \rightarrow 0} [x \ln x - x]_{\epsilon}^1 = \lim_{\epsilon \rightarrow 0} [1 \ln 1 - 1 - \epsilon \ln \epsilon + \epsilon] = -1, \end{aligned} \quad (2.130)$$

$$\begin{aligned} \int_0^{\infty} dx e^{-x} &:= \lim_{\epsilon \rightarrow \infty} \int_0^{\epsilon} dx e^{-x} = \lim_{\epsilon \rightarrow \infty} \int_0^{\epsilon} dx \frac{d}{dx} [-e^{-x}] \\ &= \lim_{\epsilon \rightarrow \infty} [-e^{-x}]_0^{\epsilon} = \lim_{\epsilon \rightarrow \infty} [-e^{-\epsilon} + 1] = 1, \end{aligned} \quad (2.131)$$

2.9. Integrations Methoden

Referenzen: [1] Abschn. (5.2-3), [2] Abschn. A.7.4

Aus allgemeinen Regeln der Differentialrechnung erhalten wir entsprechend Rechenregeln für unbestimmte Integrale.

Variablentransformation oder Substitution der Variablen. Die Kettenregel der Differentialrechnung liefert die **Substitutionsregel für Integrale**. Sei φ eine differenzierbare $[\varphi' \exists]$ und invertierbare $[\varphi^{-1} \exists]$ Funktion.

$$\int_a^b dx f(x) = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} dt \varphi'(t) f[\varphi(t)] = \int_{t_a}^{t_b} dt \frac{dx}{dt} f(x) = \left[\int dx f(x) \right]_{x=\varphi(t)}, \quad (2.132)$$

für $x = \varphi(t)$ lässt sich diese Regel leichter merken, wenn man formal $dx = dt dx/dt$ schreibt. In vielen Fällen ist die Form $t = \varphi^{-1}(x)$ leichter zu identifizieren.

Beweis. Sei dazu $F(x)$ eine Stammfunktion von $f(x)$, und $\tilde{F}(t) = F[\varphi(t)]$, dann

$$\begin{aligned} \frac{d}{dt} \tilde{F}(t) &\stackrel{(2.31)}{=} \frac{d}{d\varphi} \tilde{F}[\varphi(t)] \frac{d\varphi(t)}{dt} = f[\varphi(t)] \varphi'(t) \\ &\Rightarrow \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} dt \varphi'(t) f[\varphi(t)] = [\tilde{F}(t)]_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} = F(b) - F(a), \end{aligned} \quad (2.133)$$

da $\tilde{F}(\varphi^{-1}(x)) = F[\varphi(\varphi^{-1}(x))] = F(x)$.

Beispiel 23. Als Anwendungsbeispiel betrachten wir einige Integrale

$$\int_a^b dx (1+cx)^4 \stackrel{t=cx}{=} \int_{ca}^{cb} \frac{dt}{c} (1+t)^4 \stackrel{s=1+t}{=} \frac{1}{c} \int_{1+ca}^{1+cb} ds s^4, \quad (2.134)$$

$$\int_a^b dx \frac{1}{\sqrt{1+y}} \stackrel{t=\sqrt{1+y}}{=} \int_{\sqrt{1+a}}^{\sqrt{1+b}} \frac{2tdt}{t}. \quad (2.135)$$

Beispiel 24. Als weiteres Beispiel betrachten wir die Aufgabe, die Fläche eines Halbkreises um den Ursprung mit Radius r zu berechnen. Für den Halbkreis gilt $x^2 + y^2 = r^2$ bzw. $y(x) = \sqrt{r^2 - x^2}$ mit $x \in [-r, r]$. Damit ist die Fläche

$$A = \int_{-r}^r dx \sqrt{r^2 - x^2}. \quad (2.136)$$

Um das Integral zu berechnen, substituieren wir nun $x = \varphi(t) = r \cos t$, $t = \varphi^{-1}(x) = \arccos \frac{x}{r}$.

Das Intervall $[-r, r]$ wird dann umkehrbar eindeutig auf das Intervall $t \in [0, \pi]$ abgebildet. Weiter ist $dx = dt (-r \sin t)$. Für $x = -r$ ist offenbar eindeutig $t = \pi$ und $x = +r$ für $t = 0$. Damit folgt

$$A = \int_{\pi}^0 dt (-r \sin t) \sqrt{r^2 - r^2 \cos^2 t} = r^2 \int_0^{\pi} dt \sin^2 t. \quad (2.137)$$

Um schliesslich dieses Integral zu berechnen, bemerken wir, dass

$$\cos(2t) = \cos^2 t - \sin^2 t = 1 - 2 \sin^2 t \Rightarrow \sin^2 t = \frac{1}{2} [1 - \cos(2t)]. \quad (2.138)$$

Dies in (2.137) eingesetzt liefert

$$A = \frac{r^2}{2} \int_0^{\pi} dt [1 - \cos(2t)] = \frac{r^2}{2} \left[t - \frac{1}{2} \sin(2t) \right]_{t=0}^{t=\pi} = \frac{\pi r^2}{2}. \quad (2.139)$$

Partielle Integration. Aus der Produktregel der Differentialrechnung (2.26) folgt die **Regel von der partiellen Integration**. Sei der Integrand ein Produkt zweier Funktionen, und sei eine davon direkt als Ableitung erkennbar: $f(x) = u(x)v'(x)$. Dann gilt:

$$\int_a^x dy u(y)v'(y) = u(x)v(x) - \int_a^x dy u'(y)v(y). \quad (2.140)$$

Die Funktionen u und v müssen hierzu angemessen identifiziert werden.

Beweis. Aus $(uv)' = u'v + v'u$ folgt

$$\int^x dy [u(y)v'(y) + u'(y)v(y)] \stackrel{(2.26)}{=} \int^x dy \frac{d}{dy} [u(y)v(y)] = u(x)v(x), \quad (2.141)$$

wobei wir im letzten Schritt verwendet haben, dass wenn $\int^x dy f(y) = F(x)$ und $F'(x) = f(x)$, dass $\int^x dy F'(y) = \int^x dy \frac{dF}{dy} = F(x)$.

Beispiel 25. Als Anwendungsbeispiel betrachten wir einige Integrale

$$\int^x dy \ln y = x \ln x - \int^x dy y \frac{d \ln y}{dy} = x \ln x - x, \quad u = \ln y, \quad v = y, \quad (2.142)$$

$$\int^x dy y \exp y = x \exp x - \int^x dy \exp y \frac{dy}{dy} = (x - 1) \exp x, \quad u = y, \quad v = \exp y. \quad (2.143)$$

In der Tat findet man durch Ableiten, dass tatsächlich $[(x - 1) \exp x]' = \exp x + (x - 1) \exp x = x \exp x$ gilt. In der Tat folgt aus der Produktregel $[x(\ln x - 1)]' = (\ln x - 1) + x/x = \ln x$. Weitere Beispiele sind:

$$\begin{aligned} \int^x dy \sin^2(y) &= -\sin x \cos x + \int^x dy \cos y \cos y \\ &= -\sin x \cos x + x - \int^x dy \sin^2 y, \quad u = \sin y, \quad v = -\cos y, \end{aligned} \quad (2.144)$$

$$\begin{aligned} I_n &= \int_0^\infty dx x^n e^{-ax} = \underbrace{-\frac{1}{a} [x^n e^{-ax}]_0^\infty}_{\rightarrow 0} + \frac{1}{a} \int_0^\infty dx n x^{n-1} e^{-ax} \\ &= \frac{n}{a} I_{n-1} = \frac{n(n-1)}{a^2} I_{n-2} = \frac{n!}{a^n} I_0 = \frac{n!}{a^{n+1}}, \quad u = x^n, \quad v = -\frac{e^{-ax}}{a}, \end{aligned} \quad (2.145)$$

Partialbruchzerlegung. Sei der Integrand von der Form $\frac{P(x)}{Q(x)}$, wobei P und Q Polynome sind. Falls $\text{Grad}(P) \geq \text{Grad}(Q)$, kann ein Teil von $P(x)$ durch Division vereinfacht werden. Dann bleibt ein Integrand mit $\text{Grad}(P) < \text{Grad}(Q)$, übrig. Die rationale Funktion ist dann,

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x-a)(x-b)^n \dots (x^2+px+q)(x^2+rx+s)^m \dots}, \quad (2.146)$$

wobei die ersten Terme von $Q(x)$ reelle Nullstellen und die letzteren Terme keine reellen Nullstellen besitzen. Die Schritte sind wie folgt:

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x-a) \dots (x^2+px+q) \dots} = \frac{1}{(x-a)} + \dots + \frac{Bx+C}{(x^2+px+q)} + \dots, \quad (2.147)$$

$$\frac{A}{(x-a)} = A \frac{d}{dx} \ln |x-a|, \quad (2.148)$$

$$\begin{aligned} \frac{Bx+C}{x^2+px+q} &= \frac{B}{2} \frac{2x+p}{x^2+px+q} + \left(C - \frac{Bp}{2}\right) \frac{1}{\left(x + \frac{p}{2}\right)^2 + q - \frac{p^2}{4}} \\ &= \frac{B}{2} \frac{d}{dx} \ln |x^2+px+q| + \left(C - \frac{Bp}{2}\right) \frac{1}{\left(q - \frac{p^2}{4}\right)} \frac{d}{dx} \arctan \frac{x + \frac{p}{2}}{\sqrt{q - \frac{p^2}{4}}}. \end{aligned} \quad (2.149)$$

Beispiel 26. Als Anwendungsbeispiel betrachten wir einige Integrale

$$\begin{aligned} \int^x dy \frac{1}{y^2 - a^2} &= \int^x dy \frac{1}{(y - q)(y + a)} = \int^x dy \frac{1}{2a} \left(\frac{1}{y - a} - \frac{1}{y + a} \right) \\ &= \frac{1}{2a} \left(\ln |y - a| - \ln |y + a| \right) = \frac{1}{2a} \left| \frac{x - a}{x + a} \right|, \end{aligned} \quad (2.150)$$

$$\begin{aligned} \int^x dy \frac{y}{y^2 - a^2} &= \int^x dy \frac{y}{(y - q)(y + a)} = \int^x dy \frac{1}{2} \left(\frac{1}{y - a} + \frac{1}{y + a} \right) \\ &= \frac{1}{2} \left(\ln |y - a| + \ln |y + a| \right) = \frac{1}{2} |x^2 - a^2|, \end{aligned} \quad (2.151)$$

$$\int^x dy \frac{y^2}{y^2 - a^2} = \int^x dy \frac{y^2 - a^2 + a^2}{y^2 - a^2} = x + \frac{a}{2} \left| \frac{x - a}{x + a} \right|. \quad (2.152)$$

Ableitung nach Parameter. Der dazu können wir nach Parametern in e.g. den Integrationsgrenzen Ableiten

$$\frac{d}{dx} \int_a^x dy f(y) = +f(x), \quad \frac{d}{dx} \int_x^b dy f(y) = -f(x). \quad (2.153)$$

Die Ableitung von

$$I(a(x), b(x), x) := \int_{a(x)}^{b(x)} dy f(y, x), \quad (2.154)$$

ist dann (cf. implizite Differentiation (2.64))

$$\begin{aligned} \frac{dI}{dx} &= \frac{\partial I}{\partial a} \frac{\partial a}{\partial x} + \frac{\partial I}{\partial b} \frac{\partial b}{\partial x} + \frac{\partial I}{\partial x} \\ &= -a'(x) f[a(x), x] + b'(x) f[b(x), x] + \int_{a(x)}^{b(x)} dy \frac{\partial f(y, x)}{\partial x}. \end{aligned} \quad (2.155)$$

Falls a und b unabhängig von x sind, gibt nur der letzte Term einen Beitrag.

