Analysis II

Wolf P. Barth

Erlangen, SS 07

Version vom 27. April 2007
Mathematisches Institut der Universität, Bismarckstr. 1 1/2, D-91054 Erlangen

Inhaltsverzeichnis

0 Einführung

Stoff des ersten Semesters Analysis war die Differential- und Integralrechnung in einer Veränderlichen. Logischerweise sollte der Stoff des zweiten Semesters die Differential- und Integralrechnung in mehreren Veränderlichen sein. Genau wie im ersten Semester gehören dazu die drei großen Problemkreise

- Konvergenz
- Differentiation
- Integration

Die grundlegenden mathematischen Techniken bleiben die gleichen wie im ersten Semester. Weil es aber im $\mathbb{R}^n$ sehr viel mehr verschiedenartige Mengen gibt, als im $\mathbb{R}^1$, verlagern sich die Problemstellungen und damit die Gewichte etwas. Oft ist man happy, wenn man die Probleme auf die Theorie einer Veränderlichen zurückführen kann, wie etwa bei der partiellen Differentiation. Oft treten aber auch völlig neuartige Probleme auf. Ganz allgemein ist die Fülle des Stoffes viel umfangreicher, als in der Theorie einer Veränderlichen. Der Idealfall wäre, dass man die drei genannten Problemkreise im zweiten Semester im Wesentlichen abhandelt. Insbesondere bei der Integration wird das kritisch.

So ist man seit einigen Jahrzehnten dazu übergegangen, die Integration ganz ins dritte Semester zu verschieben, und dafür im zweiten Semester Grundlagen für die Theorie der Differentialgleichungen zu behandeln. Beispielhaft dafür ist das Buch

O. Forster, Analysis II, Rohwolt Vieweg 1977

Genau wie Forster I ist es ein Klassiker, der seither in vielen Auflagen erschienen ist. Ich persönlich halte den Forster I für das beste Mathematik-Lehrbuch in deutscher Sprache. Mit dem Forster II bin ich bei weitem nicht so vorbehaltlos einverstanden. Ich halte mich z.T. an meine Analysis-Skripten vergangener Jahre. Das Problem dabei ist, dass ich schon lange keine Analysis-Vorlesung mehr gehalten habe, und dass ich keine geTEX-te Version habe. Ich muss also diesen Begleittext zur Vorlesung ziemlich neu tippen. Wie ich das zeitlich hinkriege, weiß ich im Moment noch nicht. Bei meinen früheren Analysis-Vorlesungen habe ich mich stark an den Büchern

F. Erwe, Differential- und Integralrechnung I, II, BI 1962

orientiert. Jetzt habe ich mich bei Stoff-Auswahl, Darstellung und Beweisen sehr an das Skriptum

A. Knauf, Analysis II

gehalten. Das gilt vor allem für viele Details, weniger für die Gliederung des Stoffes. In dem genannten Skriptum finden Sie auch viele illustrative Graphiken, für deren Erstellung ich momentan keine Zeit habe. Auch bezüglich weiterer Literatur möchte ich auf das Skriptum des Kollegen Knauf verweisen.

1 Der Zahlenraum $\mathbb{R}^n$

Der Zahlenraum $\mathbb{R}^n$ besteht aus allen n-tupeln reeller Zahlen. So ein n-tupel nennt man auch Vektor. Wir wollen Vektoren immer als Spaltenvektoren

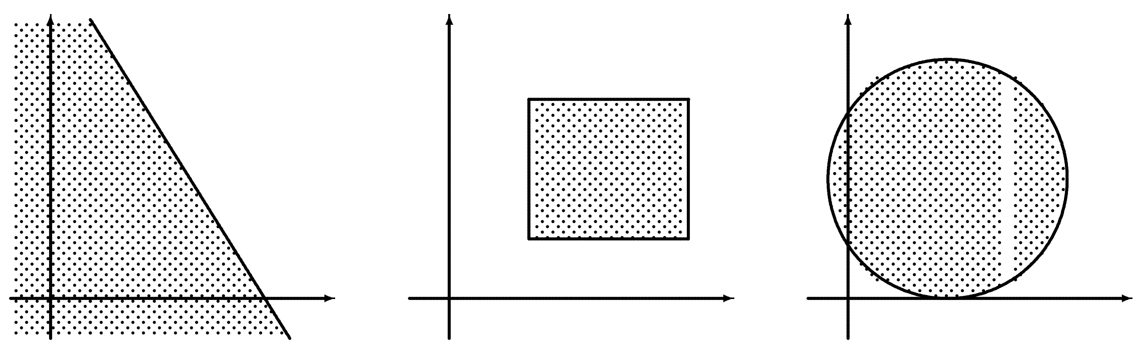
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

schreiben, d.h., immer, wenn nichts dagegenspricht.

Die algebraischen Eigenschaften dieses Zahlenraums untersucht man in der Linearen Algebra. Wir beschäftigen uns mit seinen analytischen Eigenschaften. Die Analysis im $\mathbb{R}^n$ ist vor allem deswegen schwieriger, als die eindimensionale Analysis aus den ersten beiden Semestern, weil es im $\mathbb{R}^n$ sehr viel mehr verschiedenartige Teilmengen gibt, als in der reellen Geraden. Die einfachsten davon, die immer wieder vorkommen werden, sind - außer den linearen und affinen Unterräumen - z.B.

Halbräume	$\{\mathbf{x} \in \mathbb{R}^n : a_1x_1 + a_2x_2 + ... + a_nx_n \leq c\}$	$a_1, a_2, ..., a_n, c \in \mathbb{R}$
Quader	$\{\mathbf{x} \in \mathbb{R}^n : a_1 \leq x_1 \leq b_1, ..., a_n \leq x_n \leq b_n\}$	$a_1 < b_1, ..., a_n < b_n \in \mathbb{R}$
Kugeln	$\{\mathbf{x} \in \mathbb{R}^n : (x_1 - a_1)^2 + ... + (x_n - a_n)^2 \leq r^2\}$	$a_1, a_2, ..., a_n, r \in \mathbb{R}$

Ihre zweidimensionalen Versionen sehen so aus:



Halbebene Rechteck Kreisscheibe

1.1 Funktionen und Abbildungen

Wenn man mehr als eine Dimension zur Verfügung hat, kann man Abbildungen betrachten, deren Definitionsmenge etwa im $\mathbb{R}^m$ liegt, und deren Bildmenge zum $\mathbb{R}^n$ gehört:

$$F : U \to V, \quad U \subset \mathbb{R}^m, V \subset \mathbb{R}^n.$$

Solche Abbildungen F kann man sich natürlich jetzt nur viel schwerer vorstellen, als die Funktionen der eindimensionalen Analysis. Wir wollen uns ihnen deswegen schrittweise nähern.

Der Fall $m = 1$: Kurven. Eine Abbildung eines Intervalls $[a, b] \subset \mathbb{R}$ in den $\mathbb{R}^n$ heißt *Kurve*. Man schreibt so eine Kurve wie folgt:

$$\phi : [a, b] \ni t \mapsto \phi(t) = \mathbf{x}(t) \in \mathbb{R}^n$$

Am besten stellt man sich t als einen Parameter (etwa die Zeit) vor, von dem der Bildpunkt abhängt. Dieser Bildvektor

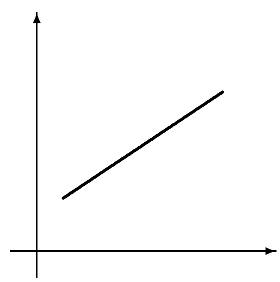
$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

ist also ein n-tupel von ganz normalen Funktionen $x_\nu(t)$, $\nu = 1, ..., n$.

Für $n = 2$ oder $n = 3$ kann man sich solche Kurven noch recht gut vorstellen. Zeichnen kann man sie eigentlich vernünftig nur für $n = 2$ (ebene Kurven). Hier einige Beispiele:

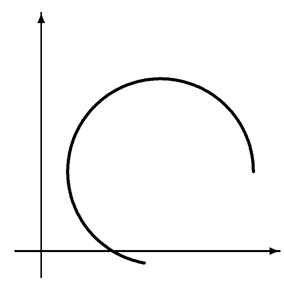
Geradenstück

$x_1 = a \cdot t + b$
$x_2 = c \cdot t + d$



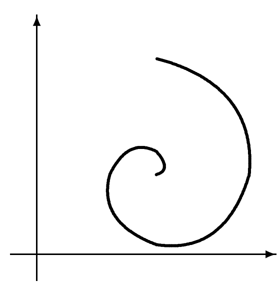
Kreisbogenstück

$x_1 = a + r \cdot cos(t)$
$x_2 = b + r \cdot sin(t)$



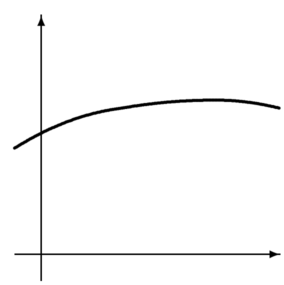
Spiralenstück

$x_1 = a + c \cdot t \cdot cos(t)$
$x_2 = b + c \cdot t \cdot sin(t)$



Funktionsgraph

$x_1 = t$
$x_2 = f(t)$



Hier ordnen sich also auch die Graphen aller in der eindimensionalen Analysis betrachteten Funktionen ein.

Raumkurven möchte ich hier nicht mehr zeichnen, aber eine besonders schöne ist die Helix (Wendeltreppe)

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} r \cdot cos(t) \\ r \cdot sin(t) \\ c \cdot t \end{pmatrix}.$$

Praktisch kommen Raumkurven vor als Bahnkurven von Teilchen, solange man sich diese Teilchen punktförmig denkt: Asteroiden im Weltraum, Kanonenkugeln in der Erdatmosphäre, Elementarteilchen im Teilchenbeschleuniger, Schmetterlinge über einer Wiese, ...

Der Fall $n = 1$: Funktionen. Eine Abbildung

$$f : \mathbb{R}^n \ni \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \mapsto \quad f(x_1, ..., x_n) \in \mathbb{R}$$

bezeichnet man üblicherweise als *Funktion*, auch wenn sie von $m > 1$ Variablen x_μ abhängt. Für $m = 1$ erhalten wir nichts neues.

Für $m = 2$ bekommen wir Funktionen $f(x_1, x_2)$ von zwei Variablen. Deren Graphen

$$y = f(x_1, x_2)$$

kann man sich immer noch vorstellen, als eine Fläche, die über der (x_1, x_2)-Ebene im Raum liegt. Zeichnen möchte ich die jetzt nicht mehr, das ist mir zu aufwendig. Aber hier einige Beispiele:

affine Ebene	$f(x_1, x_2) = a \cdot x_1 + b \cdot x_2 + c$
Halb-Sphäre	$f(x_1, x_2) = \sqrt{r^2 - (x_1^2 + x_2^2)}$
Paraboloid	$f(x_1, x_2) = x_1^2 + x_2^2$
Sattelfläche	$f(x_1, x_2) = x_1^2 - x_2^2$

Hierher gehören auch Verknüpfungen, wie sie aus der Algebra bekannt sind

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mapsto \quad x_1 + x_2, \quad x_1 - x_2, \quad x_1 \cdot x_2, \quad x_1/x_2,$$

oder aus der Analysis

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \mapsto \quad x_1^{x_2}, \quad ln^{x_1}(x_2).$$

Der Fall $m = n$: Dieser Fall kommt sehr häufig vor, unter dem Namen *Transformation*. Eine Transformation ist eine bijektive Abbildung

$$F : U \to V, \quad U, V \subset \mathbb{R}^n.$$

Hierher gehören die Vektorraum-Isomorphismen

$$\mathbf{x} \mapsto A \cdot \mathbf{x}, \quad A \in GL(n, \mathbb{R}),$$

aus der linearen Algebra, oder die Affinitäten

$$\mathbf{x} \mapsto \mathbf{a} + A \cdot \mathbf{x}.$$

Außer diesen Transformationen ist für uns die wichtigste die Transformation von euklidischen Koordinaten in Polarkoordinaten r, φ

$$\begin{pmatrix} r \\ \varphi \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} r \cdot cos(\varphi) \\ r \cdot sin(\varphi) \end{pmatrix}.$$

Diese Transformation ist meistens günstig bei rotations-symmetrischen Situationen in der Ebene. Es gibt auch eine ähnliche Transformation für rotations-symmetrische Situationen im Raum, die Transformation in Kugelkoordinaten r, φ, θ

$$\begin{pmatrix} r \\ \varphi \\ \theta \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} r \cdot cos(\varphi) \cdot cos(\theta) \\ r \cdot sin(\varphi) \cdot cos(\theta) \\ r \cdot sin(\theta) \end{pmatrix}.$$

Natürlich gibt es auch Abbildungen $U \to V$, mit $U, V \in \mathbb{R}^n$, die nicht bijektiv sind. Ziemlich häufig kommt zum Beispiel die Projektion des Raums in die (x_1, x_2)-Ebene vor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}.$$

Abbildungen $F : U \to V$ mit $U, V \subset \mathbb{R}^2$ oder $\mathbb{R}^3$ kommen auch als *Vektorfelder* vor: In jedem Punkt $\mathbf{x} \in U$ stellt man sich den Bildpunkt (= Bildvektor) $F(\mathbf{x}) \in V$ als Vektor angeheftet vor. Solche Vektorfelder braucht man zur Beschreibung des Feldes von Geschwindigkeitsvektoren einer Flüssigkeitsströmung, oder für elektrische oder magnetische Felder.

Der Allgemeinfall: Sind m und $n > 0$, so kann man sich eine Abbildung $U \to V$, $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ meist nicht mehr vorstellen. Vorkommen tun sie trotzdem. So deutet man etwa den Fall $m = 2, n = 3$ als Parametrisierung einer Fläche (in Analogie zu $m = 1, n = 3$, der Parametrisierung einer Kurve,) im Raum. Für uns wichtig sind aber nur die oben aufgeführten Fälle.

Noch etwas zur Schreibweise: Eine Abbildung $F : U \to V$ mit $U \subset \mathbb{R}^m$ und $V \subset \mathbb{R}^n$ schreibe ich immer als

$$F : U \to V, \quad F : \mathbf{x} \mapsto F(\mathbf{x}) = \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix}.$$

Dabei ist $F_\nu(\mathbf{x})$ die ν-te Komponente des Bildvektors $F(\mathbf{x}) \in V \subset \mathbb{R}^n$. Diese ν-te Komponente ist eine Funktion $F_\nu : U \to \mathbb{R}$, also eine Funktion $F_\nu(\mathbf{x}) = F_\nu(x_1, ..., x_m)$ von m Variablen. Die Abbildung F ist genau dasselbe wie die Kollektion ihrer n Komponentenfunktionen F_ν,

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}.$$

Beispiel 1.1 *Die Komponentenfunktionen der Polarkoordinatentransformation sind*

$$F_1(r, \varphi) = r \cdot cos(\varphi), \quad F_2(r, \varphi) = r \cdot sin(\varphi).$$

Auf der Halbebene $x_1 > 0$ z.B. ist sie umkehrbar durch

$$r = \sqrt{x_1^2 + x_2^2}, \quad \varphi = arctg\left(\frac{x_2}{x_1}\right).$$

Die Komponentenfunktionen dieser Umkehrabbildung G sind

$$G_1(x_1, x_2) = \sqrt{x_1^2 + x_2^2}, \quad G_2(x_1, x_2) = arctg\left(\frac{x_2}{x_1}\right).$$

Aufgabe 1.1 *Die Abbildung $F : \mathbb{R}^2 \to \mathbb{R}^2$ sei definiert durch*

$$F(t, \varphi) := \begin{pmatrix} cosh(t) \cdot cos(\varphi) \\ sinh(t) \cdot sin(\varphi) \end{pmatrix}.$$

Zeigen Sie, dass F die Geraden $t = t_0 \neq 0$ auf Ellipsen und die Geraden $\varphi = \varphi_0 \neq n\frac{\pi}{2}$, $n \in \mathbb{Z}$, auf Hyperbeln abbildet.