Beispiel 27. Als Anwendungsbeispiel betrachten wir einige Integrale

$$\begin{aligned} \int^x dy \frac{1}{(y - a)^n} &\stackrel{n \geq 1}{=} \frac{1}{n - 1} \frac{\partial}{\partial a} \int^x dy \frac{1}{(y - a)^{n-1}} = \dots = \frac{1}{(n - 1)!} \frac{\partial^{n-1}}{\partial a^{n-1}} \int^x dy \frac{1}{(y - a)} \\ &= \frac{1}{(n - 1)!} \frac{\partial^{n-1}}{\partial a^{n-1}} \ln |x - a| = \frac{(-1)^{n-1} (n - 2)!}{(n - 1)! (x - a)^{n-1}} = \frac{-1}{(n - 1)(x - a)^{n-1}}, \end{aligned} \quad (2.156)$$

$$\begin{aligned} \int^x dy \frac{By + C}{(y^2 + py + q)^n} &\stackrel{n \geq 1}{=} \frac{1}{1 - n} \frac{\partial}{\partial q} \int^x dy \frac{By + C}{(y^2 + py + q)^{n-1}} = \dots \\ &= \frac{(-1)^{n-1}}{(n - 1)!} \frac{\partial^{n-1}}{\partial q^{n-1}} \int^x dy \frac{By + C}{(y^2 + py + q)}, \end{aligned} \quad (2.157)$$

$$I_n = \int_0^\infty x^n e^{-ax} = (-1)^n \frac{d^n}{da^n} \int_0^\infty e^{-ax} = (-1)^n \frac{d^n}{da^n} \frac{1}{a} = \frac{n!}{a^{n+1}}. \quad (2.158)$$

2.10. Gewöhnliche Differentialgleichungen

Referenzen: [1] Abschn. (6.1), [2] Abschn. 5.0

Die Entwicklung von Lösungsverfahren für **Differentialgleichungen** (DG) stellt eines der wichtigsten mathematischen Hilfsmittel für die Physik dar, denn die Naturgesetze werden durch eben solche Gleichungen beschrieben. Dabei stellen uns **gewöhnliche Differentialgleichungen** vor die Aufgabe, unbekannte Funktionen einer Veränderlichen (in der klassischen Mechanik ist das e.g. die Zeit t) aus Gleichungen zu bestimmen, die diese Funktion und Ableitungen dieser Funktion enthalten. Die höchste Ordnung der Ableitung bestimmt die **Ordnung der Differentialgleichung**.

Eine Funktion $y(x)$ kann *implizit* durch eine Gleichung definiert werden (cf. Abschn. 2.4). Dann ist die Koordinate x die unabhängige Variable und die Lösung $y(x)$ die abhängige Variable. Eine Verallgemeinerung ist die Gleichung

$$\begin{aligned} G(x, y) &= 0, \\ G(y^{(n)}, y^{(n-1)}, \dots, y', y, x) &= 0, \end{aligned} \quad (2.159)$$

die die Differentialgleichung n -ter Ordnung $n \in \mathbb{N}$ definiert. Hier sei $y^{(j)} = y(x)^{(j)}$ die j -te Ableitung der Funktion y nach der Koordinate x . Je nach Anwendung kann es erforderlich sein, alle Lösungen zu kennen oder nur bestimmte, die durch die Anwendung vorgeschriebene Nebenbedingungen erfüllen.

Systeme von Differentialgleichungen n -ter Ordnung sind entsprechend gekoppelte Differentialgleichungen für mehrere Funktionen f_1, \dots, f_k , wobei die die höchste Ableitung, die in diesen Gleichungen vorkommt, die n -te ist.

In der klassischen Mechanik haben wir es i.a. mit Differentialgleichungen oder Systemen von Differentialgleichungen 2. Ordnung zu tun, denn die **Newtonsche Bewegungsgleichung** für einen Massenpunkt, auf den irgendwelche vorgegebenen Kräfte wirken, lautet

$$m\ddot{\mathbf{y}} = \mathbf{F}(t, \mathbf{y}, \dot{\mathbf{y}}) = -\gamma\dot{\mathbf{y}} - k\mathbf{y} - mg, \quad (2.160)$$

wobei $\mathbf{y} = (y_1, y_2, y_3)^T$ und wir hier konkret eine Feder betrachten, mit $\ddot{\mathbf{y}}$ der Beschleunigung, $\gamma\dot{\mathbf{y}}$ der geschwindigkeits-abhängigen Reibung, $k\mathbf{y}$ der Federkraft, und mg der Schwerkraft.

Die Bewegung eines Planeten um die sehr viel schwerere Sonne (im Ursprung), deren Bewegung wir näherungsweise vernachlässigen dürfen, ergibt mit dem Newtonschen Gravitationsgesetz die Gleichung

$$m\ddot{\mathbf{x}} = -\gamma m M \frac{\mathbf{x}}{|\mathbf{x}|^3}, \quad (2.161)$$

wobei $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Hier hängt die Kraft nicht von der Geschwindigkeit ab. Dabei ist m die Masse des Planeten, M die der Sonne und γ die Newtonsche Gravitationskonstante.

Die Bewegung eines geladenen Teilchens in einem elektromagnetischen Feld (in SI-Einheiten für das elektromagnetische Feld) ist:

$$m\ddot{\mathbf{x}} = q [\mathbf{E}(t, \mathbf{x}) + \dot{\mathbf{x}} \times \mathbf{B}(t, \mathbf{x})]. \quad (2.162)$$

Dabei ist m die Masse des Teilchens, q seine Ladung und \mathbf{E} und \mathbf{B} das durch irgendwelche anderen Ladungen und Ströme erzeugte elektrische bzw. magnetische Feld.

Begriffe

Hier definieren wir die wichtigsten Begriffe für die Diskussion von Differentialgleichungen:

- Gewöhnliche DG
eine unabhängige Variable, e.g. $y(x)$.
- Partielle DG
mehrere Variablen, e.g.

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\phi(x, y) = 0. \quad (2.163)$$

- Lineare DG
Abhängigkeit von y ist *linear*,

$$\underbrace{\left(\frac{d^2}{dt^2} + \frac{\gamma}{m} \frac{d}{dt} + \frac{k}{m}\right)}_{:=L, \text{ Differentialoperator}} y = -g, \quad (2.164)$$

wobei wir Gl. (2.160) in einer Dimension geschrieben haben. Hier kann der Differentialoperator, L , linear auf die einzelnen Funktionen angewendet werden: $L(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 L y_1 + \alpha_2 L y_2$.

- Nichtlineare DG

$$\ddot{y} + y\dot{y} + y^3 = 0. \quad (2.165)$$

- Allgemeine Lösung
Enthält unbekannte Integrationskonstanten, wie ein unbestimmtes Integral.
Normalerweise gilt: Zahl der Konstanten = Ordnung der DG.
- Partikuläre (spezielle) Lösung
Keine Unbekannten, wie ein bestimmtes Integral.
- Anfangsbedingungen

In der Physik suchen wir Lösungen solcher Bewegungsgleichungen unter Vorgabe bestimmter **Anfangsbedingungen**, d.h. man gibt zur Zeit $t_0 = 0$ die Werte für Ort $\mathbf{x}(0) = \mathbf{x}_0$ und Geschwindigkeit $\dot{\mathbf{x}}(0) = \mathbf{v}_0$ des Teilchens vor und sucht Lösungen für die Differentialgleichung, die zusätzlich diese Anfangsbedingungen erfüllen:

$$\mathbf{x}(t_0) = \mathbf{x}_0, \quad \dot{\mathbf{x}}(t_0) = \mathbf{v}_0. \quad (2.166)$$

Es braucht dann so viele Anfangsbedingungen, wie es Integrationskonstanten gibt, um eine eindeutige Lösung zu wählen.

Normalerweise: bestimme an x_0 die Werte von $y, y', \dots, y^{(n-1)}$. Dann folgt $y^{(n)}$ von der DG und $y^{(n+1)}, \dots$ von Ableitungen der DG. Die Lösung kann also eindeutig als Taylor-Reihe

$$y(x) = \exp\left[(x - x_0) \frac{d}{dx}\right] y(x_0), \quad (2.167)$$

geschrieben werden.

- Randbedingungen

Nur $y(x_{\min})$ und $y(x_{\max})$ sind bekannt, nicht aber $y(x_0), y'(x_0)$. Dann muss man $y'(x_0)$ so wählen, dass $y(x_{\max})$ erfüllt wird.

- Definitionsbereich

$(x, y, y', \dots, y^{(n-1)}) \subset \mathbb{R}^{n+1}$, so dass die DG $y^{(n)}$ gelöst werden kann.

Aus der physikalischen Fragestellung heraus erwarten wir, dass diese Differentialgleichungen stets eine Lösung besitzen (**Existenz**) und bei Vorgabe der Anfangsbedingungen (2.166) eindeutig (**Eindeutigkeit**) sind. Die Beweise entsprechender **Existenz- und Eindeutigkeitssätze** für Differentialgleichungen bei genauer Bestimmung der Eigenschaften der involvierten Funktionen (wie F in (2.159) oder die Kräfte in der Newtonschen Bewegungsgleichung) sind Klassiker der Analysis und finden sich in vielen mathematischen Lehrbüchern. In diesem Skript beschränken wir uns auf die Lösungsmethoden für die einfachsten Typen von Differentialgleichungen. Von der umfangreichen Spezialliteratur sei hier nur auf [7, 8] verwiesen.

Beispiel 28. Betrachten wir den Fall $n = 1$, mit $G(y', y, x) = 0$. Wir nehmen an $G(y', y, x) = y' - yx$, d.h. $y' = yx$. Dazu kann man y' in der (x, y) -Ebene als Tangentensteigung skizzieren; vgl. Abb. 11. Die Allgemeine Lösung dazu ist

$$y(x) = C \exp\left(\frac{x^2}{2}\right), \quad (2.168)$$

mit der Integrationskonstante C .

2.11. Allgemeine Lösung linearer Differentialgleichungen erster Ordnung

Referenzen: [1] Abschn. (6.2), [2] Abschn. 5.1

Die allgemeine Form einer linearen Differentialgleichung erster Ordnung lautet

$$\text{inhomogene DG:} \quad y'(x) + p(x)y(x) + q(x) = 0, \quad (2.169)$$

$$\text{homogene DG:} \quad y'(x) + p(x)y(x) = 0, \quad (2.170)$$

als *homogen* ist hier die Abwesenheit eines konstanten Terms $q(x)$ gemeint (vgl. homogene DG 1. Ordnung). Die *inhomogene* Gleichung kann dann als $Ly = -q(x)$ geschrieben werden, wobei L wiederum der lineare Differentialoperator ist.

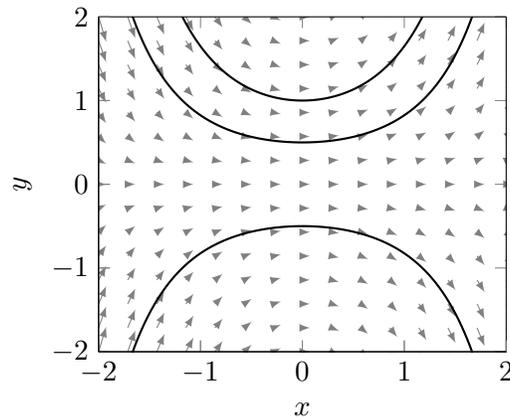


Figure 11: Beispiele für die allgemeine Lösung und einige partikuläre Lösungen der gewöhnlichen DG $G(y', y, x) = y' - yx$.

Satz 2.19.

$$\begin{aligned} \text{allgemeine Lösung der inhomogenen DG} &= \text{allgemeine Lösung der homogenen DG } (y_a) \\ &+ \text{partikuläre Lösung der inhomogenen DG } (y_p) \end{aligned} \quad (2.171)$$

Beweis. Seien Y_1 und Y_2 zwei unterschiedliche Lösungen der inhomogenen Gleichung

$$\begin{aligned} Y_1' + pY_1 + q &= 0, \\ Y_2' + pY_2 + q &= 0. \end{aligned} \quad (2.172)$$

Nach Subtraktion der Gleichungen $(Y_1 - Y_2)' + p(Y_1 - Y_2) = 0$ erkennt man, dass Unbestimmtheiten der homogenen DG genügen. Daher folgt, dass bei Kenntnis allgemeinen der Lösung, y_a , der homogenen Gleichung man die Differenz folgend ausdrücken kann

$$Y_1 - Y_2 = c_1 y_a(x). \quad (2.173)$$

Nimmt man jetzt an $Y_1 = y(x)$ sei die allgemeine Lösung der inhomogenen Gleichung und $Y_2 = y_p(x)$ irgendeine partikuläre Lösung davon, dann folgt

$$y(x) = c_1 y_a(x) + y_p(x). \quad (2.174)$$

Dieser Satz kann dann auf beliebige Ordnungen der DG erweitert werden (vgl. Abschn. 2.12).

Allgemeine Lösung der homogenen Gleichung (y_a). Für den dem homogenen Fall aus Gl. (2.170)

$$y_a'(x) + p(x)y_a(x) = 0. \quad (2.175)$$

Dividiert man diese Gleichung durch y_a , erkennt man, dass sie vom separablen Typ ist. Daher gibt es entweder die triviale Lösung

$$y'_a(x) = 0, \quad (2.176)$$

oder man bringt die DG auf die Form

$$\frac{y'_a(x)}{y_a(x)} = \frac{d}{dx} \ln |y_a(x)| = -p(x), \quad (2.177)$$

$$\ln |y_a(x)| = - \int_{x_0}^x d\tilde{x} p(\tilde{x}) + \text{const}, \quad (2.178)$$

$$|y_a(x)| = \exp \left\{ - \int_{x_0}^x d\tilde{x} p(\tilde{x}) + \text{const} \right\}. \quad (2.179)$$

Berücksichtigen wir die Anfangsbedingung $y_a(x_0) = y_0$ und integrieren diese Gleichung bzgl. x , erhalten wir nach einer einfachen Umformung die Lösung in der Form

$$y_a(x) = C \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right], \quad (2.180)$$

wobei die Integrationskonstante, $C = \pm \exp(\text{const})$ auch die triviale Lösung $y_a(x) = 0$ inkludiert.

Partikuläre Lösung der inhomogenen Gleichung (y_p). Mit Satz 2.19 ist die allgemeine Lösung der inhomogenen Gleichung durch die Summe aus der allgemeinen Lösung der homogenen Gleichung und einer beliebigen Lösung der inhomogenen Gleichung gegeben. Sei also $y_a(x) \neq 0$ eine Lösung der homogenen Gleichung. Dann können wir eine partikuläre Lösung der inhomogenen Gleichung finden, indem wir den **Ansatz der Variation der Konstanten** vornehmen.

$$y_p(x) = C(x) \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right], \quad (2.181)$$

$$y'_p(x) = C'(x) \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right] - p(x) C(x) \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right], \quad (2.182)$$

$$y'_p(x) + p(x) y_p(x) = C'(x) \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right]. \quad (2.183)$$

Die inhomogene DG (2.169) ist daher

$$y'_p(x) + p(x) y_p(x) = C'(x) \exp \left[- \int_{x_0}^x d\tilde{x} p(\tilde{x}) \right] = -q(x), \quad (2.184)$$

$$\frac{dC(x)}{dx} = -q(x) \exp \left[\int_{x_0}^x d\tilde{x} p(\tilde{x}) \right], \quad (2.185)$$

$$C(x) = - \int_{x_0}^x ds q(s) \exp \left[\int_{x_0}^s d\tilde{x} p(\tilde{x}) \right]. \quad (2.186)$$

Die allgemeine Lösung ist $y(x) = y_a(x) + y_p(x)$. Mit der Anfangsbedingung $y(x_0) = y_0$ wird $C \rightarrow y_0$ und die Lösung kann als

$$y(x) = \exp\left[-\int_{x_0}^x d\tilde{x} p(\tilde{x})\right] \left\{ y_0 - \int_{x_0}^x ds q(s) \exp\left[\int_{x_0}^s d\tilde{x} p(\tilde{x})\right] \right\}, \quad (2.187)$$

ausgedrückt werden.

Betrachten wir nun einige spezielle nicht-lineare DG, die sich lösen lassen:

Separierbare Differentialgleichung.

$$\frac{dy}{dx} = g(y)f(x). \quad (2.188)$$

Ähnlich der Lösung der homogenen DG schreiben wir

$$\frac{dy}{g(y)} = dx f(x), \quad (2.189)$$

$$\int_{y_0}^{y(x)} \frac{d\tilde{y}}{g(\tilde{y})} = \int_{x_0}^x d\tilde{x} f(\tilde{x}). \quad (2.190)$$

Das Ergebnis kann durch Ableitung nach x überprüft werden.

Beispiel 29.

$$y'(x) = -\frac{y^2}{x^2}, \quad \frac{dy}{y^2} = -\frac{dx}{x^2}, \quad \left[-\frac{1}{y}\right]_{y_0}^{y(x)} = \left[\frac{1}{x}\right]_{x_0}^x, \quad \frac{1}{y_0} - \frac{1}{y(x)} = \frac{1}{x} - \frac{1}{x_0}. \quad (2.191)$$

Differentialgleichung der Bernoulli-Form.

$$\frac{dy}{dx} = p(x)y + q(x)y^\alpha, \quad \alpha \neq 0, 1. \quad (2.192)$$

Wir führen eine neue abhängige Variable ein, $z(x) := (y(x))^{1-\alpha}$, $y(x) := (z(x))^{\frac{1}{1-\alpha}}$. Es gilt

$$y'(x) = \frac{1}{1-\alpha} (z(x))^{\frac{1}{1-\alpha}-1} z'(x) = \frac{1}{1-\alpha} (z(x))^{\frac{\alpha}{1-\alpha}} z'(x) \stackrel{(2.192)}{=} p z^{\frac{1}{1-\alpha}} + q z^{\frac{\alpha}{1-\alpha}}, \quad (2.193)$$

$$z' = (1-\alpha)pz + (1-\alpha)q, \quad (2.194)$$

wobei wir im letzten Schritt beide Seiten mit $(1-\alpha)z^{-\frac{\alpha}{1-\alpha}}$ multipliziert haben. Das ist eine lineare inhomogene DG und kann mit den obigen Methoden (vgl. Gl. (2.187)) gelöst werden.

Exaktes Differential.

$$\frac{dy}{dx} = -\frac{p(x, y)}{q(x, y)}, \quad (2.195)$$

wobei $p(x, y) = \frac{\partial F(x, y)}{\partial x}$ und $q(x, y) = \frac{\partial F(x, y)}{\partial y}$. Durch die Notation der impliziten Differentiation (vgl. Gl. (2.64))

$$dx p(x, y) + dy q(x, y) = 0, \quad (2.196)$$

$$\frac{\partial F(x, y)}{\partial x} dx + \frac{\partial F(x, y)}{\partial y} dy = dF = 0, \quad (2.197)$$

daher folgt $F = \text{const.}$

Beispiel 30.

$$y'(x) = -\frac{y}{x}, \quad F = xy, \quad y(x) = \frac{c}{x}. \quad (2.198)$$

Homogene DG erster Ordnung.

$$\frac{dy}{dx} = -\frac{p(x, y)}{q(x, y)} = f(x, y), \quad (2.199)$$

wobei $p(x, y)$ und $q(x, y)$ homogene Funktionen vom Grad m sind. Dazu definieren wir den Begriff der Homogenität einer DG. Dies ist nicht zu verwechseln mit homogenen linearen DG, womit ein verschwindender konstanter Term gemeint war (vgl. Gl. (2.169)). Angenommen, die DG besitzt die spezielle Form, dass p, q homogen zur Ordnung m sind, und f die **Homogenitätseigenschaft** erfüllt

$$p(\lambda x, \lambda y) = \lambda^m p(x, y), \quad q(\lambda x, \lambda y) = \lambda^m q(x, y), \quad f(\lambda x, \lambda y) = \frac{p(\lambda x, \lambda y)}{q(\lambda x, \lambda y)} = \frac{p(x, y)}{q(x, y)} = f(x, y). \quad (2.200)$$

Dann können wir durch die Substitution $z(x) = \frac{y(x)}{x}$ ($y(x) = xz(x)$) die Differentialgleichung in eine separable Gleichung transformieren. In der Tat ist dann wegen der Homogenitätseigenschaft (2.200) und Produktregel

$$y' = z(x) + xz'(x) \stackrel{(2.199)}{=} f(x, xz(x)) = f(1, z(x)) = \tilde{f}(z). \quad (2.201)$$

Etwas umgeformt erhalten wir die folgende Gleichung für z :

$$z' = \frac{\tilde{f}(z) - z}{x}, \quad (2.202)$$

die tatsächlich von der separablen Form ist

$$\int_{z_0}^{z(x)} \frac{d\tilde{z}}{\tilde{f}(\tilde{z}) - \tilde{z}} = \int_{x_0}^x \frac{d\tilde{x}}{\tilde{x}}. \quad (2.203)$$

Beispiel 31.

$$y'(x) = -\frac{y}{x}, \quad f(z_0) = -z, \quad -\frac{1}{2} \left[\ln z \right]_{z_0}^z = \left[\ln x \right]_{x_0}^x, \quad z = \frac{c}{x^2}, \quad y = \frac{c}{x}. \quad (2.204)$$

2.12. Lineare Differentialgleichungen zweiter Ordnung

Referenzen: [1] Abschn. (6.3.1), [2] Abschn. 5.2

Lineare Differentialgleichungen besitzen eine besonders einfache Lösungsstruktur. In der Physik kann man auch oft kompliziertere Probleme durch lineare Differentialgleichungen nähern e.g. harmonischen Oszillator als Näherung für das mathematische Pendel:

$$\ddot{\varphi}(t) + \frac{g}{l} \sin \varphi(t) = 0, \quad \implies \quad \ddot{\varphi}(t) + \frac{g}{l} \varphi(t) = 0, \quad (2.205)$$

wobei φ der Auslenkungswinkel ist, l die Pendellänge. Hier betrachten die Struktur der allgemeinen Lösungen für solche lineare DG 2. Ordnung.

Die allgemeine lineare Differentialgleichung (DG) 2. Ordnung lautet

$$\text{inhomogene DG:} \quad G(y'', y', y, x) = y''(x) + p(x)y'(x) + q(x)y(x) = f(x), \quad (2.206)$$

$$\text{homogene DG:} \quad G(y'', y', y, x) = y''(x) + p(x)y'(x) + q(x)y(x) = 0, \quad (2.207)$$

wobei

$$G(y'', y', y, x) = \underbrace{\left(\frac{d^2}{dx^2} + p(x) \frac{d}{dx} + q(x) \right)}_{=:L} y(x) = \underbrace{f(x)}_{\text{inhom. Term}}. \quad (2.208)$$

Dabei sind p , q und f vorgegebene Funktionen der unabhängigen Variablen x und $y(x)$ die gesuchte Funktion oder abhängige Variable. Hier besprechen wir nur die wichtigsten Grundlagen über die Struktur der Lösungen solcher linearer Differentialgleichungen.