Aufgabe 1.2 *Zeigen Sie, dass durch*

$$u := x^2 - y^2, \quad v := 2xy$$

der Quadrant $\{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$ bijektiv auf die Halbebene $(u, v) \in \mathbb{R}^2 : v > 0\}$ abgebildet wird, indem Sie die Umkehrabbildung angeben.

Aufgabe 1.3 *Bestimmen Sie auf dem Halbraum $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 > 0\}$ die Komponentenfunktionen*

$$r = r(x_1, x_2, x_3), \quad \varphi = \varphi(x_1, x_2, x_3), \quad \theta = \theta(x_1, x_2, x_3)$$

der Umkehrabbildung zur Transformation in Kugelkoordinaten r, φ, θ.

1.2 Abstand und Topologie

Ausgangspunkt zur Abstandsberechnung im $\mathbb{R}^n$ ist das euklidische Skalarprodukt

$$(\mathbf{x}.\mathbf{y}) = \sum_{\nu=1}^{n} x_\nu y_\nu$$

aus der Linearen Algebra. Seine *drei wichtigen Eigenschaften* sind wohlbekannt. Stellen wir sie hier noch einmal zusammen:

- *Bilinearität:* Das Skalarprodukt $(\mathbf{x}.\mathbf{y})$ ist linear in $\mathbf{x}$ und $\mathbf{y}$. (Das brauchen wir nicht noch explizit in Formeln hinzuschreiben.)
- *Symmetrie:* $(\mathbf{x}.\mathbf{y}) = (\mathbf{y}.\mathbf{x})$.
- *Positiv-Definitheit:* Es ist stets $(\mathbf{x}.\mathbf{x}) \geq 0$ und $(\mathbf{x}.\mathbf{x}) = 0$ nur wenn $\mathbf{x} = \mathbf{0}$.

Die Zahl

$$\parallel \mathbf{x} \parallel := \sqrt{(\mathbf{x}.\mathbf{x})} = \sqrt{\sum_{\nu=1}^{n} (x_\nu)^2}$$

heißt *Länge* oder *(euklidische) Norm* des Vektors $\mathbf{x} \in \mathbb{R}^n$. Auch diese Norm-Funktion hat *drei wichtige Eigenschaften*:

N1: Aus der Positiv-Definitheit folgt $\| \mathbf{x} \| \geq 0$ für alle $\mathbf{x} \in \mathbb{R}^n$, und $\| \mathbf{x} \| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$.

N2: Aus der Bilinearität folgt

$$\| c \cdot \mathbf{x} \| = |c| \cdot \| \mathbf{x} \|$$

für alle $c \in \mathbb{R}$ und $\mathbf{x} \in \mathbb{R}^n$.

- Es gilt die Cauchy-Schwarz-Ungleichung

$$|(\mathbf{x}.\mathbf{y})| \leq \| \mathbf{x} \| \cdot \| \mathbf{y} \| \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

N3: Aus der Cauchy-Schwarz-Ungleichung folgt die Dreiecks-Ungleichung

$$\| \mathbf{x} + \mathbf{y} \| \leq \| \mathbf{x} \| + \| \mathbf{y} \| \qquad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Beweis der Cauchy-Schwarz-Ungleichung: Wenn $\mathbf{y} = \mathbf{0}$ ist, steht auf beiden Seiten der Ungleichung 0, dann gilt sie also trivialerweise, wie man sagt. Wenn $\mathbf{y} \neq \mathbf{0}$ ist, dann projizieren wir den Vektor $\mathbf{x}$ orthogonal in die von $\mathbf{y}$ aufgespannte Grade $\mathbb{R} \cdot \mathbf{y}$ und erhalten

$$\mathbf{x}' = \frac{(\mathbf{x}.\mathbf{y})}{(\mathbf{y}.\mathbf{y})}\mathbf{y}.$$

(Dies ist tatsächlich die Orthogonalprojektion des Vektors $\mathbf{x}$ in die Gerade $\mathbb{R} \cdot \mathbf{y}$, denn man sieht sofort, dass der Differenzvektor $\mathbf{x}' - \mathbf{x}$ auf $\mathbf{y}$ senkrecht steht:

$$(\mathbf{x}' - \mathbf{x}.\mathbf{y}) = \frac{(\mathbf{x}.\mathbf{y})}{(\mathbf{y}.\mathbf{y})}(\mathbf{y}.\mathbf{y}) - (\mathbf{x}.\mathbf{y}) = 0.$$

Aber das brauchen wir für den Beweis gar nicht.) Wir brauchen nur die Positiv-Definitheit, ausgewertet für den Vektor $\mathbf{x}' - \mathbf{x}$:

$$\begin{aligned}
\| \mathbf{x}' - \mathbf{x} \|^2 &= \| \mathbf{x}' \|^2 - 2(\mathbf{x}'.\mathbf{x}) + \| \mathbf{x} \|^2 \\
&= \left(\frac{(\mathbf{x}.\mathbf{y})}{(\mathbf{y}.\mathbf{y})} \right)^2 \| \mathbf{y} \|^2 - 2 \cdot \frac{(\mathbf{x}.\mathbf{y})}{(\mathbf{y}.\mathbf{y})} \cdot (\mathbf{y}.\mathbf{x}) + \| \mathbf{x} \|^2 \\
&= -\frac{(\mathbf{x}.\mathbf{y})^2}{\| \mathbf{y} \|^2} + \| \mathbf{x} \|^2 \\
&\geq 0, \\
\frac{(\mathbf{x}.\mathbf{y})^2}{\| \mathbf{y} \|^2} &\leq \| \mathbf{x} \|^2, \\
(\mathbf{x}.\mathbf{y})^2 &\leq \| \mathbf{x} \|^2 \cdot \| \mathbf{y} \|^2 .
\end{aligned}$$

Die Cauchy-Schwarz-Ungleichung ergibt sich, wenn man aus dieser letzten Ungleichung die Wurzel zieht.

Beweis der Dreiecksungleichung: Aus der Cauchy-Schwarz-Ungleichung folgt

$$\begin{aligned}
\| \mathbf{x} + \mathbf{y} \|^2 &= (\mathbf{x} + \mathbf{y}.\mathbf{x} + \mathbf{y}) \\
&= \| \mathbf{x} \|^2 + 2(\mathbf{x}.\mathbf{y}) + \| \mathbf{y} \|^2 \\
&\leq \| \mathbf{x} \|^2 + 2 \cdot \| \mathbf{x} \| \cdot \| \mathbf{y} \| + \| \mathbf{y} \|^2 \\
&= (\| \mathbf{x} \| + \| \mathbf{y} \|)^2 ,
\end{aligned}$$

und das ist die quadrierte Form der Dreiecks-Ungleichung. $\square$

Beispiel 1.2 (Matrix-Norm) *Der $\mathbb{R}$-Vektorraum $M(m \times n, \mathbb{R})$ der reellen $m \times n$-Matrizen ist einfach nur ein Exemplar des Vektorraums $\mathbb{R}^{(m \cdot m)}$. Als solcher hat er eine Norm. Für eine Matrix*

$$A = (a_{\mu,\nu})_{\substack{\mu=1,\ldots,m, \\ \nu=1,\ldots,n}} \qquad \textit{ist} \quad \| A \| = \sqrt{\sum_{\mu,\nu=1}^{m,n} a_{\mu,\nu}^2}.$$

Diese Matrix-Norm hat folgende Eigenschaften, die für Abschätzungen sehr nützlich sind:

1) Für alle $A \in M(m \times n, \mathbb{R})$ und $\mathbf{x} \in \mathbb{R}^n$ ist

$$\| A \cdot \mathbf{x} \| \leq \| A \| \cdot \| \mathbf{x} \| .$$

2) Für alle $A \in M(m \times n, \mathbb{R})$ und $B \in M(n \times p, \mathbb{R})$ ist

$$\| A \cdot B \| \leq \| A \| \cdot \| B \| .$$

Beweis von 1): Die Zeilenvektoren der Matrix A seien $A_1, ..., A_m$. Dann ist also

$$A \cdot \mathbf{x} = \begin{pmatrix} A_1 \cdot \mathbf{x} \\ \vdots \\ A_m \cdot \mathbf{x} \end{pmatrix}.$$

Mit der Cauchy-Schwarz-Ungleichung

$$\| A_\mu \cdot \mathbf{x} \| \leq \| A_\mu \| \cdot \| \mathbf{x} \|$$

folgt

$$\| A \cdot \mathbf{x} \|^2 = \sum_{\mu=1}^{m} \| A_\mu \cdot \mathbf{x} \|^2 \leq \sum_{\mu=1}^{m} \| A_\mu \|^2 \cdot \| \mathbf{x} \|^2 = \| A \| \| \mathbf{x} \| .$$

Beweis von 2): Es sein $B_1, ..., B_p$ die Spaltenvektoren von B. Dann ist also

$$A \cdot B = (A \cdot B_1, ..., A \cdot B_p).$$

Mit Eigenschaft 1) folgt

$$\| A \cdot B \|^2 = \| A \cdot B_1 \|^2 + ... + \| A \cdot B_p \|^2 \leq \| A \|^2 \cdot (\| B_1 \|^2 + ... + \| B_p \|^2) = \| A \|^2 \cdot \| B \|^2 .$$

Definition 1.1 (Norm) *Es sei V ein $\mathbb{R}$-Vektorraum. Eine* Norm *auf V ist eine Abbildung*

$$V \to \mathbb{R}, \quad \mathbf{x} \mapsto \| \mathbf{x} \|,$$

mit den Eigenschaften N1, N2, N3.

Mit der oben angegebenen Norm auf $\mathbb{R}^n$ definiert man den euklidischen Abstand (die Distanz) zweier Vektoren $\mathbf{x}$ und $\mathbf{y} \in \mathbb{R}^n$ als

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

Die *drei wichtigen Eigenschaften* des Abstands sind:

M1 Es ist stets $d(\mathbf{x},\mathbf{y}) \geq 0$ und $= 0$ nur, wenn $\mathbf{x} = \mathbf{y}$. (Dies sieht man sofort mit N1.)

M2 Die Abstandsfunktion ist symmetrisch: $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$. (Dies folgt aus N2, wenn man dort $c = -1$ setzt.)

M3 Es gilt die *Dreiecksungleichung*

$$d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \text{ für alle } \mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^n.$$

Beweis der Dreiecksungleichung: Mit N3 erhalten wir

$$d(\mathbf{x},\mathbf{z}) = \| \mathbf{x}-\mathbf{z} \| = \| (\mathbf{x}-\mathbf{y}) + (\mathbf{y}-\mathbf{z}) \| \leq \| \mathbf{x}-\mathbf{y} \| + \| \mathbf{y}-\mathbf{z} \| = d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}). \qquad \square$$

Diese Abstandsfunktion $d(\mathbf{x},\mathbf{y})$ nennt man auch *Metrik*.

Definition 1.2 (Metrik) *Allgemein heißt eine Funktion $d(\mathbf{x},\mathbf{y})$ von zwei Elementen aus einer beliebigen Menge M eine Metrik auf M, wenn sie die Eigenschaften M1, M2, M3 besitzt. Die Menge M zusammen mit einer Metrik d heißt dann ein* metrischer Raum.

Mit dieser Abstandsfunktion $d(\mathbf{x},\mathbf{y})$ kann man Kugeln definieren:

Definition 1.3 *Die Menge*

$$\begin{array}{rcll} \bar{K} & := & \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}-\mathbf{a}) \leq r\} & \textit{heißt abgeschlossene,} \\ K & := & \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}-\mathbf{a}) < r\} & \textit{heißt offene} \end{array}$$

Kugel mit Mittelpunkt $\mathbf{a}$ *und Radius r.*

Den Unterschied zwischen einer abgeschlossenen und einer offenen Kugel sieht man an deren Rand: Der *Rand* der oben definierten Kugeln ist die Menge der Punkte, welche vom Mittelpunkt genau den Abstand r haben:

$$\partial(\bar{K}) = \partial(K) = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x},\mathbf{a}) = r\}.$$

Bei der abgeschlossenen Kugel $\bar{K}$ gehört der Rand dazu, bei der offenen Kugel nicht. (Das ist genau derselbe Unterschied, wie bei den Intervallen auf der rellen Achse $\mathbb{R}$: Zum abgeschlossenen Intervall gehören die Randpunkte dazu, zum offenen Intervall nicht.) Der Radius r einer Kugel soll bei uns immer > 0 sein. Für $r = 0$ bestünde die abgeschlossene Kugel vom Radius $r = 0$ nur aus ihrem Mittelpunkt, die offene Kugel vom Radius $r = 0$ wäre leer.

Mit offenen Kugeln kann man offene Mengen definieren:

Definition 1.4 *Eine Teilmenge $U \subset \mathbb{R}^n$ heißt* offen, *wenn zu jedem Punkt* $\mathbf{u} \in U$ *eine Kugel mit Mittelpunkt* $\mathbf{u}$ *und einem Radius $r > 0$ existiert, die noch ganz zu U gehört.*

Beispiel 1.3 *Sei $c \in \mathbb{R}$ beliebig. Die Halbebene*

$$H := \{\mathbf{x} \in \mathbb{R}^2 : x_1 < c\}$$

ist offen.

Beweis. *Sei* $\mathbf{u} \in H$*, also* $u_1 < c$*. Dann gibt es einen Radius* $r > 0$ *mit* $r < c - u_1$*, etwa* $r := (c - u_1)/2$*. Wir zeigen: Die Kugel* K *um* $\mathbf{u}$ *mit diesem Radius liegt ganz in* H*.*

In der Tat, sei $\mathbf{x} \in K$*, also* $d(\mathbf{x}, \mathbf{u}) < r$*. Dann ist*

$$x_1 - u_1 \leq |x_1 - u_1| = \sqrt{(x_1 - u_1)^2} \leq \sqrt{(x_1 - u_1)^2 + (x_2 - u_2)^2} = d(\mathbf{x}, \mathbf{u}) < r.$$

Daraus folgt

$$x_1 < u_1 + r < c$$

und $\mathbf{x} \in H$*.* □

Beispiel 1.4 *Dieselbe Halbebene, definiert mit dem* $\leq$*-Zeichen an Stelle des* $<$*-Zeichens*

$$\bar{H} := \{\mathbf{x} \in \mathbb{R}^2 : x_1 \leq c\}$$

ist nicht *offen.*

Beweis. *Zu* $\bar{H}$ *gehören auch alle Punkte* $\mathbf{u} \in \mathbb{R}^2$ *mit* $u_1 = c$*. Wir halten einen solchen Punkt* $\mathbf{u}$ *fest und zeigen: Es gibt keinen Radius* $r > 0$ *derart, dass die Kugel um* $\mathbf{u}$ *mit diesem Radius noch ganz in* $\bar{H}$ *liegt.*

In der Tat, sei $r > 0$*. Für den Punkt*

$$\mathbf{x} := \begin{pmatrix} u_1 + r/2 \\ u_2 \end{pmatrix}$$

ist der Abstand zu $\mathbf{u}$

$$d(\mathbf{x}, \mathbf{u}) = \sqrt{(x_1 - u_1)^2 + (x_2 - u_2)^2} = \frac{r}{2} < r,$$

Also gehört $\mathbf{x}$ *zur Kugel* K *um* $\mathbf{u}$ *mit Radius* r*. Wegen*

$$x_1 = u_1 + \frac{r}{2} = c + \frac{r}{2} > c$$

ist $\mathbf{x} \notin \bar{H}$*. Die Kugel* K *liegt also nicht ganz in* $\bar{H}$*.* □

Beispiel 1.5 *Offene Kugeln (im Sinn der Definition offener Kugeln) sind offen (im Sinn der Definition offener Mengen).*

Beweis. *Sei*

$$K := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) < R\}$$

eine offene Kugel mit Mittelpunkt $\mathbf{a}$ *und Radius* $R > 0$*. Sei* $\mathbf{u} \in K$ *ein Punkt. Wir müssen zeigen, es gibt einen Radius* $r > 0$ *derart, dass die Kugel* $K_{\mathbf{u}}$ *um* $\mathbf{u}$ *mit diesem Radius* r *ganz in der Kugel* K *liegt.*

In der Tat, weil $\mathbf{u}$ *zu* K *gehört ist* $d(\mathbf{u}, \mathbf{a}) < R$ *und*

$$r := \frac{R - d(\mathbf{u}, \mathbf{a})}{2} > 0.$$

Sei $K_{\mathbf{u}}$ *die Kugel um* $\mathbf{u}$ *mit diesem Radius* r*. Nach der Dreiecksungleichung gilt für jeden Punkt* $\mathbf{x}$ *dieser Kugel* $K_{\mathbf{u}}$

$$d(\mathbf{x}, \mathbf{a}) \leq d(\mathbf{x}, \mathbf{u}) + d(\mathbf{u}, \mathbf{a}) < \frac{R - d(\mathbf{u}, \mathbf{a})}{2} + d(\mathbf{u}, \mathbf{a}) < (R - d(\mathbf{u} - \mathbf{a})) + d(\mathbf{u} - \mathbf{a}) = R.$$

Also gehört jeder Punkt $\mathbf{x} \in K_{\mathbf{u}}$ *zu* K*, es gilt* $K_{\mathbf{u}} \subset K$*.* □

Beispiel 1.6 *Die leere Menge* $\emptyset \subset \mathbb{R}^n$ *ist offen: Mit jedem Punkt* $\mathbf{u} \in \emptyset$ *gehört eine ganze Kugel um* $\mathbf{u}$ *zu* $\emptyset$. *Das ist richtig, denn es gibt keinen einzigen Punkt* $\mathbf{u} \in \emptyset$, *für den wir das beweisen müssen.*

Der Gesamtraum $\mathbb{R}^n$ *als Teilmenge von sich selbst ist offen. Mit jedem Punkt* $\mathbf{u} \in \mathbb{R}^n$ *gehört ja auch jede Kugel um* $\mathbf{u}$ *zu* $\mathbb{R}^n$.

Beispiel 1.7 *Sind* $U, V \subset \mathbb{R}^n$ *offen, so ist auch ihre Vereinigung* $U \cup V \subset \mathbb{R}^n$ *offen.*

Beweis. Sei $\mathbf{u} \in U \cup V$, *etwa* $\mathbf{u} \in U$. *Weil* U *offen ist, gibt es ein Kugel* K *mit Zentrum* $\mathbf{u}$, *die ganz in* U *liegt. Diese Kugel* K *gehört dann auch zu* $U \cup V$. □

Genauso sieht man, dass die Vereinigung beliebig vieler (auch unendlich vieler) offener Mengen wieder offen ist.

Beispiel 1.8 *Sind* $U, V \subset \mathbb{R}^n$ *offen, so ist auch ihr Durchschnitt* $U \cap V \subset \mathbb{R}^n$ *wieder offen.*

Beweis. Sei $\mathbf{u} \in U \cap V$. *Weil* U *offen ist, gibt es eine Kugel* K_U *mit Zentrum* $\mathbf{u}$, *die ganz in* U *liegt. Und weil* V *offen ist, gibt es auch eine Kugel* K_V *um* $\mathbf{u}$, *die ganz in* V *liegt. Sei* K *diejenige der beiden Kugeln* K_U *und* K_V, *welche den kleineren Radius besitzt. Dann ist also* $K \subset K_U \subset U$ *und* $K \subset K_V \subset V$, *also* $K \subset U \cap V$. □

Genauso sieht man, dass der Durchschnitt beliebig (aber endlich) vieler offener Mengen wieder offen ist.

Definition 1.5 *Eine Teilmenge* $A \subset \mathbb{R}^n$ *heißt* abgeschlossen, *wenn ihr Komplement* $U := \mathbb{R}^n \setminus A$ *offen ist.*

Beispiel 1.9 *Sei* $c \in \mathbb{R}$ *beliebig. Die Halbebene*

$$\bar{H} := \{\mathbf{x} \in \mathbb{R}^2 : x_1 \leq c\}$$

ist abgeschlossen.

Beweis. Ihr Komplement

$$\mathbb{R}^2 \setminus \bar{H} = \{\mathbf{x} \in \mathbb{R}^2 : x_1 > c\}$$

ist eine offene Halbebene (vgl. Beispiel 1.3 bei den offenen Mengen). □

Beispiel 1.10 *Abgeschlossene Kugeln (im Sinn der Definition abgeschlossener Kugeln) sind abgeschlossen (im Sinn der Definition abgeschlossener Mengen).*

Beweis. Sei etwa

$$\bar{K} := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) \leq R\}$$

eine abgeschlossene Kugel. Wir müssen zeigen, ihr Komplement

$$U = \mathbb{R}^n \setminus \bar{K} = \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{a}) > R\}$$

ist offen. Sei dazu $\mathbf{u} \in U$ *herausgegriffen. Dann ist also* $d(\mathbf{u}, \mathbf{a}) > R$. *Wir wählen*

$$r := \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R)$$

und nehmen als $K_{\mathbf{u}}$ *die Kugel um* $\mathbf{u}$ *mit diesem Radius* r. *Wir zeigen* $K_{\mathbf{u}} \subset U$.

Sei dazu $\mathbf{x} \in K_{\mathbf{u}}$, *also*

$$d(\mathbf{x}, \mathbf{u}) < r = \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R).$$

Wenn $\mathbf{x}$ *nicht in* U, *sondern in* $\bar{K}$ *liegen würde, dann würde* $d(\mathbf{x}, \mathbf{a}) \leq R$ *gelten, und aus der Dreiecksungleichung würde folgen*

$$d(\mathbf{u}, \mathbf{a}) \leq d(\mathbf{u}, \mathbf{x}) + d(\mathbf{x}, \mathbf{a}) < r + R = \frac{1}{2}(d(\mathbf{u}, \mathbf{a}) - R) + R < (d(\mathbf{u}, \mathbf{a}) - R) + R = d(\mathbf{u}, \mathbf{a}),$$

Widerspruch! □

Beispiel 1.11 (Abgeschlossener Quader) *Ein abgeschlossener Quader* $Q \subset \mathbb{R}^n$ *ist ein Produkt*

$$Q = I_1 \times ... \times I_n$$

von abgeschlossenen Intervallen $I_\nu = [a_\nu, b_\nu] \subset \mathbb{R}$. *Es ist also*

$$Q = \{\mathbf{x} \in \mathbb{R}^n : a_\nu \leq x_\nu \leq b_\nu \text{ für } \nu = 1, ..., n\}.$$

Diese Menge $Q \subset \mathbb{R}^n$ *ist abgeschlossen: Jede Bedingung* $a_\nu \leq x_\nu \leq b_\nu$ *ist äquivalent dazu, dass* $\mathbf{x}$ *zu keinem der beiden offenen Halbräume*

$$H_\nu^- := \{\mathbf{x} \in \mathbb{R}^n : x_\nu < a_\nu\}, \quad H_\nu^+ := \{\mathbf{x} \in \mathbb{R}^n : x_\nu > b_\nu\}$$

gehört. Deswegen ist Q *das Komplement der im* $\mathbb{R}^n$ *offenen Menge*

$$\bigcup_{\nu=1}^{n} (H_\nu^- \cup H_\nu^+).$$

Schließlich noch eine

Definition 1.6 *Sei* $M \subset \mathbb{R}^n$ *eine Teilmenge. Ein Punkt* $\mathbf{x} \in \mathbb{R}^n$ *heißt* Randpunkt *von* M, *wenn jede offene Kugel* $K \subset \mathbb{R}^n$ *mit Mittelpunkt* $\mathbf{x}$ *sowohl Punkte aus* M *als auch Punkte aus* $\mathbb{R}^n \setminus M$ *enthält. Die Menge aller Randpunkte von* M *nennt man den* Rand ∂M *von* M.

Ist beispielweise M offen, so ist kein Punkt $\mathbf{x} \in M$ ein Randpunkt, denn eine ganze Kugel K um $\mathbf{x}$ gehört ja zu M, enthält deswegen keinen Punkt aus $\mathbb{R}^n \setminus M$.

Ist beispielsweise M abgeschlossen, so gilt $\partial M \subset M$. Denn $\mathbb{R}^n \setminus M$ ist offen, und nach Definition ist $\partial M = \partial(\mathbb{R}^n \setminus M)$, und deswegen ist $\partial M \cap (\mathbb{R}^n \setminus M) = \emptyset$.

Noch ein ganz konkretes Beispiel:

Beispiel 1.12 *Sei* B *die offene Einheitskugel* $\{\mathbf{x} \in \mathbb{R}^n : \| \mathbf{x} \| < 1\}$. *Ihr Rand ist die Einheitssphäre* $\{\mathbf{x} \in \mathbb{R}^n : \| \mathbf{x} \| = 1\}$.

Aufgabe 1.4 (Maximum-Norm) *Zeigen Sie, dass die Abbildung*

$$\mathbb{R}^n \to \mathbb{R}, \quad \mathbf{x} \mapsto \| \mathbf{x} \|_m := max(|x_1|, ..., |x_n|)$$

eine Norm auf $\mathbb{R}^n$ *ist.*

Aufgabe 1.5 *Zeigen Sie:*
a) $M \subset \mathbb{R}^n$ *ist offen* $\quad \Leftrightarrow \quad M \cap \partial M = \emptyset$.
b) $M \subset \mathbb{R}^n$ *ist abgeschlossen* $\quad \Leftrightarrow \quad \partial M \subset M$.
c) ∂M *ist abgeschlossen.*

Aufgabe 1.6 *Sei* M *eine offene Teilmenge des* $\mathbb{R}^2$. *Beweisen Sie, dass der Rand* ∂M *von* M *keine nichtleere offene Teilmenge des* $\mathbb{R}^2$ *enthält.*

Aufgabe 1.7 *Sei* A *eine abgeschlossene Teilmenge des* $\mathbb{R}^n$ *mit* $\mathbf{0} \notin A$. *Zeigen Sie, dass es eine positive reelle Zahl* ϵ *gibt mit*

$$\| \mathbf{a} \| \geq \epsilon \text{ für jedes } \mathbf{a} \in A.$$

Aufgabe 1.8 *a) Skizzieren Sie* $G := G_1 \cup G_2 \cup G_3 \setminus G_4$, *wobei*

$$\begin{aligned}
G_1 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1 \text{ und } x \leq 0\}, \\
G_2 &= \{(x,y) \in \mathbb{R}^2 : x^2 + \left(y - \frac{3}{4}\right)^2 \leq \left(\frac{1}{4}\right)^2 \text{ und } x \geq 0\}, \\
G_3 &= \{(x,y) \in \mathbb{R}^2 : x^2 + \left(y + \frac{3}{4}\right)^2 \leq \left(\frac{1}{4}\right)^2 \text{ und } x \geq 0\}, \\
G_4 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < \frac{1}{4} \text{ und } x \leq 0\}.
\end{aligned}$$

b) Ist G *abgeschlossen, offen?*

Aufgabe 1.9 *Zeigen Sie durch ein Beispiel, dass die Vereinigung unendlich vieler abgeschlossener Teilmengen von* $\mathbb{R}$ *nicht abgeschlossen sein muss.*

Aufgabe 1.10 *Es sei* $f : \mathbb{R} \to \mathbb{R}$ *eine beschränkte Funktion, deren Graph in* $\mathbb{R}^2$ *abgeschlossen ist. Beweisen Sie, dass für jede gegen* $0 \in \mathbb{R}$ *konvergente Folge* $(a_n)_{n\geq 1}$ *reeller Zahlen die Folge* $(f(a_n))_{n\geq 1}$ *eine gegen* $f(0)$ *konvergente Teilfolge besitzt.*

Aufgabe 1.11 *Es sei* $\mathbf{p}$ *ein Punkt und* M *eine offene Teilmenge des* $\mathbb{R}^n$. *Beweisen Sie, dass auch*

$$\{\mathbf{p}+\mathbf{x}\,|\,\mathbf{x}\in M\}$$

eine offene Teilmenge des $\mathbb{R}^n$ *ist.*

Aufgabe 1.12 *Beweisen Sie, dass* $(0,0)$ *ein Randpunkt der Teilmenge*

$$\left\{\left(x, sin(\frac{1}{x})\right)\,|\,x\in\mathbb{R},\,x>0\right\}$$

des $\mathbb{R}^2$ *ist.*

Aufgabe 1.13 **a)** *Es sei* $I\subset\mathbb{R}$ *eine beliebige Teilmenge und* V *der* $\mathbb{R}$*-Vektorraum der auf* I *stetigen und beschränkten Funktionen. Zeigen Sie, dass*

$$\parallel f\parallel_s := \sup_{x\in I}|f(x)|$$

eine Norm auf V *ist.*
b) *Nun sei* $I\subset\mathbb{R}$ *ein offenes Intervall und* $W\subset V$ *der* $\mathbb{R}$*-Vektorraum der auf* I *differenzierbaren und beschränkten Funktionen derart, dass auch* f' *auf* I *beschränkt ist. Zeigen Sie, dass*

$$\parallel f\parallel_d := \sup_{x\in I}|f(x)|+\sup_{x\in I}|f'(x)|$$

eine Norm auf W *ist.*

1.3 Konvergenz, kompakte Mengen

In diesem Abschnitt behandeln wir konvergente Folgen von Punkten im $\mathbb{R}^n$. Als erstes habe ich dabei ein Problem mit der Notation. In der eindimensionalen Analysis haben wir Folgen von Punkten immer $(x_\nu)_{\nu\in\mathbb{N}}$ geschrieben, oder so ähnlich. Hier ist aber x_ν die Bezeichnung für die ν-te Komponente des Vektors $\mathbf{x}$. Deswegen werden wir Folgen von Vektoren als $(\mathbf{x}^{(k)})_{k\in\mathbb{N}}$ schreiben. Der k-te Vektor dieser Folge ist also

$$\mathbf{x}^{(k)}=\begin{pmatrix} x_1^{(k)}\\ \vdots\\ x_n^{(k)}\end{pmatrix}\in\mathbb{R}^n.$$

Definition 1.7 *Eine Folge* $(\mathbf{x}^{(k)})_{k\in\mathbb{N}}$ *von Vektoren* $\mathbf{x}^{(k)}\in\mathbb{R}^n$ *konvergiert gegen den Vektor* $\mathbf{a}\in\mathbb{R}^n$, *wenn die Folge der Abstände* $d(\mathbf{x}^{(k)},\mathbf{a})$ *eine Nullfolge ist, d.h., wenn*

$$\lim_{k\to\infty} d(\mathbf{x}^{(k)},\mathbf{a})=0.$$

Konvergiert die Folge $(\mathbf{x}^{(k)})$ gegen $\mathbf{a} \in \mathbb{R}^n$, so ist dieser Vektor $\mathbf{a}$ durch die Folge eindeutig bestimmt:

Beweis: Sei $\mathbf{b} \in \mathbb{R}^n$ noch ein Vektor, gegen den dieselbe Folge $\mathbf{x}^{(k)}$ konvergiert. Für alle $\epsilon > 0$ gibt es also ein $N \in \mathbb{N}$ mit

$$d(\mathbf{x}^{(k)}, \mathbf{a}) < \epsilon \text{ und } d(\mathbf{x}^{(k)}, \mathbf{b}) < \epsilon \text{ falls } k > N.$$

Mit der Dreiecksungleichung folgt daraus

$$d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{x}^{(k)}) + d(\mathbf{x}^{(k)}, \mathbf{b}) < 2\epsilon$$

für alle $\epsilon > 0$. Das geht nur, wenn $d(\mathbf{a}, \mathbf{b}) = 0$ ist, also wenn $\mathbf{a} = \mathbf{b}$ ist. $\square$

Diese Bemerkung rechtfertigt (wie im ersten Semester) die Schreibweise

$$\mathbf{a} = \lim_{k \to \infty} \mathbf{x}^{(k)}$$

dafür, dass die Folge $\mathbf{x}^{(k)}$ gegen $\mathbf{a}$ konvergiert.