Man nennt die obige DG (2.206) **homogen**, wenn $f(x) = 0$ und entsprechend **inhomogen**, wenn $f(x) \neq 0$. Wir betrachten zuerst die Lösungsstruktur der **homogenen Gleichung**, nämlich $Ly = 0$. Da die Ableitungsoperation linear ist, d.h. für irgendwelche zwei Funktionen $y_1(x)$ und $y_2(x)$, die mindestens zweimal differenzierbar sind,

$$\frac{d}{dx} [C_1 y_1(x) + C_2 y_2(x)] = C_1 y_1'(x) + C_2 y_2'(x), \quad \frac{d^2}{dx^2} [C_1 y_1(x) + C_2 y_2(x)] = C_1 y_1''(x) + C_2 y_2''(x), \quad (2.209)$$

gilt, ist für zwei Lösungen y_1 und y_2 von (2.207) auch die **Linearkombination**

$$y(x) = C_1 y_1(x) + C_2 y_2(x) \quad (2.210)$$

mit $C_1, C_2 = \text{const}$ eine weitere Lösung. Es ist weiter klar, dass wir zur eindeutigen Festlegung der Lösung **Anfangsbedingungen** fordern müssen, d.h. wir verlangen von der Lösung y der DG zusätzlich, dass sie und ihre erste Ableitung bei $x = x_0$ bestimmte Werte annimmt:

$$y(x_0) \stackrel{!}{=} y_0, \quad y'(x_0) \stackrel{!}{=} v_0. \quad (2.211)$$

Nehmen wir an, wir hätten zwei Lösungen y_1 und y_2 gefunden, können wir versuchen, die Konstanten C_1 und C_2 in der Linearkombination (2.210) so zu bestimmen, dass diese Anfangsbedingungen (2.211) gelten, d.h. wir müssen das lineare Gleichungssystem

$$\begin{aligned} C_1 y_1(x_0) + C_2 y_2(x_0) &= y_0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) &= v_0, \end{aligned} \quad (2.212)$$

nach C_1 und C_2 auflösen. Multiplizieren wir die erste Gleichung mit $y_2'(x_0)$ ($y_1'(x_0)$) und die zweite mit $y_2(x_0)$ ($y_1(x_0)$) und subtrahieren die beiden entstehenden Gleichungen, finden wir

$$C_1 [y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0)] = y_2'(x_0)y_0 - y_2(x_0)v_0, \quad (2.213)$$

$$C_2 [y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0)] = y_1(x_0)v_0 - y_1'(x_0)y_0. \quad (2.214)$$

Wenn die eckige Klammer nicht verschwindet, können wir nach C_1 und C_2 auflösen:

$$C_1 = \frac{y_2'(x_0)y_0 - y_2(x_0)v_0}{y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0)}, \quad (2.215)$$

$$C_2 = \frac{y_1(x_0)v_0 - y_1'(x_0)y_0}{y_1(x_0)y_2'(x_0) - y_2(x_0)y_1'(x_0)}. \quad (2.216)$$

Wir untersuchen nun noch, wann die Bedingung, dass der Nenner in (2.215) und (2.216) für die Lösungen y_1 und y_2 der DG nicht verschwindet, erfüllt ist. Es handelt sich um die **Determinante der Koeffizientenmatrix** des linearen Gleichungssystems (2.212). Um diesen Ausdruck näher zu untersuchen, definieren wir für die beiden Lösungen y_1 und y_2 die **Wronski-Determinante**,

$$W(x) = W(y_1, y_2) = \det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} = y_1(x)y_2'(x) - y_2(x)y_1'(x). \quad (2.217)$$

Berechnen wir die erste Ableitung, finden wir mit Hilfe der Produktregel nach einiger Rechnung

$$W'(x) = y_1(x)y_2''(x) - y_2(x)y_1''(x). \quad (2.218)$$

Jetzt verwenden wir, dass y_1 und y_2 Lösungen der homogenen DG (2.207) sind und setzen

$$y_j''(x) = -A(x)y_j'(x) - B(x)y_j(x), \quad j \in \{1, 2\} \quad (2.219)$$

in (2.218) ein. Das ergibt

$$W'(x) = -A(x)[y_1(x)y_2'(x) - y_2(x)y_1'(x)] = -A(x)W(x). \quad (2.220)$$

Dies ist eine lineare Differentialgleichung 1. Ordnung, die sich durch Separation der Variablen lösen lässt. Teilen wir also (2.220) durch $W(x)$, erhalten wir

$$\frac{W'(x)}{W(x)} = -A(x). \quad (2.221)$$

Integration dieser Gleichung nach x von $x = x_0$ bis x , liefert

$$\ln \left(\frac{W(x)}{W(x_0)} \right) = - \int_{x_0}^x d\tilde{x} A(\tilde{x}). \quad (2.222)$$

Lösen wir dies nach $W(x)$ auf, finden wir schliesslich

$$W(x) = W(x_0) \exp \left[- \int_{x_0}^x d\tilde{x} A(\tilde{x}) \right]. \quad (2.223)$$

Das bedeutet aber, dass entweder $W(x) = 0 = \text{const}$ ist (nämlich wenn $W(x_0) = 0$) oder $W(x) \neq 0$ für alle $x \in \mathbb{R}$ gilt.

Untersuchen wir deshalb weiter, was es für die Lösungen y_1 und y_2 der DG bedeutet, wenn $W(x) = 0$ für alle x gilt. Aus der Definition der Wronski-Determinante (2.217) folgt dann

$$y_1(x)y_2'(x) - y_2(x)y_1'(x) = 0 \Rightarrow \frac{y_1'(x)}{y_1(x)} = \frac{y_2'(x)}{y_2(x)}. \quad (2.224)$$

Auch diese Gleichung können wir wieder bzgl. x von x_0 bis x integrieren, und das ergibt

$$\ln \left(\frac{y_1(x)}{y_1(x_0)} \right) = \ln \left(\frac{y_2(x)}{y_2(x_0)} \right) \Rightarrow y_2(x) = \frac{y_2(x_0)}{y_1(x_0)} y_1(x). \quad (2.225)$$

Das bedeutet aber, dass $W(x_0) = 0$ genau dann, wenn $y_2(x) = C y_1(x)$ mit $C = y_2(x_0)/y_1(x_0) = \text{const}$ ist.

Die Anfangsbedingungen (2.211) sind also durch die Linearkombination (2.210) genau dann immer erfüllbar, wenn die Lösungen y_1 und y_2 **linear unabhängig** sind und daher $W(x_0) \neq 0$ ist. Um also die **allgemeine Lösung der homogenen DG** zu finden, müssen wir nur zwei linear unabhängige Lösungen finden. Die allgemeine Lösung ist dann durch die allgemeine Linearkombination (2.210) gegeben.

Beispiel 32. Die beiden Funktionen $f_1 = x$ und $f_2 = 1$ sind linear unabhängig, da ja gilt:

$$W(x) = \det \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix} = -1 \neq 0. \quad (2.226)$$

Die beiden Funktionen $f_1 = x^2$ und $f_2 = \lambda x^2$ mit $\lambda \in \mathbb{R}$ sind linear abhängig, da ja gilt:

$$W(x) = \det \begin{pmatrix} x^2 & \lambda x^2 \\ 2x & 2\lambda x \end{pmatrix} = 0. \quad (2.227)$$

Kommen wir nun auf die inhomogene Gleichung (2.206) zurück. Nehmen wir wieder an, dass $y_1^{(\text{inh})}(x)$ und $y_2^{(\text{inh})}(x)$ Lösungen der inhomogenen Gleichung sind. Wegen der Linearität der Ableitungsoperation und der Linearität der linken Seite der inhomogenen Gleichung, erfüllt dann offenbar

$y_1^{(\text{inh})}(x) - y_2^{(\text{inh})}(x)$ die *homogene DG*. Das bedeutet aber, dass bei Kenntnis von zwei linear unabhängigen Lösungen $y_1^{(\text{hom})}(x)$ und $y_2^{(\text{hom})}(x)$ der *homogenen DG* diese Differenz durch eine Linearkombination dieser Lösungen gegeben sein muss. Es ist also die *allgemeine* Lösung der inhomogenen Gleichung durch

$$y(x) = C_1 y_1^{(\text{hom})}(x) + C_2 y_2^{(\text{hom})}(x) + y_p^{(\text{inh})}(x), \quad (2.228)$$

gegeben. Haben wir also die allgemeine Lösung der homogenen Gleichung gefunden, genügt es, nur eine einzige partikuläre Lösung der inhomogenen DG zu kennen, um alle Lösungen in der Form (2.228) angeben zu können. Es ist klar, dass auch hier die Anfangsbedingungen (2.211) durch die entsprechende Berechnung der Integrationskonstanten C_1 und C_2 stets erfüllbar sind, wenn nur $y_1^{(\text{hom})}(x)$ und $y_2^{(\text{hom})}(x)$ linear unabhängig sind und also die Wronski-Determinante $W(x_0) \neq 0$ ist.

Strategie zur Bestimmung der allgemeinen Lösung

(i) Die *allgemeine* Lösung der inhomogenen Gleichung, $Ly = f$, ist gegeben durch

$$y(x) = y_a^{(\text{hom})}(x) + y_p^{(\text{inh})}(x). \quad (2.229)$$

(ii) $Ly = 0$ führt zu zwei Integrationskonstanten (da 2. Ordnung). Die allgemeine Lösung ist von der Form

$$y_a(x) = C_1 y_1(x) + C_2 y_2(x). \quad (2.230)$$

(iii) y_1, y_2 sind unabhängig und bilden eine Basis (Fundamentalsystem) und $W(x) \neq 0$. Im allgemeinen Fall einer DG n -ter Ordnung, ist das Fundamentalsystem ein Satz von n linear unabhängigen partikulären Lösungen:

$$F = \left\{ y_1(x), \dots, y_n(x) \right\}. \quad (2.231)$$

(iv) Sei $y_1(x)$ eine Lösung der homogenen DG. Die zweite Lösung $y_2(x)$ kann durch Variation der Konstanten (vgl. 2.11) bestimmt werden.

(v) Seien $y_1(x), y_2(x)$ unabhängige Lösungen der linearen (homogenen) DG. Eine partikuläre Lösung, y_p , der inhomogenen DG ($Ly = f$) kann wieder durch Variation der Konstanten gefunden werden.

Beispiel 33 (iv). Betrachte die DG 2. Ordnung

$$y'' + \frac{1}{x}y' - \frac{1}{x^2}y = 0. \quad (2.232)$$

eine Lösung ist $y_1(x) := x$ ($y_1' = 1, y_1'' = 0$). Wir wählen den Ansatz

$$y_2(x) = C(x)x, \quad y_2'(x) = C(x) + C'(x)x, \quad y_2''(x) = 2C'(x) + C''(x)x, \quad (2.233)$$

und setzen diesen in die DG ein

$$2C'(x) + C''(x)x + \frac{1}{x}C(x) + C'(x) - \frac{1}{x}C(x) = 0 \quad (2.234)$$

$$C''(x) = -\frac{3}{x}C'(x). \quad (2.235)$$

Sei $z := C'(x)$, dann

$$\frac{dz}{z} = -3\frac{dx}{x} \Leftrightarrow d \ln |z| = -3(d \ln |x|). \quad (2.236)$$

Daher folgt $z = x^{-3}$, da eine partikuläre Lösung genügt. Wir integrieren C

$$C(x) = \int^x d\tilde{x} \frac{dC}{d\tilde{x}} = -\frac{1}{2x^2}, \quad (2.237)$$

und setzen das Ergebnis in den Ansatz (2.233) von y_2 ein

$$y_2 = C(x)x = \frac{\text{const}}{x}, \quad (2.238)$$

Da $W(x) \sim x \times \frac{-1}{x^2} - \frac{1}{x} \times 1 = -\frac{2}{x} \neq 0$ für $x \neq 0$, ist das Fundamentalsystem $F = \{x, \frac{1}{x}\}$ und die allgemeine Lösung:

$$y_a(x) = C_1x + \frac{1}{x}C_2. \quad (2.239)$$

Beispiel 34 (v). Um eine partikuläre Lösung zu finden, wählen wir als Ansatz zwei Funktionen

$$y_p(x) = C_1(x)y_1(x) + C_2(x)y_2(x), \quad (2.240)$$

$$y_p'(x) = C_1'(x)y_1(x) + C_2'(x)y_2(x) + C_1(x)y_1'(x) + C_2(x)y_2'(x), \quad (2.241)$$

$$y_p''(x) = C_1(x)y_1''(x) + C_2(x)y_2''(x) + C_1'(x)y_1'(x) + C_2'(x)y_2'(x), \quad (2.242)$$

wo wir im zweiten Schritt

$$C_1'(x)y_1(x) + C_2'(x)y_2(x) = 0 \quad (2.243)$$

wählen konnten, da noch immer eine freie Funktion übrig bleibt. Wir setzen in die DG, $(y_p'' + py_p' + qy_p = f)$, ein und benutzen die homogenen Lösungen $y_1'' + py_1' + qy_1 = y_2'' + py_2' + qy_2 = 0$, sodass folgt, dass

$$C_1'(x)y_1'(x) + C_2'(x)y_2'(x) = f(x). \quad (2.244)$$

Konstruiert man nun das resultierende System

$$y_1'(2.243) - y_1(2.244) \implies C_2' = \frac{-y_1f}{y_1'y_2 - y_2'y_1}, \quad (2.245)$$

$$y_2'(2.243) - y_2(2.244) \implies C_1' = \frac{-y_2f}{y_1y_2' - y_2y_1'}, \quad (2.246)$$

können nun die variierenden Funktionen C_1, C_2 integriert werden. Die jeweiligen Nenner sind stets endlich, da aufgrund von (iii) folgt, dass $W(x) \neq 0$.