Man kann die Theorie der konvergenten Folgen von Vektoren im $\mathbb{R}^n$ genauso aufbauen, wie wir das im ersten Semster für Folgen reeller Zahlen taten. Man muss nur überall den Absolutbetrag $|x_\nu - a|$ durch den Abstand $d(\mathbf{x}^{(k)}, \mathbf{a})$ ersetzen. Und weil alles ganz genauso geht, wie in der Dimension eins, hat sich so um 1960 herum die Gewohnheit eingebürgert, in der Analysis-Vorlesung gar nicht erst die eindimensionale Analysis zu behandeln, sondern gleich alles im $\mathbb{R}^n$ zu machen. Das war damals der Zug der Zeit. Ich finde soetwas eine didaktische Katastrophe.

Wir wollen den umgekehrten Weg gehen, und alles soweit wie nur irgend möglich auf die uns nunmehr (hoffentlich) vertraute eindimensionale Analysis zurückführen. Dazu dient der folgende Satz.

Satz 1.1 *Eine Folge $(\mathbf{x}^{(k)})$ von Vektoren im $\mathbb{R}^n$ konvergiert genau dann gegen einen Vektor $\mathbf{a} \in \mathbb{R}^n$, wenn für $\nu = 1, ..., n$ die Folge der ν-ten Komponenten der Vektoren $\mathbf{x}^{(k)}$ gegen a_ν konvergiert, also*

$$\lim_{k \to \infty} \mathbf{x}^{(k)} = \mathbf{a} \quad \Leftrightarrow \quad \lim_{k \to \infty} x_\nu^{(k)} = a_\nu \text{ für } \nu = 1, ..., n.$$

Dem Beweis schicken wir ein Abschätzungs-Lemma voraus.

Satz 1.2 (Abschätzungs-Lemma) *Für $\mathbf{x} \in \mathbb{R}^n$ gilt*

$$\max_{\nu=1}^{n} |x_\nu| \leq \| \mathbf{x} \| \leq \sqrt{n} \cdot \max_{\nu=1}^{n} |x_\nu|.$$

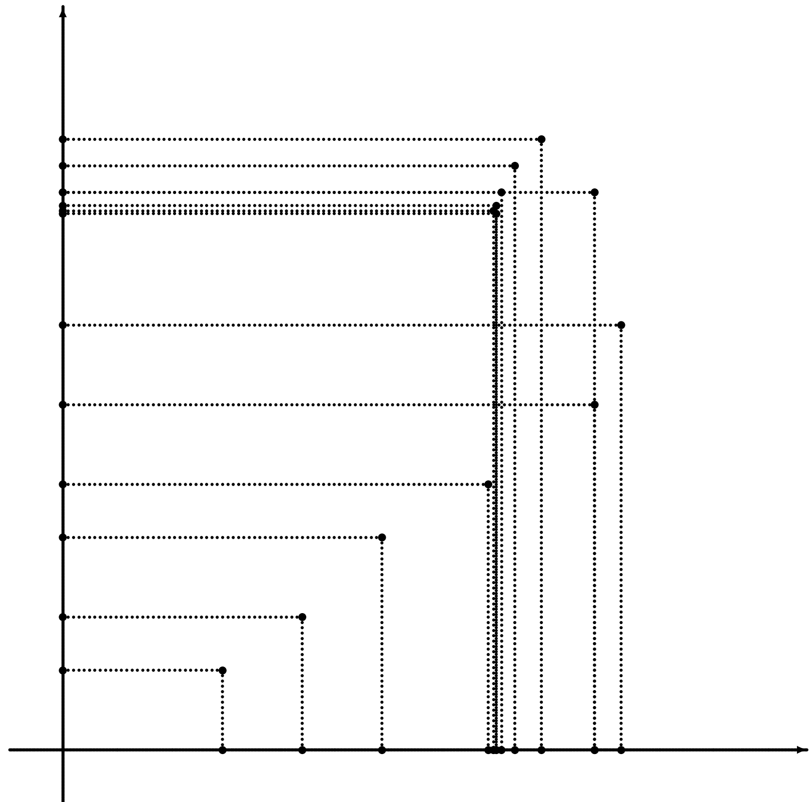
Beweis. Für $\nu = 1, ..., n$ ist

$$|x_\nu| = \sqrt{x_\nu^2} \leq \sqrt{\sum_{\mu=1}^{n} x_\mu^2} = \| \mathbf{x} \| .$$

Das war die erste Ungleichung, und jetzt kommt die zweite:

$$\| \mathbf{x} \| = \sqrt{\sum_{\nu=1}^{n} x_\nu^2} \leq \sqrt{n \cdot \max_{\nu=1}^{n} x_\nu^2} = \sqrt{n} \cdot \max_{\nu=1}^{n} |x_\nu|.$$

$\square$



Eine Folge im $\mathbb{R}^2$ und ihre beiden Komponentenfolgen

Beweis von Satz 1.1). a) Es sei $\mathbf{a} = \lim \mathbf{x}^{(k)}$ vorausgesetzt. Nach dem Abschätzungs-Lemma ist für $\nu = 1, ..., n$ stets $|x_\nu^{(k)} - a_\nu| \leq d(\mathbf{x}^{(k)}, \mathbf{a})$, also

$$0 \leq \lim_{k\to\infty} |x_\nu^{(k)} - a_\nu| \leq \lim_{k\to\infty} d(\mathbf{x}^{(k)}, \mathbf{a}) = 0.$$

Somit konvergiert die ν-te Komponentenfolge $(x_\nu^{(k)})_{k\in\mathbb{N}}$ gegen a_ν.

b) Jetzt nehmen wir umgekehrt an, dass für $\nu = 1, ..., n$ gilt

$$\lim_{k\to\infty} x_\nu^{(k)} = a_\nu.$$

Zu jedem $\epsilon > 0$ gibt es also Schranken $N_1, ..., N_n$ mit

$$|x_\nu^{(k)} - a_\nu| < \epsilon \text{ für } k > N_\nu.$$

Für $k \geq \max_{\nu=1}^n N_\nu$ ist also

$$d(\mathbf{x}^{(k)}, \mathbf{a}) \leq \sqrt{n} \cdot \max_{\nu=1}^{n} |x_\nu^{(k)} - a_\nu| < \sqrt{n} \cdot \epsilon.$$

Das bedeutet

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{a}. \qquad \square$$

Satz 1.3 (Konvergenz und offene Mengen) *Für eine Teilmenge $U \subset \mathbb{R}^n$ sind äquivalent:*
i) U ist offen.
ii) Ist $\mathbf{x} \in U$ und $(\mathbf{x}^{(k)})_{k\in\mathbb{N}}$ eine Folge von Vektoren die gegen $\mathbf{x}$ konvergiert, so enthält U alle Folgenvektoren $\mathbf{x}^{(k)}$ bis auf endlich viele davon.

Beweis. i) $\Rightarrow$ ii): Nach Definition der offenen Mengen gibt es eine Kugel K um $\mathbf{x}$ mit $K \subset U$. Es sei r deren Radius. Nach Definition der Konvergenz gibt es ein $N(r) \in \mathbb{N}$ mit

$$d(\mathbf{x}^{(k)}, \mathbf{a}) < r \text{ für } k > N(r).$$

Für $k > N(r)$ gehören also alle Folgen-Vektoren $\mathbf{x}^{(k)}$ zur Kugel K, und damit zu U.

ii) $\Rightarrow$ i): Es sei $\mathbf{x} \in U$. Zu zeigen ist, dass es eine Kugel um $\mathbf{x}$ mit einem Radius $r > 0$ gibt, die ganz zu U gehört. Wäre dies nicht der Fall, dann gäbe es zu jedem $r = 1/k$, $k \in \mathbb{N}$, einen Vektor $\mathbf{x}^{(k)} \in \mathbb{R}^n$ mit $d(\mathbf{x}, \mathbf{x}^{(k)}) < 1/k$, der nicht zu U gehört. Die Folge $\mathbf{x}^{(k)}$ konvergiert gegen $\mathbf{x}$, und damit haben wir einen Widerspruch zu ii). $\square$

Satz 1.4 (Konvergenz und abgeschlossene Mengen) *Für eine Menge $A \subset \mathbb{R}^n$ sind äquivalent:*
i) A ist abgeschlossen.
ii) Ist $(\mathbf{x}^{(k)})_{k\in\mathbb{N}}$ eine Folge von Vektoren in A, die gegen einen Vektor $\mathbf{x} \in \mathbb{R}^n$ konvergiert, so gehört $\mathbf{x}$ zu A.

Beweis. i) $\Rightarrow$ ii): Nach Definition der abgeschlossenen Mengen ist das Komplement $U := \mathbb{R}^n \setminus A$ offen. Würde $\mathbf{a}$ zu diesem Komplement U gehören, so würden nach Satz 1.3 auch alle Folgenvektoren $\mathbf{x}^{(k)}$, bis auf endlich viele, zu U gehören. Nach Voraussetzung ist dies aber nicht der Fall. Also gehört $\mathbf{a}$ nicht zu U, sondern zu A.

ii) $\Rightarrow$ i): Wir zeigen Eigenschaft ii) aus Satz 1.3 für $U := \mathbb{R}^n \setminus A$. Sei also $\mathbf{x}^{(k)} \in \mathbb{R}^n$ eine Folge, die gegen $\mathbf{x} \in U$ konvergiert. Wenn Eigenschaft ii) aus Satz 1.3 nicht gilt, dann gibt es eine unendliche Teilfolge der Folge $(\mathbf{x}^{(k)})$, deren Punkte zu A gehören. Auch diese Teilfolge konvergiert gegen $\mathbf{x} \notin A$, im Widerspruch zu ii). $\square$

Definition 1.8 *Eine Teilmenge $B \subset \mathbb{R}^n$ heißt beschränkt, wenn es eine Kugel $K \subset \mathbb{R}^n$ gibt, die B enthält.*

Nach dieser Definition sind z.B. alle Kugeln beschränkt, aber keine Gerade ist es. Eine Umformulierung der Definition ist: Die Menge B ist beschränkt, wenn es ein Schranke R gibt mit $\| \mathbf{b} \| < R$ für alle $\mathbf{b} \in B$.

Satz 1.5 (Konvergente und beschränkte Folgen) *a) Jede konvergente Folge im $\mathbb{R}^n$ ist beschränkt.*
b) (Bolzano-Weierstraß) Jede beschränkte Folge im $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.

Beweis. a) Die Folge $(\mathbf{x}^{(k)})$ konvergiere gegen $\mathbf{a}$. Dann gibt es also eine Kugel $K_{\mathbf{a}}$ um $\mathbf{a}$ und ein $N \in \mathbb{N}$ mit $\mathbf{x}^{(k)} \in K_{\mathbf{a}}$ für $k > N$. Außerdem gibt es eine Kugel K', welche die endlich vielen Folgen-Vektoren $\mathbf{x}^{(1)}, ..., \mathbf{x}^{(N)}$ enthält. Wir können eine Kugel K wählen, welche die beiden Kugeln $K_{\mathbf{a}}$ und K' enthält. In dieser Kugel K liegen dann alle Vektoren $\mathbf{x}^{(k)}$ der Folge.

b) Die Folge $(\mathbf{x}^{(k)})$ sei beschränkt, es gelte etwa $\| \mathbf{x}^{(k)} \| < R$ für alle k. Nach dem Abschätzungs-Lemma ist dann auch jede Komponentenfolge $(x_\nu^{(k)})_{k\in\mathbb{N}}$ eine beschränkte Folge reeller Zahlen:

$$|x_\nu^{(k)}| = |x_\nu^{(k)} - 0| \leq d(\mathbf{x}^{(k)}, \mathbf{0}) = \| \mathbf{x}^{(k)} \|$$

Nach dem Satz von Bolzano-Weierstraß für Folgen reeller Zahlen aus dem ersten Semester besitzt die erste Komponentenfolge $(x_1^{(k)})_{k\in\mathbb{N}}$ eine konvergente Teilfolge. Nachdem wir zur entsprechenden Teilfolge der Folge $\mathbf{x}^{(k)}$ übergehen, können wir also annehmen, dass die erste Komponentenfolge $(x_1^{(k)})$ konvergiert.

Diesen Schritt wiederholen wir jetzt für die Folge $(x_2^{(k)})$ der zweiten Komponenten. Nachdem wir nochmal zu einer Teilfolge übergehen, können wir also annehmen, dass die beiden Komponentenfolgen $(x_1^{(k)})$ und $(x_2^{(k)})$ konvergieren. Und wenn wir diesen Schritt n mal durchgeführt haben, haben wir also eine Teilfolge der ursprünglichen Folge $\mathbf{x}^{(k)}$ gefunden, deren n Komponentenfolgen $(x_\nu^{(k)})$, $\nu = 1, ..., n$, alle konvergieren. Nach Satz 1.1 konvergiert also diese Teilfolge auch selbst. □

Die folgende Definition brauchen wir vor allem für den Satz vom Maximum im $\mathbb{R}^n$:

Definition 1.9 *Eine Teilmenge $M \subset \mathbb{R}^n$ heißt* kompakt, *wenn sie* beschränkt *und* abgeschlossen *ist.*

Nach dieser Definition sind beispielsweise alle abgeschlossenen Kugeln kompakt. Offene Kugeln sind nicht kompakt, weil sie nicht abgeschlossen sind. Geraden sind nicht kompakt, weil sie nicht beschränkt sind.

Satz 1.6 (Folgen-Kompaktheit) *Für eine Teilmenge $M \subset \mathbb{R}^n$ sind äquivalent:*
i) M ist kompakt.
ii)Jede Folge $(\mathbf{x}^{(k)})_{k\in\mathbb{N}}$ von Vektoren $\mathbf{x}^{(k)} \in M$ besitzt eine konvergente Teilfolge mit Grenzwert in M.

Beweis. i) $\Rightarrow$ ii): Sei also $(\mathbf{x}^{(k)})$ eine Folge von Vektoren aus M. Weil M beschränkt ist, ist auch die Folge beschränkt, und besitzt nach Satz 1.5b) eine konvergente Teilfolge. Diese Teilfolge liegt auch wieder in M. Weil M auch abgeschlossen ist, gehört der Grenzwert dieser Folge nach Satz 1.4 selbst auch zu M.

ii) $\Rightarrow$ i): Wir müssen zeigen, dass M beschränkt und abgeschlossen ist.

Beschränktheit: Wenn M nicht beschränkt ist, dann gibt es zu jeder natürlichen Zahl $k \in \mathbb{N}$ einen Punkt $\mathbf{x}^{(k)} \in M$ mit $\| \mathbf{x}^{(k)} \| > k$. Keine Teilfolge von $\mathbf{x}^{(k)}$ ist beschränkt und deswegen ist auch keine dieser Teilfolgen konvergent, im Widerspruch zu ii).

Abgeschlossenheit: Wir weisen Eigenschaft ii) aus Satz 1.4 nach. Sei also $\mathbf{x}^{(k)}$ eine Folge von Punkten in M, die gegen $\mathbf{x} \in \mathbb{R}^n$ konvergiert. Nach ii) besitzt diese Folge eine Teilfolge, die gegen einen Punkt in M konvergiert. Aber $\mathbf{x}$ ist auch der Grenzwert dieser Teilfolge, und damit gehört $\mathbf{x}$ zu M. □

Schließlich noch eine sehr allgemeine Charakterisierung kompakter Mengen:

Satz 1.7 (Heine-Borel) *Für eine Teilmenge $M \subset \mathbb{R}^n$ sind äquivalent:*
i) M ist kompakt.
ii) Ist M enthalten in einer Vereinigung $\bigcup_{i\subset I} U_i$ offener Mengen $U_i \subset \mathbb{R}^n$, dann gibt es eine endliche Teilmenge $J \subset I$ mit $M \subset \bigcup_{i\in J} U_i$.

Beweis i) $\Rightarrow$ ii): Weil M kompakt, und damit beschränkt ist, ist M enthalten in einem Quader

$$Q = I_1 \times ... \times I_n, \quad I_\nu = [a_\nu, b_\nu] \subset \mathbb{R}.$$

Jedes Intervall I_ν halbieren wir als

$$I_\nu = I'_\nu \cup I''_\nu \text{ mit } m_\nu := \frac{1}{2}(a_\nu + b_\nu),\ I'_\nu := [a_\nu, m_\nu],\ I''_\nu := [m_\nu, b_\nu].$$

Mit den Produkten der $2n$ Intervalle $I'_1, I''_1, ..., I'_n, I''_n$ kann man 2^n Quader halber Kantenlängen im R^n definieren, deren Vereinigung der urspr̈ugliche Quader Q ist. Für diese 2^n Quader möchte ich mir jetzt keine Notation überlegen, weil mir die zu kompliziert ist. Nennen wir sie einfach die 2^n Quader der ersten Generation.

Kommen wir zur Behauptung. Wenn sie falsch ist, gibt es keine endliche Teilmenge $J \subset I$ so, dass Q in der Vereinigung $\bigcup_{i \in J} U_i$ enthalten ist. Dann kann es auch nicht für jeden Teilquader der ersten Generation ein solche endliche Teilmenge $J \subset U$ geben derart, dass der Teilquader in $\bigcup_{i \in J} U_i$ enthalten ist. Mindestens ein Teilquader $Q^{(1)}$ der ersten Generation ist in keiner solchen endlichen Vereinigung enthalten.

Wenden wir unsere Aufmerksamkeit diesem Teilquader zu und zerlegen ihn in 2^n Teilquader der zweiten Generation mit wieder halbierten Kantenlängen. Und wieder muss es mindestens einen Teilquader $Q^{(2)}$ der zweiten Generation geben, der in keiner endlichen Vereinigung $\bigcup_{i \in J} U_i$ enthalten ist. Dies Verfahren iterieren wir und erhalten eine unendliche Folge von Teilquadern

$$Q \supset Q^{(1)} \supset Q^{(2)} \supset ... \supset Q^{(k)} \supset ...$$

mit jeweils halbierten Kantenlängen so, dass kein Teilquader $Q^{(k)}$ in einer endlichen Vereinigung $\bigcup_{i \in J} U_i$ enthalten ist. Diese Teilquader werden schnell ziemlich klein. Das quantifizieren wir jetzt:

Es sei

$$d := \max_{\nu=1}^{n}(b_\nu - a_\nu)$$

die maximale Kantenlänge des Quaders Q. Dann ist

$$\frac{d}{2^k}$$

die maximale Kantenlänge des Teilquaders $Q^{(k)}$. Nach dem Abschätzungs-Lemma ist

$$\| \mathbf{x} - \mathbf{y} \| \leq \frac{d}{2^k}\sqrt{n}$$

für je zwei Punkte $\mathbf{x}, \mathbf{y} \in Q^{(k)}$. Nun fixieren wir in jedem Quader $Q^{(k)}$ einen Punkt $\mathbf{x}^{(k)} \in Q^{(k)}$, etwa den Mittelpunkt. Für $l, m \geq k$ ist dann $\mathbf{x}^{(l)}, \mathbf{x}^{(m)} \in Q^{(k)}$ und

$$\| \mathbf{x}^{(l)} - \mathbf{x}^{(m)} \| \leq \frac{d}{2^k}\sqrt{n}.$$

Damit ist die Folge $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$ eine Cauchy-Folge und konvergiert gegen einen Punkt $\mathbf{x} \in \mathbb{R}^n$. Weil M abgeschlossen ist, liegt $\mathbf{x}$ in M. Für alle $l \geq k$ ist $\mathbf{x}^{(l)} \in Q^{(l)} \subset Q^{(k)}$. Weil $Q^{(k)}$ abgeschlossen ist, gehört $\mathbf{x} = \lim \mathbf{x}^{(l)}$ zu $Q^{(k)}$.

Jetzt wird es eng: Für alle k ist $\mathbf{x} \in Q^{(k)}$ und damit für alle $\mathbf{y} \in Q^{(k)}$

$$\| \mathbf{y} - \mathbf{x} \| \leq \frac{d}{2^k}\sqrt{n}.$$

Der Quader $Q^{(k)}$ ist enthalten in der abgeschlossenen Kugel um $\mathbf{x}$ vom Radius $d\sqrt{n}/2^k$. Sei U_i eine offene Menge, die $\mathbf{x}$ enthält. Dann gibt es eine offene Kugel von einem Radius $r > 0$ um $\mathbf{x}$, die ganz in U_i enthalten ist. Ist k so groß, dass

$$\frac{d}{2^k}\sqrt{n} < r,$$

dann ist der Quader $Q^{(k)}$ enthalten in der *einzigen* offenen Menge U_i. Das ist der lang ersehnte Widerspruch zu unserer kunstvollen Konstruktion der Teilquader.

ii) $\Rightarrow$ i): Wir müssen zeigen: M ist beschränkt und abgeschlossen.

Beschränkt: Wir betrachten die offenen Kugeln

$$U_i := \{\mathbf{x} \in R^n : \parallel \mathbf{x} \parallel < i\} \text{ mit } i \in \mathbb{N}.$$

Offensichtlich ist

$$M \subset \mathbb{R}^n = \bigcup_{i \in \mathbb{N}} U_i.$$

Nach Eigenschft ii) liegt M also schon in der Vereinigung von endlich vielen dieser Kugeln, damit ind der größten Kugel von diesen endlich vielen, und M ist beschränkt.

Abgeschlossen: Sei $\mathbf{x}^{(k)} \in M$ eine Folge, die gegen einen Punkt $\mathbf{x} \in \mathbb{R}^n$ konvergiert. Zu zeigen ist $\mathbf{x} \in M$. Angenommen, das wäre nicht der Fall. Für alle $\mathbf{y} \in M$ ist dann

$$d_{\mathbf{y}} := \parallel \mathbf{x} - \mathbf{y} \parallel > 0.$$

Und $\mathbf{y}$ liegt nicht in der abgeschlossen Kugel vom Radius $d_{\mathbf{y}}/2$. Jeder Punkt $\mathbf{y} \in M$ gehört zu einer offenen Menge

$$U_r := \{\mathbf{y} \in \mathbb{R}^n : \parallel \mathbf{x} - \mathbf{y} \parallel > r\}$$

mit $r > 0$. Das heißt

$$M \subset \bigcup_{r>0} U_r.$$

Nach ii) liegt dann M schon in der Vereinigung von endlich vielen Mengen U_r. Ist ρ der kleinste der endlich vielen zugehörigen Radien r, so folgt, dass M die abgeschlossene Kugel um $\mathbf{x}$ von diesem Radius ρ nicht schneidet. Aber das steht im Widerspruch zu

$$\lim \mathbf{x}^{(k)} = \mathbf{x}.$$ □

Aufgabe 1.14 *a) Zeigen Sie, dass die Vereinigung endlich vieler kompakter Mengen $K_i \subset \mathbb{R}^n$ wieder kompakt ist.*
b) Es seien $A \subset \mathbb{R}^m$, $B \subset \mathbb{R}^n$ kompakte Mengen. Zeigen Sie dass auch $A \times B \subset \mathbb{R}^{m+n}$ kompakt ist.
c) Es sei C eine kompakte und A eine abgeschlossene Teilmenge des $\mathbb{R}^n$. Beweisen Sie, dass auch $C \cap A$ eine kompakte Teilmenge des $\mathbb{R}^n$ ist.

Aufgabe 1.15 *Die Teilmenge $M \subset \mathbb{R}^n$ sei kompakt. Weiter seien $A_i \subset M$, $i \in I$, Teilmengen, die im R^n abgeschlossen sind, so, dass*

$$\bigcap_{i \in I} A_i = \emptyset.$$

Zeigen Sie: Es gibt eine endliche Teilmenge $J \subset I$ so, dass schon

$$\bigcap_{i \in J} A_i = \emptyset.$$

1.4 Stetigkeit

Die Stetigkeit einer Abbildung $F : U \to V$, $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ kann man ganz genauso definieren, wie die Stetigkeit einer Funktion einer reellen Variablen im ersten Semster, sogar auf zwei Weisen:

Sei dazu $\mathbf{x} \in U$ und $\mathbf{y} = F(\mathbf{x}) \in V$.

Definition 1.10 (ϵ-δ-Stetigkeit)) *Die Abbildung F heißt stetig in $\mathbf{x}$, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert mit*

$$\mathbf{x}' \in U,\ d(\mathbf{x}', \mathbf{x}) < \delta \quad \Rightarrow \quad d(F(\mathbf{x}'), F(\mathbf{x})) < \epsilon.$$

Definition 1.11 (Folgen-Stetigkeit) *Die Abbildung F heißt stetig in $\mathbf{x}$, wenn für jede Folge $(\mathbf{x}^{(k)})_{k \in \mathbb{N}}$ von Punkten $\mathbf{x}^{(k)} \in U$ mit*

$$\lim_{k \to \infty} \mathbf{x}^{(k)} = \mathbf{x}$$

gilt

$$\lim_{k \to \infty} F(\mathbf{x}^{(k)}) = \mathbf{y}.$$

Beide Definitionen sind - wie im ersten Semester auch - äquivalent:

Satz 1.8 *Die Abbildung F ist in $\mathbf{x}$ genau dann folgenstetig, wenn sie ϵ-δ-stetig ist.*

Beweis. „ϵ-δ-Stetigkeit $\Rightarrow$ Folgenstetigkeit“: Sei F ϵ-δ-stetig in $\mathbf{x}$. Sei $\mathbf{x}^{(k)}$ eine Folge in U, die gegen $\mathbf{x}$ konvergiert. Wir zeigen, dass die Bildfolge $F(\mathbf{x}^{(k)})$ gegen $F(\mathbf{x}) = \mathbf{y}$ konvergiert. Dazu müssen wir uns ein $\epsilon > 0$ vorgeben. Nach Voraussetzung gibt es ein $\delta > 0$ mit $d(\mathbf{x}', \mathbf{x}) < \delta \ \Rightarrow \ d(F(\mathbf{x}'), \mathbf{y}) < \epsilon$. Weil die Folge $\mathbf{x}^{(k)}$ gegen $\mathbf{x}$ konvergiert, gibt es zu diesem δ ein N mit

$$k > N \quad \Rightarrow \quad d(\mathbf{x}^{(k)}, \mathbf{x}) < \delta.$$

Für $k > N$ ist dann auch

$$d(F(\mathbf{x}^{(k)}), \mathbf{y}) < \epsilon.$$

Damit ist die Folgen-Stetigkeit nachgewiesen.

„Folgenstetigkeit $\Rightarrow$ ϵ-δ-Stetigkeit“: Diese Richtung geht mit Widerspruch. Nehmen wir also an, die Abbildung F sei folgenstetig in $\mathbf{x}$, aber nicht ϵ-δ-stetig. Dass die ϵ-δ-Eigenschaft nicht erfüllt ist, heißt:

Es gibt ein $\epsilon > 0$ derart, dass
zu jedem $\delta > 0$ ein
$\mathbf{x}' \in U$ existiert mit $d(\mathbf{x}', \mathbf{x}) < \delta$
und $d(F(\mathbf{x}'), \mathbf{y}) > \epsilon$.

Spielen wir das einmal durch für $\delta := 1/k$. Das $\mathbf{x}' \in U$ mit $d(\mathbf{x}', \mathbf{x}) < \delta = 1/k$, das dazu existiert, nennen wir $\mathbf{x}^{(k)}$. Dann ist $d(\mathbf{x}^{(k)}, \mathbf{x}) < 1/k$ eine Nullfolge, und die Folge $\mathbf{x}^{(k)}$ konvergiert gegen $\mathbf{x}$. Für die Bildvektoren $F(\mathbf{x}^{(k)})$ der Vektoren dieser Folge gilt immer $d(F(\mathbf{x}^{(k)}), \mathbf{y}) > \epsilon$. Dabei ist das ϵ das vom Anfang des Beweises, also im Beweis fest. Deswegen kann die Folge $F(\mathbf{x}^{(k)})$ unmöglich gegen $\mathbf{y}$ konvergieren, im Widerspruch zur vorausgesetzten Folgenstetigkeit. $\square$

Beispiel 1.13 *Jede lineare Abbildung $F : \mathbb{R}^n \to \mathbb{R}^m$ ist stetig. Ist nämlich $A \in M(m \times n, \mathbb{R})$ die darstellende Matrix dieser Abbildung, dann folgt mit der Matrix-Norm (Beispiel 1.2)*

$$\| F(\mathbf{x} - \mathbf{x}') \| = \| A \cdot (\mathbf{x} - \mathbf{x}') \| \leq \| A \| \cdot \| \mathbf{x} - \mathbf{x}' \| < \epsilon,$$

wenn

$$\| \mathbf{x} - \mathbf{x}' \| < \delta := \frac{\epsilon}{\| A \|}.$$

Falls hier allerdings $\| A \| = 0$ sein sollte, dann ist A die Nullmatrix, F ist die Null-Abbildung und als konstante Abbildung stetig.

Satz 1.9 (Hintereinanderschalten stetiger Abbildungen) *Es seien $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$, $W \subset \mathbb{R}^p$ Mengen und*

$$F : U \to V, \ G : V \to W$$

stetige Abbildungen. Dann ist die zusammengesetzte Abbildung

$$G \circ F : U \to W$$

auch wieder stetig.

Beweis. Wir benützen die Folgenstetigkeit. Sei also $\mathbf{x} \in U$ und $\mathbf{x}^{(k)}$ eine Folge von Punkten in U mit Grenzwert $\mathbf{x}$. Wegen der Stetigkeit von F konvergiert die Bildfolge $\mathbf{y}^{(k)} := F(\mathbf{x}^{(k)})$ gegen $\mathbf{y} = F(\mathbf{x})$. Wegen der Stetigkeit von G konvergiert deren Bildfolge $G(\mathbf{y}^{(k)})$ gegen $G(\mathbf{y})$. Insgesamt konvergiert also

$$(G \circ F)(\mathbf{x}^{(k)}) = G(F(\mathbf{x}^{(k)})) = G(\mathbf{y}^{(k)})$$

gegen

$$G(\mathbf{y}) = G(F(\mathbf{x})) = (G \circ F)(\mathbf{x}).$$

Damit ist die Folgenstetigkeit der Abbildung $G \circ F$ nachgewiesen. □

Mit dem nächsten Satz spielen wir die Stetigkeit von Abbildungen auf die Stetigkeit ihrer Komponentenfunktionen zurück.

Satz 1.10 (Stetigkeit und Komponentenfunktionen) *Eine Abbildung $F : U \to V$, $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$, ist genau dann stetig im Punkt $\mathbf{x} \in U$, wenn alle ihre n Komponentenfunktionen $F_\nu : U \to \mathbb{R}$ in diesem Punkt $\mathbf{x}$ stetig sind.*

Beweis. Wir betrachten eine Folge $(\mathbf{x}^{(k)})$ von Vektoren $\mathbf{x}^{(k)} \in U$ und die Bildfolge $F(\mathbf{x}^{(k)}) \in V$. Nach Satz 1.1 konvergiert diese Bildfolge genau dann gegen $F(\mathbf{x})$ wenn die n Komponentenfolgen $F(\mathbf{x}^{(k)})_\nu = F_\nu(\mathbf{x}^{(k)})$ gegen $F_\nu(\mathbf{x})$ konvergieren. Also ist die Abbildung F folgenstetig in $\mathbf{x}$, genau dann, wenn ihre n Komponentenfolgen in $\mathbf{x}$ folgenstetig sind. □

Beispiel 1.14 *Beispiele für stetige Funktionen: Die Koordinatenfunktionen $\mathbf{x} \to x_\nu$ sind stetig, denn wenn eine Folge $\mathbf{x}^{(k)}$ gegen $\mathbf{x}$ konvergiert, so konvergiert nach Satz 1.1 jede ihrer Komponentenfunktionen $x_\nu^{(k)}$ gegen x_ν.*

Sind $f, g : U \to \mathbb{R}$ stetige Funktionen, so zeigt man mit der Folgenstetigkeit, und den Grenzwert-Rechenregeln aus dem ersten Semster, dass auch die Funktionen

$$f \pm g, \quad f \cdot g, \quad f/g \text{ wo } g \neq 0$$

wieder stetig sind.