2.13. Lineare Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten

Referenzen: [1] Abschn. (6.3.2), [2] Abschn. 5.5

Vereinfachen wir nun die Problemstellung und setzen $p(x), q(x) \rightarrow \text{const}$:

$$y''(x) + py'(x) + qy(x) = f(x). \quad (2.247)$$

In diesem Fall können die Lösungen y_1, y_2 der homogenen DG ($Ly = 0$) ermittelt werden. Mit dem Ansatz $y(x) = e^{\lambda x}$ wird die homogene DG zu

$$\begin{aligned} (\lambda^2 + p\lambda + q)e^{\lambda x} &= 0, \\ \lambda^2 + p\lambda + q &= 0, \end{aligned} \quad (2.248)$$

da $e^{\lambda x} \neq 0$. Dies ist die **charakteristische Gleichung**. Es gibt drei unterschiedliche Fälle, abhängig von den Wurzeln der charakteristischen Gleichung (2.248)

(i) **Alle Wurzeln der charakteristischen Gleichung sind ungleiche reelle Zahlen.**

Falls $q < \frac{p^2}{4}$, gibt es zwei Lösungen: $\lambda = \lambda_{12} = \lambda_{\pm} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$. Wir können $y_1(x) = e^{\lambda_1 x}$, $y_2(x) = e^{\lambda_2 x}$ voneinander unabhängig wählen, *viz.*

$$\frac{y_2}{y_1} = e^{(\lambda_2 - \lambda_1)x} = e^{-2x\sqrt{\frac{p^2}{4} - q}} \neq \text{const}. \quad (2.249)$$

Das Fundamentalsystem ist also

$$F = \left\{ e^{\lambda_1 x}, e^{\lambda_2 x}, \dots, e^{\lambda_n x} \right\}, \quad (2.250)$$

falls alle Wurzeln λ_i ungleich sind.

(ii) **Einige reelle Wurzeln der charakteristischen Gleichung sind k -fach.**

Falls $q = \frac{p^2}{4}$, gibt es die (zweifach entartete) Lösung $y_1(x) = e^{-\frac{p}{2}x}$ mit $\lambda_{12} = -\frac{p}{2}$. Die zweite, unabhängige, Lösung $y_2(x)$ muss erst ermittelt werden. Dazu benutzen wir die Variation der Konstanten:

$$\begin{aligned} y_2(x) &= C(x)e^{-\frac{p}{2}x}, \\ y_2'(x) &= C'(x)e^{-\frac{p}{2}x} - \frac{p}{2}C(x)e^{-\frac{p}{2}x}, \\ y_2''(x) &= C''(x)e^{-\frac{p}{2}x} - pC'(x)e^{-\frac{p}{2}x} + \frac{p^2}{4}C(x)e^{-\frac{p}{2}x}. \end{aligned} \quad (2.251)$$

Nach Einsetzen in Gl. (2.247) folgt

$$y_2'' + py_2' + \frac{p^2}{4}y_2 = \left[C'' - pC' + \frac{p^2}{4}C + pC' - \frac{p^2}{2}C + \frac{p^2}{4}C \right] e^{-\frac{p}{2}x} = 0. \quad (2.252)$$

Daraus folgt, dass $C'' = 0$, $C' = \alpha$, und $C = \alpha x + \beta$. Nach Einsetzen in den Ansatz sieht man, dass der α Term eine neue Lösung ergibt: $y_2(x) = xe^{-\frac{p}{2}x}$, und das (allgemeine) Fundamentalsystem ist

$$F = \left\{ e^{-\frac{p}{2}x}, xe^{-\frac{p}{2}x} \right\}, \quad (2.253)$$

$$F = \left\{ \underbrace{e^{\lambda_1 x}, xe^{\lambda_1 x}, \dots, x^{k-1}e^{\lambda_1 x}}_{k\text{-fache Wurzel } \lambda_1}, \underbrace{e^{\lambda_2 x}, \dots, e^{\lambda_n x}}_{\text{einfache Wurzeln}} \right\}. \quad (2.254)$$

(iii) **Einige Wurzeln der charakteristischen Gleichung sind komplex.**

Falls $q > \frac{p^2}{4}$, benutzen wir den Ansatz

$$\begin{aligned} y_1(x) &= e^{-\frac{p}{2}x} \cos\left(\sqrt{q - \frac{p^2}{4}}x\right), \\ y_2(x) &= e^{-\frac{p}{2}x} \sin\left(\sqrt{q - \frac{p^2}{4}}x\right). \end{aligned} \quad (2.255)$$

Das allgemeine Fundamentalsystem mit den komplexen Wurzeln $\lambda_1 = a + ib$, $\lambda_2 = a - ib$ ist dann

$$F = \left\{ \underbrace{e^{ax} \cos bx, e^{ax} \sin bx}_{\text{komplexe Wurzeln } \lambda_1, \lambda_2}, \underbrace{e^{\lambda_3 x}, \dots, e^{\lambda_n x}}_{\text{einfache Wurzeln}} \right\}. \quad (2.256)$$

Physikalisch entspricht dies gedämpften ($p > 0$) oder verstärkenden ($p < 0$) Schwingungen.

Beweis. OBdA folgt für y_1 (gleich für y_2), dass

$$\begin{aligned} y_1' &= -\frac{p}{2}e^{-\frac{p}{2}x} \cos[\dots] - \sqrt{q - \frac{p^2}{4}}e^{-\frac{p}{2}x} \sin[\dots], \\ y_1'' &= \frac{p^2}{4}e^{-\frac{p}{2}x} \cos[\dots] + p\sqrt{q - \frac{p^2}{4}}e^{-\frac{p}{2}x} \sin[\dots] + \left(\frac{p^2}{4} - q\right)e^{-\frac{p}{2}x} \cos[\dots], \end{aligned} \quad (2.257)$$

$$\begin{aligned} y_1'' + py_1' + qy_1 &= e^{-\frac{p}{2}x} \cos[\dots] \left[\frac{p^2}{4} + \frac{p^2}{4} - q - \frac{p^2}{2} + q \right] \\ &\quad + e^{-\frac{p}{2}x} \sin[\dots] \left[p\sqrt{q - \frac{p^2}{4}} - p\sqrt{q - \frac{p^2}{4}} \right] = 0. \end{aligned} \quad (2.258)$$

Für alle Fälle (i)–(iii) können zwei linear unabhängige Lösungen gefunden werden. Damit haben wir die allgemeine Lösung der homogenen Gleichung gefunden. Der nächste Schritt ist die partikuläre Lösung der inhomogenen Gleichung $Ly = f$. Diese Lösung kann wiederum durch die Variation der Konstanten (vgl. Bsp. 34) gefunden werden. Praktisch findet man eine partikuläre Lösung aber auch durch Erraten eines angemessenen Ansatzes.

Beispiel 35 (Getriebener Harmonischer Oszillator). *Durch die Wahl der DG*

$$y'' + 2ky' + w_0^2 y = b \sin(wx), \quad (2.259)$$

sind $q = w_0^2$, $p = 2k$, $\frac{p^2}{4} = k^2$, wobei die Terme $2ky'$ eine Dämpfung (für $k > 0$) und $b\sin(wx)$ eine Treibung beschreiben. Falls $k^2 < w_0^2$ diskutieren wir Fall (iii), mit einer einfach-komplexen Wurzel. Die allgemeine Lösung der homogenen DG ist daher

$$y_a(x) = C_1 e^{-kx} \cos\left(\sqrt{w_0^2 - k^2}x\right) + C_2 e^{-kx} \sin\left(\sqrt{w_0^2 - k^2}x\right). \quad (2.260)$$

Für die partikuläre Lösung, $y_p(x)$, nehmen wir den Ansatz:

$$\begin{aligned} y_p(x) &= A \cos(wx) + B \sin(wx), \\ y_p'(x) &= -wA \sin(wx) + wB \cos(wx), \\ y_p''(x) &= -w^2 A \cos(wx) - w^2 B \sin(wx), \end{aligned} \quad (2.261)$$

$$\begin{aligned} y_p'' + py_p' + qy_p &= \cos(wx) \left[-w^2 A + 2kwB + w_0^2 A \right] \\ &\quad + \sin(wx) \left[-w^2 B + 2kwA + w_0^2 B \right] = b \sin(wx). \end{aligned} \quad (2.262)$$

In dem man jeweils nur Terme vergleicht die proportional zu $\cos(wx)$ oder $\sin(wx)$ sind kann man die Unbekannten A und B finden

$$B = \frac{(w^2 - w_0^2)A}{2kw}, \quad (w_0^2 - w^2)B - 2kwA = b. \quad (2.263)$$

Folglich sind

$$\begin{aligned} A &= -\frac{2kwb}{(w^2 - w_0^2)^2 + 4k^2w^2}, \\ B &= -\frac{(w^2 - w_0^2)b}{(w^2 - w_0^2)^2 + 4k^2w^2}. \end{aligned} \quad (2.264)$$

Daher ist die allgemeine Lösung der vollen DG (2.247)

$$\begin{aligned} y(x) &= e^{-kx} \left[C_1 \cos\left(\sqrt{w_0^2 - k^2}x\right) + C_2 \sin\left(\sqrt{w_0^2 - k^2}x\right) \right] \\ &\quad - \frac{b}{(w^2 - w_0^2)^2 + 4k^2w^2} \left[2kw \cos(wx) + (w^2 - w_0^2) \sin(wx) \right]. \end{aligned} \quad (2.265)$$

Sollten die Koeffizienten nicht konstant sein, kann man oft Lösung durch einen Potenzreihenansatz finden, e.g. $y(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$.

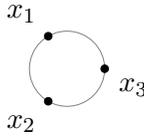


Figure 12: Diskretisierte lineare Differentialgleichung auf S^1 , dem Einheitskreis.

3. Lineare Algebra

3.1. Vektorräume, lineare Abbildungen

Referenzen: [1] Abschn. (3.3.1), [2] Abschn. 1.1

Typische Beispiele für lineare Operationen sind:

$$\text{Lineare Differentialgleichung:} \quad Ly(x) = f(x), \quad (3.1)$$

$$\text{Lineare Abbildung:} \quad M\mathbf{v} = \mathbf{w}, \quad (3.2)$$

dabei sei M eine lineare Abbildung oder Matrix, und \mathbf{v}, \mathbf{w} Vektoren.