Daraus folgt, dass alle Polynome

$$\sum a_{\nu_1,...,\nu_n} x_1^{\nu_1} \cdot ... \cdot x_n^{\nu_n}$$

in n Veränderlichen stetig sind. Polynome in $n > 1$ Veränderlichen sind sehr gefährliche Tiere. Es ist sogar schon kompliziert, solche Polynome überhaupt nur richtig hinzuschreiben. Ein besonders einfaches Polynom aber ist z.B.

$$f(x_1, ..., x_n) = x_1^2 + ... + x_n^2 = \parallel \mathbf{x} \parallel^2 .$$

Mit den elementaren Funktionen aus der eindimensionalen Analysis und den Koordinatenfunktionen kann man durch Hintereinanderschaltung alle Funktionen zusammensetzen, die man so braucht. Als Beispiel betrachten wir die Norm

$$\parallel \mathbf{x} \parallel = (g \circ f)(\mathbf{x})$$

mit

$$f : \mathbb{R}^n \to [0, \infty[\subset \mathbb{R}, \quad f(\mathbf{x}) = x_1^2 + ... + x_n^2$$

und

$$g : [0, \infty[\to \mathbb{R}, \quad g(y) = \sqrt{y}.$$

Damit ist gezeigt, dass die Funktion

$$\mathbb{R}^n \to \mathbb{R}, \quad \mathbf{x} \to \parallel \mathbf{x} \parallel$$

stetig ist.

Für den nächsten Satz erinnern wir uns an die Definition von Bild- und Urbild-Mengen. Sei dazu $F : U \to \mathbb{R}^n$ eine Abbildung. Ist $M \subset U$ eine Teilmenge, so heißt

$$F(M) = \{F(\mathbf{x}) : \ \mathbf{x} \in M\}$$

die Bildmenge von M. Ist $A \subset \mathbb{R}^n$ eine Teilmenge, so heißt

$$F^{-1}(A) = \{\mathbf{x} \in U : \ F(\mathbf{x}) \in A\}$$

die Urbildmenge von A.

Satz 1.11 (Fundamental-Eigenschaften stetiger Abbildungen) *Es sei $M \subset \mathbb{R}^m$ und $F : M \to \mathbb{R}^n$ stetig.*
a) Es sei $M \subset \mathbb{R}^m$ abgeschlossen. Ist $A \subset \mathbb{R}^n$ abgeschlossen, so ist das Urbild $F^{-1}(A) \subset M$ auch wieder abgeschlossen in $\mathbb{R}^m$.
b) Es sei $M \subset \mathbb{R}^m$ offen. Ist $B \subset \mathbb{R}^n$ offen, so ist das Urbild $F^{-1}(B) \subset M$ auch wieder offen in $\mathbb{R}^m$.
c) Ist $M \subset \mathbb{R}^m$ kompakt, so ist das Bild $F(M) \subset \mathbb{R}^n$ auch wieder kompakt.

Beweis. a) Es sei $\mathbf{x}^{(k)} \in F^{-1}(A)$ eine Folge, die gegen einen Punkt $\mathbf{x} \in \mathbb{R}^m$ konvergiert. Weil M abgeschlossen ist, gehört der Grenzwert $\mathbf{x}$ zu M. Damit ist $F(\mathbf{x})$ definiert. Weil F stetig ist, konvergiert die Bildfolge $F(\mathbf{x}^{(k)}) \in \mathbb{R}^n$ gegen $F(\mathbf{x})$. Weil A abgeschlossen ist, gehört $F(\mathbf{x})$ zu A und $\mathbf{x}$ zu $F^{-1}(A)$. Damit ist Eigenschaft ii) aus Satz 1.4 für die Menge $F^{-1}(A)$ nachgewiesen.

b) Es sei $\mathbf{x} \in F^{-1}(B)$ und $\mathbf{x}^{(k)} \in R^n$ eine Folge, die gegen $\mathbf{x}$ konvergiert. Weil M offen ist, gibt es ein $k_1 \in \mathbb{N}$ so, dass alle $\mathbf{x}^{(k)}$, $k \geq k_1$, zu M gehören. Weil F stetig ist, konvergiert die Bildfolge $F(\mathbf{x}^{(k)})$, $k \geq k_1$, gegen $F(\mathbf{x}) \in B$. Weil B offen ist, gibt es ein $k_2 \geq k_1$ so, dass alle Bildpunkte $F(\mathbf{x}^{(k)})$, $k \geq k_2$, zu B gehören. Dann gehören alle Vektoren $\mathbf{x}^{(k)}$, $k \geq k_2$, zu $F^{-1}(B)$ und Eigenschaft ii) aus Satz 1.3 ist für die Menge $F^{-1}(B)$ nachgewiesen.

c) Wir zeigen die Heine-Borel Überdeckungseigenschaft aus Satz 1.7 für die Bildmenge $F(M)$. Seien also offene Mengen $U_i \subset \mathbb{R}^n$, $i \in I$, gegeben mit

$$F(M) \subset \bigcup_{i \in I} U_i.$$

Dann ist $A_i := \mathbb{R}^n \setminus U_i$ abgeschlossen und $F^{-1}(A_i) \subset M$ auch abgeschlossen nach a). Wir betrachten die offenen Mengen $V_i := \mathbb{R}^m \setminus F^{-1}(A_i) \subset \mathbb{R}^m$. Für jedes $i \in I$ ist $F^{-1}(U_i) = M \cap V_i$. Daraus folgt

$$M \subset \bigcup_{i \in I} V_i.$$

Nach Heine-Borel gibt es eine endliche Teilmenge $J \subset I$ mit

$$M \subset \bigcup_{i \in J} V_i \quad \text{und} \quad F(M) \subset \bigcup_{i \in J} U_i.$$

□

Dieser Satz 1.11 ist furchtbar abstrakt. Aber man kann viel mit ihm machen. Schauen wir uns einige Anwendungen an.

Beispiel 1.15 *Sei $f : \mathbb{R}^n \to \mathbb{R}$ eine stetige Funktion. Wenden wir Satz 1.11a) auf die abgeschlossene Menge $A = \{0\} \subset \mathbb{R}$ an, so sehen wir*

$$f^{-1}(A) = \{\mathbf{x} \in \mathbb{R}^n : \, f(\mathbf{x}) = 0\}$$

ist eine abgeschlossene Menge. Die Nullstellenmenge einer stetigen Funktion ist also immer abgeschlossen! Nehmen wir etwa eine Koordinatenfunktion x_ν, so sehen wir: Jede Koordinaten-Hyperebene $\{x_\nu = 0\}$ ist abgeschlossen. Oder nehmen wir $f = \parallel \mathbf{x} \parallel - r$, so finden wir: Die Sphäre

$$f^{-1}(\{0\}) = \{\mathbf{x} \in \mathbb{R}^n : \parallel \mathbf{x} \parallel = r\}$$

ist abgeschlossen.

Beispiel 1.16 *Wenden wir jetzt Satz 1.11b) auf ein offenes Intervall $]-\infty, c[\subset \mathbb{R}$ an, so sehen wir: Jede Menge*

$$\{\mathbf{x} \in \mathbb{R}^n : \, f(\mathbf{x}) < c\},$$

also die Menge der Punkte, wo f Werte $< c$ hat, ist offen. Ein Spezialfall sind die offenen Kugeln ($f(\mathbf{x}) = \parallel \mathbf{x} - \mathbf{a} \parallel$, $c = r > 0$).

Ein Spezialfall von Satz 1.11c) ist der Satz vom Maximum:

Satz 1.12 (vom Maximum) *Eine stetige Funktion $f : K \to \mathbb{R}$ auf einer kompakten Menge K nimmt auf K ihr Maximum an.*

Beweis. Die Bildmenge $f(K) \subset \mathbb{R}$ ist kompakt.

Dann ist sie auf jeden Fall beschränkt. Deswegen existiert $m := sup(f(K))$, das Supremum dieser Bildmenge.

Außerdem ist $f(K) \subset \mathbb{R}$ abgeschlossen. Nach Definition des Supremums aus dem ersten Semester gibt es eine Folge $(y^{(k)})$ von Punkten $y^{(k)} \in f(K)$, die gegen das Supremum konvergiert. Deswegen gehört m zur Menge $f(K)$ und ist deren Maximum.

Wegen $m \in f(K)$ gibt es also ein $\mathbf{x} \in K$ mit $f(\mathbf{x}) = m$. Wegen $m \geq y$ für alle $y \in f(K)$ ist also $f(\mathbf{x}) \geq f(\mathbf{x}') = y$ für alle $\mathbf{x}' \in K$. Die Funktion f nimmt also in $\mathbf{x} \in K$ ihren größten Wert an. □

Wie in der Dimension 1 definiert man:

Definition 1.12 (Gleichmäßige Stetigkeit) *Es sei $M \subset \mathbb{R}^m$. Die Abbildung $F : M \to \mathbb{R}^n$ heißt* gleichmäßig stetig, *wenn zu jeden $\epsilon > 0$ ein $\delta(\epsilon) > 0$ existiert, so, dass für alle $\mathbf{x}_1, \mathbf{x}_2 \in M$ mit $d(\mathbf{x}_1, \mathbf{x}_2) < \delta(\epsilon)$ gilt $d(F(\mathbf{x}_1), F(\mathbf{x}_2)) < \epsilon$.*

Satz 1.13 *Ist $M \subset \mathbb{R}^m$ kompakt, so ist jede stetige Abbildung $F : M \to \mathbb{R}^n$ gleichmäßig stetig.*

Beweis. Sei $\epsilon > 0$ vorgegeben. Zu jedem $\mathbf{x} \in M$ gibt es dann ein $\delta(\mathbf{x}, \epsilon)$ mit

$$\xi \in M,\ d(\xi, \mathbf{x}) < \delta(\mathbf{x}, \epsilon) \quad \Rightarrow \quad d(F(\xi), F(\mathbf{x})) < \frac{\epsilon}{2}.$$

Für $\mathbf{x} \in M$ sei $U_{\mathbf{x}}$ die offene Kugel um $\mathbf{x}$ mit dem Radius $r_{\mathbf{x}} := \delta(\mathbf{x}, \epsilon)/2$. Dann ist

$$M \subset \bigcup_{\mathbf{x} \in M} \{\mathbf{x}\} \subset \bigcup_{\mathbf{x} \in M} U_{\mathbf{x}}.$$

Nach Heine-Borel gibt es endlich viele Punkte $\mathbf{x}_i \in M$, $i \in J$, mit

$$M \subset \bigcup_{i \in J} U_{\mathbf{x}_i}.$$

Wir definieren

$$\delta(\epsilon) := \min_{i \in J} r_i.$$

Nun seien $\mathbf{y}_1, \mathbf{y}_2 \in M$ gegeben mit $d(\mathbf{y}_1, \mathbf{y}_2) < \delta(\epsilon)$. Der Punkt $\mathbf{y}_1 \in M$ liegt in einer der Kugeln $U_{\mathbf{x}_i}, i \in J$. Wegen

$$d(\mathbf{y}_2, \mathbf{x}_i) \leq d(\mathbf{y}_2, \mathbf{y}_1) + d(\mathbf{y}_1, \mathbf{x}_i) < \delta(\epsilon) + \frac{1}{2}\delta(\mathbf{x}_i, \epsilon) \leq \delta(\mathbf{x}_i, \epsilon)$$

gehört mit $\mathbf{y}_1$ auch $\mathbf{y}_2$ zur Kugel um $\mathbf{x}_i$ vom Radius $\delta(\mathbf{x}_i, \epsilon)$. Daraus folgt

$$d(F(\mathbf{y}_1), F(\mathbf{y}_2)) \leq d(F(\mathbf{y}_1), F(\mathbf{x}_i)) + d(F(\mathbf{x}_i), F(\mathbf{y}_2)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

Aufgabe 1.16 *Gegeben sei die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ durch*

$$f(x, y) = \begin{cases} exp\left(-\frac{y^2}{x}\right) & falls \quad x \neq 0 \\ 1 & falls \quad x = 0 \end{cases}.$$

a) Zeigen Sie, dass für jedes $s \in \mathbb{R}$ die Funktion

$$\mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x, sx)$$

stetig an der Stelle 0 ist.
b) Ist f stetig an der Stelle $(0,0)$?

Aufgabe 1.17 *Gegeben sei die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(0,0) = 0$ und $f(x,y) = \frac{x^2 y}{x^4+y^2}$ für jedes $(x,y) \in \mathbb{R}^2$ mit $(x,y) \neq (0,0)$. Zeigen Sie:*
a) Für jedes $(a,b) \in \mathbb{R}$ mit $(a,b) \neq (0,0)$ ist die Funktion $g : \mathbb{R} \to \mathbb{R}$, $t \mapsto f(ta, tb)$, stetig an der Stelle $t = 0$.
b) f ist nicht stetig an der Stelle $(0,0)$.

Aufgabe 1.18 *Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ sei definiert durch*

$$f(x,y) := \begin{cases} 0 & \text{für} \quad (x,y) = (0,0) \\ \dfrac{x \cdot |y|}{x^2+y^2} & \text{für} \quad (x,y) \neq (0,0) \end{cases}$$

Ist die Funktion f stetig in $(0,0)$?

Aufgabe 1.19 *Gegeben sei die Funktion*

$$f(x,y) = e^{\frac{x}{x^2+y^2}} \quad (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

a) Zeigen Sie unter Verwendung von Polarkoordinaten, dass f in jeder Kreisscheibe um den Nullpunkt alle positiven Werte annimmt.
b) Untersuchen Sie, ob f in $(0,0)$ stetig fortsetzbar ist.

Aufgabe 1.20 *Es sei C eine nichtleere, kompakte Teilmenge des $\mathbb{R}^2$ und $f : \mathbb{R}^2 \to \mathbb{R}^2$ eine stetige Funktion mit $f(x) > 0$ für jedes $x \in C$. Beweisen Sie, dass es ein $\epsilon \in \mathbb{R}$ gibt mit $\epsilon > 0$ und*

$$f(C) \cap]-\epsilon, \epsilon[= \emptyset.$$

1.5 Metrische Räume

Definition 1.13 *Es seien* $X \in \mathbb{R}^m$ *und* $Y \subset \mathbb{R}^n$ *Teilmengen. Wir bezeichnen mit* $C^0(X, Y)$ *die (ziemlich große) Menge aller stetigen Abbildungen* $F : X \to Y$*. Insbesondere ist* $C^0(X, \mathbb{R}^n)$ *der* Vektorraum *aller stetigen Abbildungen* $F : X \to \mathbb{R}^n$*.*

Ist F stetig, so ist auch die Funktion

$$\| F(\mathbf{x}) \| = \sqrt{F_1(\mathbf{x})^2 + ... + F_n(\mathbf{x})^2}$$

stetig auf X. Ist hier X kompakt, so nimmt die stetige Funktion

$$\| F(\mathbf{x}) \|$$

auf X ihr Maximum an (Satz 1.12).