Beispiel 36. Betrachten wir eine diskretisierte lineare Differentialgleichung bzw. Differenzgleichung

$$Ly(x) := \frac{dy}{dx} \rightarrow \frac{y(x+a) - y(x-a)}{2a}, \quad (3.3)$$

wobei $x_2 = x_1 + a$, $x_3 = x_1 - a$; siehe Abb. 12. Das resultierende Gleichungssystem hat die folgende Form

$$\begin{cases} \frac{y(x_1+a) - y(x_1-a)}{2a} = f(x_1) \\ \frac{y(x_2+a) - y(x_2-a)}{2a} = f(x_2) \\ \frac{y(x_3+a) - y(x_3-a)}{2a} = f(x_3) \end{cases} \Leftrightarrow \begin{cases} +\frac{1}{2a}y(x_2) - \frac{1}{2a}y(x_3) = f(x_1) \\ -\frac{1}{2a}y(x_1) + \frac{1}{2a}y(x_3) = f(x_2) \\ +\frac{1}{2a}y(x_1) - \frac{1}{2a}y(x_2) = f(x_3) \end{cases}, \quad (3.4)$$

und kann in eine vektorisierte Schreibweise $M\mathbf{v} = \mathbf{w}$ übertragen werden

$$\begin{pmatrix} 0 & +\frac{1}{2a} & -\frac{1}{2a} \\ -\frac{1}{2a} & 0 & +\frac{1}{2a} \\ +\frac{1}{2a} & -\frac{1}{2a} & 0 \end{pmatrix} \begin{pmatrix} y(x_1) \\ y(x_2) \\ y(x_3) \end{pmatrix} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \end{pmatrix}. \quad (3.5)$$

Beispiel 37. Die Partialbruchzerlegung bzw. lineare Gleichungssysteme der Form

$$\begin{aligned} \frac{1}{(1+x)^2x} &= \frac{A}{(1+x)^2} + \frac{B}{(1+x)} + \frac{C}{x} \\ &= \frac{1}{(1+x)^2x} \left\{ xA + x(1+x)B + (1+x)^2C \right\} \\ &= \frac{1}{(1+x)^2x} \left\{ x(A+B+2C) + x^2(B+C) + C \right\}, \end{aligned} \quad (3.6)$$

können zur Bestimmung von A, B, C auch in eine vektorisierte Schreibweise, $M\mathbf{v} = \mathbf{w}$, übertragen werden

$$\begin{cases} A + B + 2C & = 0 \\ B + C & = 0 \\ C & = 1 \end{cases} \Leftrightarrow \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.7)$$

Die Vektoren dieser Konstruktion bilden einen **Vektorraum**, V . Die Rechenoperationen in V , müssen nun noch weiter spezifiziert werden. Dabei ist es aus mathematischer Sicht unerheblich, um welche konkreten Objekte es sich bei dem, was wir *Vektoren* nennen, eigentlich handelt. Es werden einfach *Spielregeln*, die **Axiome**, aufgestellt, die formale Operationen festlegen. Für die **Addition** von Vektoren sind dies die folgenden Axiome:

(v1) **Assoziativgesetz**. Für beliebige drei Vektoren $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$ gilt stets, dass

$$(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3). \quad (3.8)$$

(v2) **Neutrales Element** der Addition. Es existiert ein neutrales Element, das wir mit $\mathbf{0} \in V$ bezeichnen. Es zeichnet sich dadurch aus, dass für alle Vektoren gilt, dass $\mathbf{v} \in V$

$$\mathbf{v} + \mathbf{0} = \mathbf{v}. \quad (3.9)$$

(v3) **Inverses Element**. Zu jedem Vektor $\mathbf{v} \in V$ existiert ein bzgl. der Addition $(-\mathbf{v}) \in V$, für das

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0}. \quad (3.10)$$

(v4) **Kommutativgesetz**. Für beliebige zwei Vektoren $\mathbf{v}_1, \mathbf{v}_2 \in V$ gilt, dass

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1. \quad (3.11)$$

Es führt für diese Physikvorlesung zu weit, den gesamten Apparat der linearen Algebra, aus diesen Axiomen zu beweisen. Dies ist Gegenstand der entsprechenden Mathematikvorlesung. Als ein Beispiel möge der folgende Satz dienen:

Satz 3.1. *Das neutrale Element der Addition $\mathbf{0}$ ist eindeutig.*

Beweis. Nehmen wir dazu an, es sei auch $\mathbf{0}'$ ein neutrales Element der Addition. Wegen der Eigenschaft (3.9) des neutralen Elements $\mathbf{0}$ gilt

$$\mathbf{0}' = \mathbf{0}' + \mathbf{0} \stackrel{(3.11)}{=} \mathbf{0} + \mathbf{0}'. \quad (3.12)$$

Nun ist aber $\mathbf{0}'$ voraussetzungsgemäss ebenfalls ein neutrales Element ist, und gemäss (3.9) gilt demnach

$$\mathbf{0}' = \mathbf{0} + \mathbf{0}' = \mathbf{0}, \quad (3.13)$$

d.h. dass notwendigerweise die beiden neutralen Elemente der Addition gleich sind. \square

Eine algebraische Struktur, die auf einer Menge definiert ist wird als **Gruppe** bezeichnet, wenn die Axiome (3.8)–(3.10) erfüllt sind. Hier wurden diese Axiome für die Addition definiert. So einfach diese Axiome auch anmuten mögen, so liefern sie doch eine sehr reichhaltige mathematische Struktur. Die Gruppentheorie ist ein riesiges Teilgebiet der Mathematik, und Teile der Gruppentheorie spielen auch für die Physik eine herausragende Rolle, wie im Verlauf des Physikstudiums noch klar wird. Nimmt man zu diesen drei Axiomen noch das Kommutativgesetz (3.11) hinzu, spricht man von einer **Abelschen Gruppe**. Wir können also die Axiome der Vektoraddition auch kurz zusammenfassen, indem wir bemerken, dass die Vektoraddition eine Abelsche Gruppe auf der Menge der Vektoren V bildet.

Um die so definierte additive Gruppe zu einem Vektorraum zu erweitern, benötigen wir eine zweite Operation, die Multiplikation von Vektoren mit einem Skalar. Diese erfüllt die folgenden Axiome für alle $\mathbf{v}, \mathbf{w} \in V$ und $\lambda, \mu \in \mathbb{R}$:

$$(s1) \quad (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}.$$

$$(s2) \quad \lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}.$$

$$(s3) \quad \lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}.$$

$$(s4) \quad \text{Neutrales Element}$$

$$1\mathbf{v} = \mathbf{v}.$$

Dabei seien die Skalare Elemente eines **Zahlenkörpers**, der in unserem Fall stets der Körper der reellen Zahlen \mathbb{R} ist.

Mit diesen Rechenregeln ist der Begriff des Vektorraumes bereits vollständig bestimmt. Als wichtige Begriffe ergeben sich daraus, die **Linearkombinationen** und **Basen** und im Zusammenhang damit auch der **Dimension** eines Vektorraums. Wir werden unten als weitere wichtige Strukturen die **linearen Abbildungen** verschiedener Art kennenlernen, deren Untersuchung weitere Eigenschaften der Vektorräume charakterisieren. Es ist übrigens ohnehin ein Merkmal der modernen axiomatischen Methode der Mathematik, Strukturen einzuführen und dann entsprechende diese Strukturen charakterisierende Abbildungen zwischen diesen Strukturen zu analysieren.

Nach mehrmaliger Addition und Skalarmultiplikation, kann man eine *Linearkombination* bilden:

$$\sum_{i=1}^k a_i \mathbf{v}_i \in V, \quad (3.14)$$

deren Vektoren **linear unabhängig** sind, falls gilt, dass

$$\sum_{i=1}^k a_i \mathbf{v}_i = \mathbf{0} \Leftrightarrow a_i = 0 \quad \forall i \in \{1, \dots, k\}. \quad (3.15)$$

Falls dagegen $a_j \neq 0$ gefunden werden kann, kann \mathbf{v}_j als Linearkombination von $\mathbf{v}_i, i \neq j$ ausgedrückt werden, *viz.* $\mathbf{v}_j = -\frac{1}{a_j} \sum_{i \neq j} a_i \mathbf{v}_i$.

Falls die Vektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$

- (i) linear unabhängig sind,
- (ii) den Vektorraum V aufspannen indem $\forall \mathbf{v} \in V$ als Linearkombination von $\mathbf{e}_1, \dots, \mathbf{e}_n$ ausgedrückt werden kann,

dann bilden sie eine **Basis** bzw. ein **Fundamentalsystem** (vgl. DG). Die **Dimension** des Vektorraums ist dann $\dim V = n$, und gleich der maximalen Anzahl an linear unabhängigen Vektoren.

Nach der Wahl einer Basis kann jeder Vektor eindeutig als $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$ ausgedrückt werden.

Beweis.

$$\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i = \sum_{i=1}^n w_i \mathbf{e}_i, \Rightarrow \mathbf{0} = \mathbf{v} - \mathbf{v} = \sum_{i=1}^n (v_i - w_i) \mathbf{e}_i, \Rightarrow v_i = w_i, \quad \forall i = 1, \dots, n. \quad (3.16)$$

Die Wahl der Basis selbst ist daher nicht eindeutig.

Beispiel 38. Falls $\{\mathbf{e}_1, \mathbf{e}_2\}$ eine Basis ist, ist $\{\mathbf{e}'_1 = \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}'_2 = \mathbf{e}_1 - \mathbf{e}_2\}$ eine andere Möglichkeit, wie durch die Betrachtung von (i) und (ii) gezeigt werden kann.

Der Raum, \mathbb{R}^3 , ist auch ein Vektorraum mit den Basisvektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \rightarrow \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z \rightarrow \mathbf{i}, \mathbf{j}, \mathbf{k}$.

3.2. Skalarprodukt

Referenzen: [1] Abschn. (3.3.2), [2] Abschn. 1.2

In vielen Vektorräumen kann die Struktur eines *Skalarprodukts* definiert werden. Dieses ist eine Abbildung $V \times V \rightarrow \mathbb{R}$, $(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \cdot \mathbf{w}$, mit dem folgende Axiome für alle $\mathbf{v}, \mathbf{w}, \mathbf{w}_1, \mathbf{w}_2 \in V$, erfüllt werden:

- (i) **Linearität**
Für alle $\lambda_1, \lambda_2 \in \mathbb{R}$ gilt $\mathbf{v} \cdot (\lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2) = \lambda_1 \mathbf{v} \cdot \mathbf{w}_1 + \lambda_2 \mathbf{v} \cdot \mathbf{w}_2$.
- (ii) **Kommutativgesetz**
 $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.
- (iii) Mit der Norm $|\mathbf{v}| := \sqrt{\mathbf{v} \cdot \mathbf{v}}$ gilt $|\mathbf{v}|^2 := \mathbf{v} \cdot \mathbf{v} \geq 0$.
- (iv) Es ist $|\mathbf{v}|^2 = 0$ genau dann wenn, $\mathbf{v} = \mathbf{0}$ ist.

Eine Abbildung $V \times V \rightarrow V$, welche die Axiome (i) und (ii) erfüllt, heisst **symmetrische Bilinearform**. Gilt auch noch Axiom (iii), heisst die Bilinearform **positiv semidefinit**. Ist schliesslich auch Axiom (iv) erfüllt, heisst die Bilinearform **positiv definit**. Aus diesen Axiomen resultieren folgende Ungleichungen:

Satz 3.2 (Cauchy-Schwarzche Ungleichung).

$$|\mathbf{v} \cdot \mathbf{w}| \leq |\mathbf{v}| |\mathbf{w}|. \quad (3.17)$$

Beweis.

$$0 \stackrel{\text{(iii)}}{\leq} (\lambda \mathbf{v} + \mathbf{w}) \cdot (\lambda \mathbf{v} + \mathbf{w}) \stackrel{\text{(i)}}{=} \lambda^2 |\mathbf{v}|^2 + 2\lambda \mathbf{v} \cdot \mathbf{w} + |\mathbf{w}|^2. \quad (3.18)$$

Wählt man nun $\lambda = -\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|^2}$ und multipliziert mit $|\mathbf{v}|^2$, dann

$$\begin{aligned} 0 &\leq (\mathbf{v} \cdot \mathbf{w})^2 - 2(\mathbf{v} \cdot \mathbf{w})^2 + |\mathbf{v}|^2 |\mathbf{w}|^2, \\ (\mathbf{v} \cdot \mathbf{w})^2 &\leq |\mathbf{v}|^2 |\mathbf{w}|^2, \\ |\mathbf{v} \cdot \mathbf{w}| &\leq |\mathbf{v}| |\mathbf{w}|. \end{aligned} \quad (3.19)$$

Satz 3.3 (Dreiecksungleichung).

$$|\mathbf{v} + \mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}|. \quad (3.20)$$

Beweis. Der Beweis folgt aus der positiven Definitheit des Skalarproduktes:

$$\begin{aligned} |\mathbf{v} + \mathbf{w}|^2 &= (\mathbf{v} + \mathbf{w})^2 \\ &= \mathbf{v}^2 + \mathbf{w}^2 + 2\mathbf{v} \cdot \mathbf{w} \\ &\leq \mathbf{v}^2 + \mathbf{w}^2 + 2|\mathbf{v} \cdot \mathbf{w}| \\ &\stackrel{\text{(3.17)}}{\leq} \mathbf{v}^2 + \mathbf{w}^2 + 2|\mathbf{v}| |\mathbf{w}| = (|\mathbf{v}| + |\mathbf{w}|)^2, \end{aligned} \quad (3.21)$$

wobei wir im letzten Schritt die Cauchy-Schwarz Ungleichung (3.17) benutzt haben. Weil immer nur positive Zahlen quadriert werden, folgt

$$|\mathbf{v} + \mathbf{w}| \leq |\mathbf{v}| + |\mathbf{w}|. \quad (3.22)$$

Zwei Vektoren sind *orthogonal* zueinander, falls $\mathbf{v} \cdot \mathbf{w} = 0$. Eine Basis ist orthogonal, falls $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ für $\forall i \neq j$. Ist aber auch $|\mathbf{e}_i| = 1 \forall i \in \{1, \dots, n\}$, ist dies eine *orthonormale Basis*. In einer orthonormalen Basis gilt

$$\mathbf{v} \cdot \mathbf{w} = \left(\sum_{i=1}^n v_i \mathbf{e}_i \right) \cdot \left(\sum_{j=1}^n w_j \mathbf{e}_j \right) = \sum_{i,j=1}^n v_i w_j \mathbf{e}_i \cdot \mathbf{e}_j = \sum_{i=1}^n v_i w_i, \quad (3.23)$$

wobei $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ durch das sogenannte Kronecker-Delta beschrieben werden kann:

$$\delta_{ij} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases} \quad (3.24)$$

Einen Winkel zwischen zwei Vektoren kann man jetzt durch

$$\cos \theta := \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}| |\mathbf{w}|}, \quad (3.25)$$

definieren; wegen der Cauchy-Schwarzschen Ungleichung (3.17) gilt $-1 \leq \cos \theta \leq 1$.

Damit können wir die Definition von linearen Abbildungen bzw. Matrizen weiter präzisieren:

$$M : V \rightarrow V, \quad \mathbf{v} \mapsto \mathbf{w} = M\mathbf{v}. \quad (3.26)$$

Dazu benutzen wir die orthonormale Basis $\{\mathbf{e}_i\} : \mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$. Durch die Linearität folgt, dass

$$\mathbf{w} = \sum_{k=1}^n w_k \mathbf{e}_k = M\mathbf{v} = \sum_{j=1}^n v_j M\mathbf{e}_j. \quad (3.27)$$

Nimmt man das Skalarprodukt mit \mathbf{e}_i und benutzt dass $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, dann

$$w_i = \sum_{j=1}^n v_j \underbrace{\mathbf{e}_i \cdot M\mathbf{e}_j}_{=: M_{ij}} = \sum_{j=1}^n M_{ij} v_j. \quad (3.28)$$

Indem i Zeilen und j Spalten indiziert, lautet die Rechenvorschrift für $\mathbf{w} = M\mathbf{v}$ ausführlich geschrieben

$$\begin{pmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} M_{11} & \dots & M_{1j} & \dots & M_{1n} \\ \vdots & & \vdots & & \vdots \\ M_{i1} & \dots & M_{ij} & \dots & M_{in} \\ \vdots & & \vdots & & \vdots \\ M_{n1} & \dots & M_{nj} & \dots & M_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} M_{11}v_1 + \dots + M_{1n}v_n \\ \vdots \\ M_{i1}v_1 + \dots + M_{in}v_n \\ \vdots \\ M_{n1}v_1 + \dots + M_{nn}v_n \end{pmatrix}. \quad (3.29)$$

Beispiel 39. *Besondere Beispiele sind:*

$$\text{Nullmatrix:} \quad 0_{n \times n} := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \quad 0_{n \times n} \mathbf{v} = \mathbf{0} \quad \forall \mathbf{v}, \quad (3.30)$$

$$\text{Einheitsmatrix:} \quad \mathbb{1}_{n \times n} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \mathbb{1}_{n \times n} \mathbf{v} = \mathbf{v} \quad \forall \mathbf{v}. \quad (3.31)$$

Matrizen bilden selbst einen abstrakten Vektorraum:

$$(M + N)_{ij} = M_{ij} + N_{ij}, \quad (3.32)$$

$$(aM)_{ij} = aM_{ij}, \quad (3.33)$$

$$(0)_{ij} = 0, \quad \forall i, j, \quad (3.34)$$

$$(-M)_{ij} = -M_{ij}, \quad (3.35)$$

mit der zusätzlichen Struktur der *Komposition* oder Matrix-Multiplikation

$$(M \circ N)_{ij} = \sum_{k=1}^n M_{ik} N_{kj}. \quad (3.36)$$

Beweis.

$$(M \circ N)_{ij} = \mathbf{e}_i \cdot (M \circ N) \mathbf{e}_j = \mathbf{e}_i \cdot (M \underbrace{(N \mathbf{e}_j)}_{\sum_k \mathbf{e}_k N_{kj}}) = \sum_k \underbrace{\mathbf{e}_i \cdot M \mathbf{e}_k}_{M_{ik}} N_{kj}. \quad (3.37)$$

Die Existenz dieser Strukturen (Summe, Skalarmultiplikation, Komposition) definiert eine Algebra.

3.3. Transponierung, Spur, Determinante

Referenzen: [1] Abschn. (3.1), [2] Abschn. 1.6

Nach Wahl einer Basis kann eine lineare Abbildung als Matrix dargestellt werden:

$$M \mathbf{v} = \mathbf{w}, \quad M_{ij} := (M)_{ij} := \mathbf{e}_i \cdot M \mathbf{e}_j = \langle e_i | M | e_j \rangle, \quad (3.38)$$

wobei die letzte Notation in der Physik als *Dirac-Notation* bekannt ist. Im folgenden beschränken wir uns auf quadratische Matrizen.

Definition 3.4 (Transponierung). *Die transponierte Matrix wird als*

$$(M^T)_{ij} := (M)_{ji} = M_{ji}, \quad (3.39)$$

definiert. Für Vektoren gilt

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}^T = (v_1, \dots, v_n), \quad (3.40)$$

wobei hier die linke Seite als Spaltenvektor und die rechte Seite als Reihenvektor bezeichnet wird. Es gilt dann:

$$\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^n v_i w_i = (v_1, \dots, v_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{v}^T \mathbf{w}. \quad (3.41)$$

Eine Matrix ist

$$\text{symmetrisch} \quad M^T = M, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (3.42)$$

$$\text{antisymmetrisch} \quad M^T = -M, \quad \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (3.43)$$

Satz 3.5. (ohne Beweis)

$$(M^T)^T = M, \quad (3.44)$$

$$(M + N)^T = M^T + N^T, \quad (3.45)$$

$$(M \circ N)^T = N^T \circ M^T. \quad (3.46)$$

Beweis. Wir beweisen den letzten Satz mit der Definition (3.36)

$$(M \circ N)_{ij}^T = \sum_k M_{jk} N_{ki} = \sum_k (N^T)_{ik} (M^T)_{kj}. \quad (3.47)$$

Die transponierten Matrizen spielen eine wichtige Rolle bei Drehungen. Dabei sei eine Drehung eine lineare Abbildung, die das Skalarprodukt zwischen zwei Vektoren invariant lässt:

$$\begin{aligned} M\mathbf{x} \cdot M\mathbf{y} &= (M\mathbf{x})^T M\mathbf{y} = \mathbf{x}^T M^T M\mathbf{y} \stackrel{!}{=} \mathbf{x}^T \mathbf{y}, \quad \forall \mathbf{x}, \mathbf{y} \\ M^T M &\stackrel{!}{=} \mathbb{1}. \end{aligned} \quad (3.48)$$

Eine Matrix mit dieser Eigenschaft wird **orthogonal** genannt. Für Basisvektoren $\{\mathbf{e}_i\}$ einer orthonormalen Basis bedeutet das, dass mit einer Transformation $\mathbf{e}'_i := M\mathbf{e}_i$ folglich auch $\mathbf{e}'_j \cdot \mathbf{e}'_j = \mathbf{e}_j \cdot \mathbf{e}_j = \delta_{ij}$ falls $M^T M = \mathbb{1}$ und damit die Basisvektoren $\{\mathbf{e}'_i\}$ eine andere orthonormierte Basis aufspannen.

Definition 3.6 (Spur). Neben der Determinante gibt es noch eine zweite charakteristische Größe für eine quadratische Matrix, die so genannte Spur einer Matrix. Diese Matrixfunktion bezeichnet die Summe über alle Diagonalelemente. Sie wird abgekürzt mit $\text{Sp}A$ (für "Spur") oder $\text{Tr} A$ (für "Trace") bezeichnet:

$$\text{Tr} M := \sum_{i=1}^n M_{ii}. \quad (3.49)$$

Satz 3.7.

$$\text{Tr} M^T = \text{Tr} M, \quad \text{Tr} M^T = \sum_i (M^T)_{ii} = \sum_i M_{ii} = \text{Tr} M, \quad (3.50)$$

$$\text{Tr} (MN) = \text{Tr} (NM), \quad \text{Tr} (MN) = \sum_{i,k} M_{ik} N_{ki} = \sum_{i,k} N_{ki} M_{ik} = \text{Tr} (NM). \quad (3.51)$$

Die letzte Eigenschaft lässt sich verallgemeinern

$$\begin{aligned} \text{Tr} (M_1, M_2, \dots, M_n) &= \sum_{ijl\dots k} (M_1)_{ij} (M_2)_{jl} \dots (M_n)_{ki} \\ &= \sum_{ijl\dots k} (M_2)_{jl} \dots (M_n)_{ki} (M_1)_{ij} \\ &= \text{Tr} (M_2, \dots, M_n, M_1), \end{aligned} \quad (3.52)$$

sodass die Spur eines Produkts von Matrizen sich nicht ändert, wenn die Reihenfolge der Matrizen zyklisch verändert wird. Ausserdem ist die für $n \times n$ Matrizen $\text{Tr} \mathbb{1}_{n \times n} = n$.

Satz 3.8. Die Spur ist invariant unter einer orthogonalen Basistransformation.