Definition 1.14 *Ist* $X \in \mathbb{R}^m$ *kompakt und* $F : X \to \mathbb{R}^n$ *stetig, so heißt*

$$\| F \|_X := \max_{\mathbf{x} \in X} \| F(\mathbf{x}) \|$$

Die Maximum-Norm *von* F *auf* X*.*

Satz 1.14 *Diese Maximum-Norm ist eine Norm, d.h., sie hat die folgenden Eigenschaften :*

N1: *Es ist* $\| F \| \geq 0$ *und* $\| F \| = 0$ *nur dann, wenn* $F \equiv 0$ *die Nullabbildung ist.*

N2: *Für* $c \in \mathbb{R}$ *ist* $\| c \cdot F \| = |c| \cdot \| F \|$*. (Hier ist* $c \cdot F$ *die Abbildung* $\mathbf{x} \mapsto c \cdot F(\mathbf{x})$*.)*

N3: *Für je zwei Abbildungen* $F, G \in C^0(X, \mathbb{R}^n)$ *gilt die Dreiecksungleichung*

$$\| F + G \| \leq \| F \| + \| G \| .$$

Beweis. Die drei Aussagen ergeben sich, wenn man für jeden Bildvektor $F(\mathbf{x}) \in \mathbb{R}^n$ die entsprechenden Aussagen aus Abschnitt 1.2 anwendet. Ich möchte nur den Beweis für die Dreiecksungleichung ausführen:

Für jeden Vektor $\mathbf{x} \in X$ ist mit der Dreiecksungleichung aus Abschnitt 1.2

$$\| F(\mathbf{x}) + G(\mathbf{x}) \| \leq \| F(\mathbf{x}) \| + \| G(\mathbf{x}) \| .$$

Dann folgt

$$\begin{aligned} \| F + G \|_X &= \max_{\mathbf{x} \in X} \| F(\mathbf{x}) + G(\mathbf{x}) \| \\ &\leq \max_{\mathbf{x} \in X} (\| F(\mathbf{x}) \| + \| G(\mathbf{x}) \|) \\ &\leq \max_{\mathbf{x} \in X} \| F(\mathbf{x}) \| + \max_{\mathbf{x} \in X} \| G(\mathbf{x}) \| \\ &= \| F \|_X + \| G \|_X . \end{aligned}$$

□

Definition 1.15 *Wie oben sei $X \in \mathbb{R}^m$ kompakt. Für je zwei Abbildungen sei*

$$d(F,G)_X := \| F - G \|_X \ .$$

Satz 1.15 *Die Funktion $d(F,G)_X$ ist eine Metrik auf $C^0(X, \mathbb{R}^m)$, d.h., es gelten die folgenden Regeln:*

M1: Es ist stets $d(F,G) \geq 0$ und $d(F,G) = 0$ nur, wenn $F = G$, d.h., $F(\mathbf{x}) = G(\mathbf{x})$ für alle $\mathbf{x} \in X$.

M2: Für $F, G \in C^0(X, \mathbb{R}^m)$ ist $d(F,G) = d(G,F)$.

M3: Es gilt die Dreiecksungleichung

$$d(F,H) \leq d(F,G) + d(G,H) \text{ für alle } F,G,H \in C^0(X, \mathbb{R}^m).$$

Beweis. Die drei Regeln folgen sofort aus den entsprechenden Eigenschaften der Norm. Wieder möchte ich nur die Dreiecksungleichung ausführen: Es ist mit der Dreiecksungleichung für die Norm

$$d(F,H) = \| F - H \| = \| (F-G) + (G-H) \| \leq \| F - G \| + \| G - H \| = d(F,G) + d(G,H). \qquad \square$$

Definition 1.16 *Es sei M eine Menge und $d : M \times M \to \mathbb{R}$ eine Abbildung, welche die drei Eigenschaften aus Satz 1.15 besitzt. Dann heißt d eine* Metrik *auf M und (M,d) heißt ein* metrischer Raum.

Der Begriff des metrischen Raums ist so allgemein, dass man damit sehr viel leeres Stroh dreschen kann. Andererseits ist er der natürliche Rahmen für die Konvergenz. In einem metrischen Raum (M,d) kann man alles genauso definieren wie im $\mathbb{R}^n$:

- Eine Kugel um $u \in M$ ist eine Menge $\{v \in M : d(u,v) < r\}$;
- Eine Folge $u_\nu \in M$ konvergiert gegen $u \in M$, wenn zu jedem $\epsilon > 0$ ein $N(\epsilon) \in \mathbb{N}$ existiert so, dass für $\nu > N(\epsilon)$ alle Punkte u_ν in der ϵ-Kugel um u liegen. Das heißt: Die Zahlen $d(u_\nu, u)$ bilden eine Nullfolge.
- Eine Menge $U \subset M$ heißt offen, wenn mit jeden $u \in U$ auch eine Kugel um U von einem Radius $r > 0$ ganz zu U gehört. Eine Menge $A \subset M$ heißt abgeschlossen, wenn $M \setminus A$ offen ist.
- Wie für $M = \mathbb{R}^n$ mit der euklidischen Metrik zeigt man: $A \subset M$ ist genau dann abgeschlossen, wenn für jede Folge (u_ν) von Punkten $u_\nu \in A$, die in M konvergiert, der Grenzwert $u = \lim(u_\nu)$ wieder in A liegt.

Beispiel 1.17 *Wir betrachten den metrischen Raum $C^0([a,b], \mathbb{R}^n)$ mit seiner Metrik*

$$d(F,G) = \max_{x \in [a,b]} \{\| F(x) - G(x) \|\}.$$

Die Folge F_ν von stetigen Abbildungen konvergiert gegen $G : [a,b] \to \mathbb{R}^n$, wenn zu jedem $\epsilon > 0$ ein $N(\epsilon)$ existiert mit $d(F_\nu, G) < \epsilon$ für $\nu > N(\epsilon)$. Das heißt explizit: Für alle $x \in [a,b]$ ist

$$\| F_\nu(x) - G(x) \| < \epsilon.$$

Im Fall $n = 1$ sagten wir dazu: Die Funktionenfolge f_ν konvergiert gleichmäßig auf $[a,b]$ gegen g. Deswegen kann man die Konvergenz in $C^0([a,b], \mathbb{R}^n)$, oder allgemeiner in $C^0(X,Y), X \subset \mathbb{R}^m, Y \subset \mathbb{R}^n$, auch gleichmäßige Konvergenz nennen.

Beispiel 1.18 *Die Menge $Y \subset \mathbb{R}^n$ sei abgeschlossen. Dann ist*

$$C^0(X, Y) \subset C^0(X, \mathbb{R}^n)$$

eine abgeschlossene Teilmenge.

Beweis. Es sei $F_\nu : X \to Y$ eine Folge von Abbildungen, die in $C^0(X, \mathbb{R}^n)$ gegen $G : X \to \mathbb{R}^n$ konvergiert. Für jedes $\mathbf{x} \in X$ konvergiert dann die Folge von Vektoren $F_\nu(\mathbf{x}) \in \mathbb{R}^n$ gegen $G(\mathbf{x})$. Alle Vektoren $F_\nu(\mathbf{x})$ liegen in Y. Diese Menge ist abgeschlossen in $\mathbb{R}^n$. also ist auch $G(\mathbf{x}) = \lim F_\nu(\mathbf{x}) \in Y$. Damit gehört $G = \lim F_\nu$ zu $C^0(X, Y)$. □

Wie im $\mathbb{R}^n$ gibt es auch in jedem metrischen Raum (M, d) den Begriff der Cauchy-Folge:

Definition 1.17 *Die Folge (u_ν) in M heißt Cauchy-Folge, wenn zu jedem $\epsilon > 0$ ein $N(\epsilon) \in \mathbb{N}$ existiert, so, dass für $\mu, \nu > N(\epsilon)$ gilt:*

$$d(u_\mu, u_\nu) < \epsilon.$$

Jetzt folgt die einzige nicht-triviale Aussage in diesem Abschnitt:

Satz 1.16 *Jede Cauchy-Folge in $C^0(X, \mathbb{R}^n)$ konvergiert.*

Beweis. Es sei (F_ν) eine Cauchy-Folge stetiger Abbildungen $F_\nu : X \to \mathbb{R}^n$. Für jedes $\mathbf{x} \in X$ ist

$$\| F_\nu(\mathbf{x}) - F_\mu(\mathbf{x}) \| \leq \max_{\mathbf{x} \in X} \| F_\mu(\mathbf{x}) - F_\nu(\mathbf{x}) \| = d(F_\mu, F_\nu).$$

Daraus folgt, dass die Bildvektoren $F_\nu(\mathbf{x}) \in \mathbb{R}^n$ eine Cauchy-Folge bilden. Der Raum $\mathbb{R}^n$ ist vollständig. Deswegen konvergiert die Folge $(F_\nu(\mathbf{x}))$ gegen einen Vektor im $\mathbb{R}^n$. Den nennen wir

$$G(\mathbf{x}) := \lim_{\nu \to \infty} F_\nu(\mathbf{x}).$$

Seien $\mu > \nu > N(\epsilon)$. Dann ist für alle $\mathbf{x} \in X$ wegen der Stetigkeit der euklidischen Norm im $\mathbb{R}^n$

$$\| G(\mathbf{x}) - F_\nu(\mathbf{x}) \| = \lim_{\mu \to \infty} \| F_\mu(\mathbf{x}) - F_\nu(\mathbf{x}) \| \leq \epsilon.$$

Deswegen konvergiert die Folge F_ν gegen die Abbildung G bezüglich der Metrik d, wenn wir noch zeigen: $G \in C^0(X, \mathbb{R}^n)$, d.h., G ist stetig.

Der Beweis dafür ist das gleiche $\epsilon/3$-Argument wie beim Beweis dafür, dass die Grenzfunktion einer gleichmäßig konvergenten Folge stetiger Funktionen wieder stetig ist. Wiederholen wir diesen Beweis und geben uns ein festes $\mathbf{x}_0 \in X$ und ein $\epsilon > 0$ vor. Wir wählen $\nu > N(\epsilon/3)$ und haben für F_ν:

$$\| F_\nu(\mathbf{x}) - G(\mathbf{x}) \| < \frac{\epsilon}{3} \text{ für alle } \mathbf{x} \in G.$$

Weil F_ν stetig in $\mathbf{x}_0$ ist, gibt es ein $\delta > 0$ so, dass für alle $\mathbf{x} \in X$ gilt:

$$\| \mathbf{x} - \mathbf{x}_0 \| < \delta \quad \Rightarrow \quad \| F_\nu(\mathbf{x}) - F\nu(\mathbf{x}_0) \| < \frac{\epsilon}{3}.$$

Damit folgt

$$\| G(\mathbf{x}) - G(\mathbf{x}_0) \| \leq \underbrace{\| G(\mathbf{x}) - F_\nu(\mathbf{x}) \|}_{<\epsilon/3} + \underbrace{\| F_\nu(\mathbf{x}) - F_\nu(\mathbf{x}_0) \|}_{<\epsilon/3} + \underbrace{\| F_\nu(\mathbf{x}_0) - G(\mathbf{x}_0) \|}_{<\epsilon/3} < \epsilon.$$

□

Definition 1.18 *Der metrische Raum (M,d) heißt vollständig, wenn jede Cauchyfolge aus M in M konvergiert.*

Beispiel 1.19 *Der $\mathbb{R}^n$ mit der euklidischen Metrik $d(\mathbf{x},\mathbf{y}) = \| \mathbf{x}-\mathbf{y} \|$ ist vollständig. Ist $X \subset \mathbb{R}^m$ kompakt, so ist der Funktionenraum $C^0(X,\mathbb{R}^n)$ mit der Metrik $d(F,G) := \| F-G \|_X$ vollständig.*

Beispiel 1.20 (Exponentialfunktion für Matrizen) *Wir betrachten den $\mathbb{R}$-Vektorraum $M = M(n\times n,\mathbb{R})$ der reellen $n\times n$-Matrizen. Für jede Matrix $A \in M$ konvergiert die Reihe*

$$exp(\| A \|) = \sum_{k=0}^{\infty} \frac{1}{k!} \| A \|^k .$$

Daraus folgt, dass die Partialsummen der Matrizenreihe

$$exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

eine Cauchy-Folge bilden, denn für jeden Reihenabschnitt gilt

$$\| \sum_{k=l}^{m} \frac{1}{k!} A^k \| \leq \sum_{k=l}^{m} \frac{1}{k!} \| A \|^k .$$

Weil M mit der Matrizen-Norm vollständig ist, ist für alle $A \in M$ die Matrix $exp(A) \in M$ definiert.
Ist

$$K = \{A \in M : \| A \| \leq R\} \subset M$$

eine kompakte Kugel, so ist nach Satz 2.16 der Vektorraum $C^0(K,M)$ vollständig. Die Reihenabschnitte der Exponentialreihe kann man für $A \in K$ simultan abschätzen

$$\| \sum_{k=l}^{m} \frac{1}{k!} A^k \|_K \leq \sum_{k=l}^{m} \frac{1}{k!} \| A \|_K^k \leq \sum_{k=l}^{m} \frac{1}{k!} R^k .$$

Deswegen ist die Abbildung

$$K \to M, \quad A \mapsto exp(A)$$

stetig. Weil man M durch immer größer werdende Kugeln ausschöpfen kann, ist die Exponential-Abbildung

$$exp : M(n\times n,\mathbb{R}) \to M(n\times n,\mathbb{R})$$

stetig.

Satz 1.17 *Es sei (M,d) ein vollständiger metrischer Raum und $A \subset M$ abgeschlossen. Dann ist auch der metrische Raum (A,d) vollständig.*

Beweis. Sei (u_ν) eine Cauchy-Folge in A. Dann ist diese Folge auch als Folge in M eine Cauchy-Folge, und damit konvergent gegen ein $u \in M$. Weil A abgeschlossen ist, gehört $u = \lim(u_\nu)$ zu A. $\square$

Beispiel 1.21 *Sei $X \subset \mathbb{R}^m$ kompakt und $Y \subset \mathbb{R}^n$ abgeschlossen. Dann ist $C^0(X,Y)$ vollständig.*

Aufgabe 1.21 *Es sei $I \subset \mathbb{R}$ ein offenes Intervall und W der $\mathbb{R}$-Vektorraum der auf I stetig differenzierbaren Funktionen, die auf I beschränkt sind und dort auch eine beschränkte Ableitung f' besitzen. In Aufgabe 1.13 wurde gezeigt, dass*

$$\| f \|_d := sup_{x\in I}|f(x)| + sup_{x\in I}|f'(x)|$$

eine Norm auf W ist. Zeigen Sie: Der Vektorraum W mit der Metrik

$$d(f,g) := \| f - g \|_d$$

ist vollständig.

Aufgabe 1.22 (Hilbertraum) *Es sei l^2 die Menge aller quadrat-summierbaren reellen Folgen, d.h., die Menge aller reellen Folgen $\alpha = (a_\nu)_{\nu\in\mathbb{N}}$, für welche die Reihe*

$$\sum_{\nu=1}^{\infty} a_\nu^2$$

konvergiert. Zeigen Sie:
a) *Die Menge l^2 ist ein $\mathbb{R}$-Vektorraum und*

$$\| \alpha \| := \sqrt{\sum_{\nu=1}^{\infty} a_\nu^2}$$

ist eine Norm auf diesem Vektorraum.
b) *Der Vektorraum l^2 mit der Metrik*

$$d(\alpha, \beta) := \| \alpha - \beta \|$$

ist vollständig.