Beweis. Sei $\mathbf{e}'_i = O\mathbf{e}_i$ eine orthogonale Basistransformation sodass $O^T O = O^T O = \mathbb{1}$. In einer neuen Basis können wir die Matrix dann folgendermassen ausdrücken

$$\begin{aligned} M'_{ij} &= \mathbf{e}'_i \cdot M\mathbf{e}'_j = (\mathbf{e}'_i)^T M\mathbf{e}'_j = (O\mathbf{e}_i)^T M O\mathbf{e}_j \\ &= \mathbf{e}_i^T O^T M O\mathbf{e}_j = (O^T M O)_{ij} \\ \text{Tr}[M'] &= \text{Tr}[O^T M O] = \text{Tr}[O O^T M] = \text{Tr}[M]. \end{aligned} \quad (3.53)$$

Definition 3.9 (Determinante). Die Determinante $\det M \in \mathbb{R}$ ist eine Abbildung

$$\det M = \Delta(\mathbf{m}_1, \dots, \mathbf{m}_n) \quad : \quad \underbrace{V \times V \times \dots \times V}_{n\text{-mal}} \rightarrow \mathbb{R}, \quad (3.54)$$

wozu wir die Matrix M zur Vereinfachung als Spaltenvektoren $M = (\mathbf{m}_1, \dots, \mathbf{m}_n)$ ausgedrückt haben. Die Determinante hat dann die folgenden Eigenschaften

$$\text{Antisymmetrie} \quad \Delta(\mathbf{m}_1, \dots, \mathbf{m}_i, \dots, \mathbf{m}_j, \dots, \mathbf{m}_n) = -\Delta(\mathbf{m}_1, \dots, \mathbf{m}_j, \dots, \mathbf{m}_i, \dots, \mathbf{m}_n), \quad (3.55)$$

$$\text{Linearität} \quad \Delta(\mathbf{m}_1, \dots, \lambda\mathbf{m}_i, \dots, \mathbf{m}_n) = \lambda\Delta(\mathbf{m}_1, \dots, \mathbf{m}_i, \dots, \mathbf{m}_n), \quad (3.56)$$

$$\begin{aligned} \Delta(\mathbf{m}_1, \dots, \mathbf{m}_i + \mathbf{m}_j, \dots, \mathbf{m}_n) &= \Delta(\mathbf{m}_1, \dots, \mathbf{m}_i, \dots, \mathbf{m}_n) \\ &\quad + \Delta(\mathbf{m}_1, \dots, \mathbf{m}_j, \dots, \mathbf{m}_n), \end{aligned} \quad (3.57)$$

$$\text{Normierung} \quad \Delta(\mathbf{e}_1, \dots, \mathbf{e}_n) = \det \mathbb{1} = 1. \quad (3.58)$$

Beispiel 40. Betrachte die allgemeine Determinante einer 2×2 Matrix:

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= \Delta(a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2, a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2) \\ &= a_{11}a_{12}\Delta(\mathbf{e}_1, \mathbf{e}_1) + a_{21}a_{12}\Delta(\mathbf{e}_2, \mathbf{e}_1) + a_{11}a_{22}\Delta(\mathbf{e}_1, \mathbf{e}_2) + a_{21}a_{22}\Delta(\mathbf{e}_2, \mathbf{e}_2) \\ &= (a_{11}a_{22} - a_{21}a_{12})\Delta(\mathbf{e}_1, \mathbf{e}_2) = a_{11}a_{22} - a_{21}a_{12}. \end{aligned} \quad (3.59)$$

Im dritten Schritt haben wir die Antisymmetrie von $\Delta(\mathbf{e}_1, \mathbf{e}_1) = -\Delta(\mathbf{e}_1, \mathbf{e}_1) = 0$ und $\Delta(\mathbf{e}_1, \mathbf{e}_2) = -\Delta(\mathbf{e}_2, \mathbf{e}_1)$ benutzt.

Satz 3.10 (Levi-Civita-Symbol). Sei $\varepsilon_{i_1, \dots, i_n}$ das Levi-Civita-Symbol sodass gilt

$$\varepsilon_{i_1, \dots, i_n} = \begin{cases} +1 & \text{falls } (i_1, i_2, \dots, i_n) \text{ aus } (1, 2, \dots, n) \text{ durch eine} \\ & \text{gerade Anzahl an Permutationen entsteht,} \\ -1 & \text{falls } (i_1, i_2, \dots, i_n) \text{ aus } (1, 2, \dots, n) \text{ durch eine} \\ & \text{ungerade Anzahl an Permutationen entsteht,} \\ 0 & \text{falls zwei oder mehrere Indizes ident sind.} \end{cases} \quad (3.60)$$

Beispiel 41.

$$\varepsilon_{12} = +1, \quad \varepsilon_{21} = -1, \quad \varepsilon_{11} = 0, \quad \varepsilon_{22} = 0. \quad (3.61)$$

Dann gilt für die Berechnung der Determinante die so-genannte Leibniz-Formel:

$$\det M = \sum_{i_1, \dots, i_n} \varepsilon_{i_1, \dots, i_n} M_{i_1, 1} \dots M_{i_n, n}. \quad (3.62)$$

Beweis.

$$\begin{aligned} \Delta(\mathbf{m}_1, \dots, \mathbf{m}_n) &= \sum_{i_1, \dots, i_n} \Delta(M_{i_1, 1} \mathbf{e}_{i_1}, \dots, M_{i_n, n} \mathbf{e}_{i_n}) \\ &\stackrel{(3.55)}{=} \sum_{i_1, \dots, i_n} M_{i_1, 1} \dots M_{i_n, n} \Delta(\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_n}) \\ &\stackrel{(3.56)}{=} \sum_{i_1, \dots, i_n} M_{i_1, 1} \dots M_{i_n, n} \varepsilon_{i_1, \dots, i_n} \underbrace{\Delta(\mathbf{e}_1, \dots, \mathbf{e}_n)}_{=1, \text{Norm (3.58)}}. \end{aligned} \quad (3.63)$$

Satz 3.11. Nun betrachten wir die Determinante der transponierten Matrix,

$$\det M^T = \det M. \quad (3.64)$$

Beweis. Da gilt, dass $(M^T)_{ij} = M_{ji}$ folgt

$$\det M^T = \sum_{i_1, \dots, i_n} \varepsilon_{i_1, \dots, i_n} (M^T)_{i_1, 1} (M^T)_{i_2, 2} \dots (M^T)_{i_n, n} = \sum_{i_1, \dots, i_n} \varepsilon_{i_1, \dots, i_n} M_{1, i_1} M_{2, i_2} \dots M_{n, i_n}. \quad (3.65)$$

Nun liefern gemäss der Definition des Levi-Civita-Symbols nur solche Terme einen von 0 verschiedenen Beitrag zur Determinante, für die (i_1, \dots, i_n) eine **Permutation** (d.h. Anordnung in geänderter Reihenfolge) von $(1, 2, \dots, n)$ ist, und es wird über alle möglichen dieser Anordnungen summiert. Man kann demnach im letzten Ausdruck in (3.65) genauso gut die Produkte auch bzgl. der hinteren Indizes “sortieren” und über alle Permutationen von $(1, 2, \dots, n)$ der vorderen Indizes summieren. Das ist aber gerade wieder die Determinante der ursprünglichen Matrix. \square

Satz 3.12.

$$\det(M \circ N) = \det M \det N. \quad (3.66)$$

Beweis. Die Matrix $M = (\mathbf{m}_1, \dots, \mathbf{m}_n)$ hat die Determinante $\det M = \Delta(\mathbf{m}_1, \dots, \mathbf{m}_n)$. Dadurch sind

$$\begin{aligned}
 M \circ N &= \left(\sum_{i_1} \mathbf{m}_{i_1} N_{i_1,1}, \sum_{i_2} \mathbf{m}_{i_2} N_{i_2,2}, \dots, \sum_{i_n} \mathbf{m}_{i_n} N_{i_n,n} \right), & (3.67) \\
 \det(M \circ N) &\stackrel{(3.56)}{=} \sum_{i_1, \dots, i_n} \Delta(\mathbf{m}_{i_1}, \dots, \mathbf{m}_{i_n}) N_{i_1,1} \dots N_{i_n,n} \\
 &\stackrel{(3.55)}{=} \sum_{i_1, \dots, i_n} \varepsilon_{i_1, \dots, i_n} \Delta(\mathbf{m}_1, \dots, \mathbf{m}_n) N_{i_1,1} \dots N_{i_n,n} \\
 &= \underbrace{\left(\sum_{i_1, \dots, i_n} \varepsilon_{i_1, \dots, i_n} N_{i_1,1} \dots N_{i_n,n} \right)}_{\det N} \underbrace{\Delta(\mathbf{m}_1, \dots, \mathbf{m}_n)}_{\det M}. & (3.68)
 \end{aligned}$$

Satz 3.13. *Die Determinante ist invariant unter einer orthogonalen Basistransformation.*

Beweis.

$$\begin{aligned}
 \det M' &= \det(O^T M O) = \det O^T \det M \det O \\
 &= \det O^T \det O \det M = \underbrace{\det(O^T O)}_{=1} \det M. & (3.69)
 \end{aligned}$$

Zur praktischen Bestimmung von Determinanten gibt es zwei Methoden:

Methode 1: Laplacescher Entwicklungssatz.

$$\det M = \sum_{i=1}^n M_{ij} (-1)^{i+j} \det \hat{M}_{ij}, \quad (3.70)$$

dabei sei j beliebig aber fix, $(-1)^{i+j} \det \hat{M}_{ij}$ der *Kofaktor* und $\det \hat{M}_{ij}$ wie M_{ij} aber ohne i -te Zeile und j -te Spalte. Man kann auch i fixieren und über j summieren.

Beweis. Aus der Definition folgt

$$\begin{aligned}
 \det M &= \sum_{i=1}^n \varepsilon_{i_1, \dots, i_n} M_{1,i_1} M_{2,i_2} \dots M_{n,i_n} \\
 &= \sum_{i_j=1} M_{i_j,j} \underbrace{\sum_{i_1, \dots, \cancel{i_j}, \dots, i_n} \varepsilon_{i_1, \dots, i_n} M_{i_1,1} \dots \cancel{M}_{i_j,j} \dots M_{i_n,n}}_{\pm \det \hat{M}_{ij}}, & (3.71)
 \end{aligned}$$

wobei der Beitrag mit Index $i_j \rightarrow \cancel{i_j}$ bzw. $M_{i_j,j} \rightarrow \cancel{M}_{i_j,j}$ weggelassen wurde.

Wir fixieren die Spalte $j = 1$ und summieren über i

Beispiel 42.

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} &= 1 \det \begin{pmatrix} 5 & 6 \\ 8 & 0 \end{pmatrix} - 4 \det \begin{pmatrix} 2 & 3 \\ 8 & 0 \end{pmatrix} + 7 \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} \\ &= 1(5 \cdot 0 - 8 \cdot 6) - 4(2 \cdot 0 - 3 \cdot 8) + 7(2 \cdot 6 - 5 \cdot 3) = -48 + 96 - 21 = 27. \end{aligned} \quad (3.72)$$

Methode 2: Gauß-Algorithmus. Zu jeder Spalte (bzw. Zeile) kann man eine beliebige Linearkombination andere Spalten (bzw. Zeilen) hinzuaddieren, um die Determinante zu vereinfachen.

Beweis. Aufgrund der Linearität der Determinante (3.56) können wir schreiben, dass

$$\begin{aligned} \Delta \left(\mathbf{m}_1, \dots, \mathbf{m}_i + \sum_{k \neq i} c_k \mathbf{m}_k, \dots, \mathbf{m}_n \right) &= \Delta \left(\mathbf{m}_1, \dots, \mathbf{m}_i, \dots, \mathbf{m}_n \right) \\ &+ \sum_{k \neq i} c_k \Delta \left(\mathbf{m}_1, \dots, \mathbf{m}_k, \dots, \mathbf{m}_n \right), \end{aligned} \quad (3.73)$$

wobei der letzte Term, die Summe über k , aufgrund von Antisymmetrie verschwindet, da \mathbf{m}_k jeweils gleich einem anderen der n Vektoren ist.

Beispiel 43.

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix} = 1(-3)(-9) = 27. \quad (3.74)$$

Die Determinante ist dann in oberer Dreiecksform und kann mittels des Produkts über die Diagonalelemente, $\det M = \prod_{i=1}^n M_{ii}$, berechnet werden. Dies ergibt sich aus iterativer Anwendung des Laplaceschen Entwicklungssatzes (3.70). Falls eine Spalte oder Zeile verschwindet, ist $\det = 0$. Bei grossen Matrizen (e.g. $n \geq 4$) ist Methode 2 effektiver.

References

- [1] C. B. Lang and N. Pucker, *Mathematische Methoden in der Physik*. Springer Berlin Heidelberg, Berlin, Heidelberg, 3 ed., 2016.
- [2] H. v. Hees, *Mathematische Ergänzungen zur Theoretischen Physik 1*. 2018.
- [3] R. Adams and C. Essex, *Calculus: A Complete Course*. Pearson, 10 ed., 2022.
- [4] *Analysis Eins*. Serlo Education, 1 ed., 2017.
- [5] J. Weidmann, *Linear Operators in Hilbert Space*. Springer Verlag, New York, Berlin, Heidelberg, 1980.
- [6] M. Laine, *Mathematische Methoden der Physik I*. 2013.
- [7] L. Collatz, *Differentialgleichungen*. Teubner, 1990.
- [8] R. Bronson, *Differential Equations*. Schaum's Easy Outlines. McGraw-Hill, New York, Chicago, San Francisco, 2003.