Aufgabe 1.23 *Berechnen Sie $exp(t \cdot A)$ für $t \in \mathbb{R}$ und die Matrizen $A =$*

$$N := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad R := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad J := \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Aufgabe 1.24 *a) Es seien A und B zwei reelle $n \times n$-Matrizen, die kommutieren:*

$$A \cdot B = B \cdot A.$$

Zeigen Sie:

$$exp(A + B) = exp(A) \cdot exp(B)$$

und folgern Sie daraus, dass jede Matrix $exp(A)$ invertierbar ist.
b) Es sei A eine beliebige und B eine invertierbare $n \times n$-Matrix. Zeigen Sie:

$$B^{-1} \cdot exp(A) \cdot B = exp(B^{-1} \cdot A \cdot B).$$

1.6 Der Fixpunktsatz von Banach

Banach war ein polnischer Mathematiker. In Warschau gibt es ein Banach-Zentrum für Mathematik.

Definition 1.19 *Es seien (M, d) und (M', d') zwei metrische Räume. Eine Abbildung $F : M \to M'$ heißt Lipschitz-stetig mit Lipschitz-Konstante L, wenn für je zwei Punkte $u, v \in M$ gilt*

$$d'(F(u), F(v)) \leq L \cdot d(u, v).$$

Beispiel 1.22 *Seien $X \subset \mathbb{R}^m$ und $Y \subset \mathbb{R}^n$, jeweils versehen mit der euklidischen Metrik. Ist $F : X \to Y$ Lipschitz-stetig, dann ist F auch stetig.*

Beweis. Wir setzen $\delta(\epsilon) := \epsilon/L$. Sind $\mathbf{u}, \mathbf{v} \in X$ mit $\| \mathbf{u} - \mathbf{v} \| < \delta$, dann ist

$$\| F(\mathbf{u}) - F(\mathbf{v}) \| \leq L \cdot \| \mathbf{u} - \mathbf{v} \| < L \cdot \frac{\epsilon}{L} = \epsilon.$$

Beispiel 1.23 *Es sei $]a, b[\subset \mathbb{R}$ ein offenes Intervall und $f :]a, b[\to \mathbb{R}$ differenzierbar. Ist*

$$f'(x) \leq L \quad \text{für alle} \quad x \in]a, b[,$$

Dann ist f Lipschitz-stetig mit Lipschitz-Konstante L.

Beweis. Es seien $x < y \in]a, b[$. Nach dem Mittelwertsatz der Differentialrechnung gibt es ein $\xi \in]x, y[$ mit

$$|f(x) - f(y)| = |f'(\xi)| \cdot |x - y| \leq L \cdot |x - y|.$$

□

Beispiel 1.24 *Die Funktion $f(x) = x^2$ ist Lipschitz-stetig auf jedem endlichen Intervall $[a, b] \subset \mathbb{R}$, nicht aber auf ganz $\mathbb{R}$.*

Beweis. Auf $[a, b]$ ist

$$|f'(x)| = |2x| \leq L := \max\{2|a|, 2|b|\}$$

und nach Beispiel 1.23 ist f Lipschitz-stetig auf $[a, b]$ mit Lipschitz-Konstante L.

Sei nun $x = 0$ und $y > 0, y \in \mathbb{R}$, beliebig. Wenn y groß genug gewählt wird, dann ist

$$\frac{|f(x) - f(y)|}{|x - y|} = \frac{2y^2}{y} = 2y$$

größer als jede vorgegebene Konstante L. □

Definition 1.20 *Sei (M, d) ein metrischer Raum. Eine Abbildung $F : M \to M$ heißt* kontrahierend, *wenn sie Lipschitz-stetig ist mit einer Lipschitz-Konstante $L < 1$.*

Beispiel 1.25 *Ist $f :]a, b[\to]a, b[$ differenzierbar mit $|f'(x)| \leq L$ für alle x, wo die Konstante $L < 1$ ist, dann ist f kontrahierend.*

Definition 1.21 *Ein Punkt $u \in M$ heißt* Fixpunkt *der Abbildung $F : M \to M$, wenn $F(u) = u$.*

Beispiel 1.26 *Die Funktion*

$$f : \mathbb{R} \to \mathbb{R}, \quad f(x) = x^2$$

hat genau die beiden Fixpunkte $x = 0$ und $x = 1$.

Satz 1.18 (Fixpunktsatz von Banach) *Es sei (M, d) ein vollständiger metrischer Raum und $F : M \to M$ kontrahierend. Dann besitzt F genau einen Fixpunkt $u \in M$.*

Beweis. Nach Voraussetzung ist $F : M \to M$ Lipschitz-stetig mit einer Lipschitz-Konstante $L < 1$.

a) Eindeutigkeit: Es seien $u \neq v \in M$ zwei verschiedene Fixpunkte von F. Dann ist also

$$d(u, v) = d(F(u), F(v)) \leq L \cdot d(u, v) < d(u, v)$$

wegen $d(u, v) > 0$. Aber dies ist ein Widerspruch.

b) Existenz: Wir wählen einen beliebigen Punkt $u_0 \in M$ und betrachten die rekursiv definierte Folge

$$u_1 := F(u_0),\ u_2 := F(u_1),\ ..., u_{n+1} := F(u_n),\ ...$$

in M. Durch Induktion zeigt man

$$d(u_2, u_1) \leq L \cdot d(u_1, u_0), ..., d(u_{n+1}, u_n) \leq L^n \cdot d(u_1, u_0), ...$$

Mit der geometrischen Summenformel folgt daraus für $m < n \in \mathbb{N}$

$$d(u_n, u_m) \leq d(u_{m+1}, u_m) + ... + d(u_n, u_{n-1}) \leq (L^m + ... + L^{n-1}) \cdot d(u_1, u_0) = L^m \cdot \frac{1 - L^{n-m}}{1 - L} \cdot d(u_1, u_0).$$

Wegen $0 \leq L < 1$ ist $0 < 1 - L^{n-m} \leq 1$ und

$$L^m \cdot \frac{1 - L^{n-m}}{1 - L} \leq \frac{L^m}{1 - L}.$$

Damit haben wir bewiesen:

$$d(u_n, u_m) \leq L^m \cdot \frac{d(u_1, u_0)}{1 - L}.$$

Deswegen ist die Folge (u_n) eine Cauchy-Folge in M.

Weil M vollständig vorausgesetzt ist, hat die Folge (u_n) einen Grenzwert $u \in M$. Bei vorgegebenem $\epsilon > 0$ ist also $d(u_n, u) < \epsilon$, wenn n groß genug ist. Aber dann ist auch

$$d(F(u_n), F(u)) \leq L \cdot d(u_n, u) < L \cdot \epsilon < \epsilon.$$

Die Folge $F(u_n)$ konvergiert gegen $F(u)$. Nun ist die Folge

$$(F(u_n))_{n \geq 0} = (u_{n+1})_{n \geq 0} = (u_n)_{n \geq 1}$$

eine Teilfolge der Folge (u_n) und hat wie diese den Grenzwert u. Damit ist

$$F(u) = \lim(F(u_n)) = \lim(u_n) = u,$$

also ist u der gesuchte Fixpunkt. □

Für die Approximation des Fixpunkts u durch die Punkte u_n folgt aus dem Beweis von Satz 1.18 die Abschätzung

Satz 1.19 (Zusatz zu Satz 1.18) *Mit der Notation aus Satz 1.18 gilt*

$$d(u_m, u) \le \frac{L^m}{1-L} \cdot d(u_1, u_0).$$

Beweis. Im Beweis von Satz 1.18 haben wir für $m < n \in \mathbb{N}$ gesehen

$$d(u_m, u_n) \le \frac{L^m}{1-L} \cdot d(u_1, u_0).$$

Zu jedem $\epsilon > 0$ gibt es ein $n > m$ mit $d(u_n, u) < \epsilon$ und

$$d(u_m, u) \le d(u_m, u_n) + d(u_n, u) \le \frac{L^m}{1-L} \cdot d(u_1, u_0) + \epsilon.$$

Weil hier ϵ beliebig klein gewählt werden kann, folgt

$$d(u_m, u) \le \frac{L^m}{1-L} \cdot d(u_1, u_0).$$

□

Fixpunkte einer Abbildung sind etwas ziemlich äsotherisches. Aber in einem Vektorraum V kann man das Problem, eine Gleichung $F(u) = v$ zu lösen, umschreiben in ein Fixpunktproblem. Sei etwa $F : V \to V$ gegeben, $v \in V$ fest, und eine Lösung $u \in V$ für die Gleichung $F(u) = v$ gesucht. Diese Lösung u ist auch eine Lösung der Gleichung

$$F(u) - v = 0, \quad \text{bzw.} \quad F(u) - v + u = u,$$

also ein Fixpunkt der Abbildung $u \mapsto F(u) - v + u$.

Beispiel 1.27 ($\sqrt{2}$-Approximation) *Gesucht sei die Lösung $x \in [1, 2]$ der Gleichung $x^2 = 2$, bzw. eine Approximation dieser Lösung. Es handelt sich also um eine Nullstelle der Gleichung*

$$f(x) = 0 \quad \textit{mit} \quad f(x) = x^2 - 2,$$

bzw. um einen Fixpunkt der Funktion

$$g(x) \quad \textit{mit} \quad g(x) = x^2 - 2 + x.$$

Nun ist $g'(x) = 2x + 1 \ge 3$ im Intervall $[1, 2]$ und die Abbildung g ist nicht kontrahierend. Aber man kann die Funktion $f(x) = x^2 - 2$ ersetzen durch $c \cdot f(x)$ mit $0 \ne c \in \mathbb{R}$, die Nullstelle $\sqrt{2}$ bleibt gleich. Dann wird g ersetzt durch die neue Funktion

$$g(x) = c \cdot (x^2 - 2) + x \quad \textit{mit} \quad g'(x) = 2cx + 1.$$

Für $x > 0$ ist

$$g'(x) < 1 \quad \Leftrightarrow \quad c < 0.$$

Und für $1 \le x \le 2$ ist

$$g'(x) > -1 \Leftrightarrow 2cx > -2 \Leftrightarrow cx > -1 \Leftrightarrow c > -\frac{1}{2}.$$

Wählen wir etwa

$$c := -\frac{1}{4},$$

dann ist

$$g(x) = -\frac{1}{4}(x^2 - 2) + x = \frac{1}{2} + x - \frac{1}{4}x^2, \quad g'(x) = 1 - \frac{x}{2}.$$

Für $1 < x < 2$ ist $0 < g'(x) < 1/2$. Die Funktion g ist auf dem Intervall $[1, 2]$ streng monoton steigend mit

$$g(1) = \frac{5}{4}, \quad g(2) = \frac{3}{2}.$$

Deswegen ist die Abbildung $g : [1, 2] \to [1, 2]$ kontrahierend. Nach dem Fixpunktsatz besitzt g einen Fixpunkt in $[1, 2]$ (die Zahl $\sqrt{2}$, das wissen wir schon). Aber um $\sqrt{2}$ zu approximieren können wir etwa die induktiv definierte Folge

$$x_0 := 2, \quad x_{n+1} = g(x_n) = \frac{1}{2} + x_n - \frac{1}{4}x_n^2$$

wählen. Auf $[1, 2]$ ist $|g'(x)| \leq 1/2$. Somit ist g hier Lipschitz-stetig mit $L = 1/2$. Wegen

$$x_1 = \frac{3}{2}, \quad |x_1 - x_0| = \frac{1}{2},$$

folgt aus Satz 1.19 für die Güte der Approximation

$$|x_n - \sqrt{2}| \leq \frac{(1/2)^n}{1 - 1/2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^n.$$

Beispiel 1.28 (Newton-Verfahren) *Das Newton-Verfahren verwendet man, um eine Nullstelle u einer gegebenen Funktion $f(x)$ numerisch zu approximieren. Die Idee besteht darin, mit einem geeigneten Punkt x_0 zu beginnen, und die Funktion f zu ersetzen durch ihre Tangente im Punkt $(x_0, f(x_0))$. Diese Tangente hat die Gleichung*

$$y = f(x_0) + f'(x_0) \cdot (x - x_0).$$

Sie schneidet die x-Achse dort, wo $y = 0$, d.h.,

$$f'(x_0) \cdot (x - x_0) = -f(x_0), \quad x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Dies führt auf das Iterationsverfahren

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}.$$

Diese Iteration hat gewisse Tücken (s. [Knauf, p. 72]). Ich möchte hier eine Situation präzisieren, wo sie zum Ziel führt.

Dazu setzen wir voraus:

- *Es sei $[a, b] \subset \mathbb{R}$ ein Intervall mit $f(a) < 0, f(b) > 0$.*
- *die Funktion f sei zweimal stetig differenzierbar auf einem offenen Intervall, welches das Intervall $[a, b]$ enthält.*

- *Für $a \le x \le b$ gelte $f'(x) > 0$ und $f''(x) > 0$.*

Aus dem Zwischenwertsatz folgt, dass f im Intervall $[a,b]$ mindestens eine Nullstelle hat. Wegen $f'(x) > 0$ ist f auf diesem Intervall streng monoton steigend, und die Nullstelle u von f in $[a,b]$ ist eindeutig bestimmt.

Zum Iterationsverfahren gehört die Abbildung

$$F : x \mapsto x - \frac{f(x)}{f'(x)}.$$

Wegen $f'(x) > 0$ ist $F(x)$ definiert für alle $x \in [a,b]$. Ein Fixpunkt von F erfüllt

$$F(x) = x \Leftrightarrow x - \frac{f(x)}{f'(x)} = x \Leftrightarrow f(x) = 0 \Leftrightarrow x = u.$$

Wir suchen die Nullstelle von f also als Fixpunkt von F.

Dazu müssen wir zunächst wissen, wann F kontrahierend ist. Weil F zweimal differenzierbar ist, können wir mit dem MWS der Differentialrechnung für $ax \le < y \le b$ abschätzen

$$\left|\frac{F(y)-F(x)}{y-x}\right| = |F'(\xi)| = \left|1 - \frac{f'(\xi)^2 - f(\xi)f''(\xi)}{f'(\xi)^2}\right| = |f(\xi)| \cdot \frac{f''(\xi)}{f'(\xi)^2}$$

mit einem Zwischenpunkt ξ, $x < \xi < y$. Wenn eine Konstante $L < 1$ existiert mit

$$|f(x)| \cdot \frac{f''(x)}{f'(x)^2} < L$$

für alle $x \in [a,b]$, dann ist F kontrahierend. Die stetige Funktion $f''(x)/f'(x)^2$ nimmt auf diesem Intervall ein Maximum m an. Wenn $|f(x)| < L/m$ ist, also wenn a und b nahe genug bei der Nullstelle u sind, dann ist diese Bedingung sicher erfüllt.

Etwas wesentliches kann allerdings noch schiefgehen: Die Abbildung F braucht das Intervall $[a,b]$ nicht in sich abzubilden. Um das zu reparieren, zeigen wir:

Für $u < x \le b$ ist $u < F(x) \le x$.

Beweis. Wegen $x > u$ ist $f(x) > 0$. Weil $f' > 0$ vorausgesetzt ist, haben wir sicher $F(x) < x$. Wir müssen noch zeigen

$$F(x) = x - \frac{f(x)}{f'(x)} > u.$$

Äquivalent dazu ist

$$x - u > \frac{f(x)}{f'(x)} \quad \textit{bzw.} \quad f(x) < (x-u) \cdot f'(x).$$

Wegen $f(u) = 0$ ist

$$f(x) = \int_u^x f'(t)dt = (x-u) \cdot f'(\tau), \quad u < \tau < x.$$

Weil $f'' > 0$ vorausgesetzt ist, ist auch f' streng monoton steigend und $f'(\tau) < f'(x)$. □

Auch wenn wir die Nullstelle u noch nicht kennen, wissen wir jetzt, dass $F : [u,b] \to [u,b]$ eine kontrahierende Abbildung ist. Wir können mit einem beliebigen Startwert x_0, $u < x_0 \le b$ beginnen, rekursiv definieren

$$x_{n+1} := F(x_n),$$

und wissen aus dem Beweis des Banachschen Fixpunkt-Satzes, dass die Folge x_n gegen die Nullstelle u konvergiert.

Für praktische Rechnungen sind natürlich Abschätzungen für $|x_n - u|$ sehr wichtig. Ich möchte darauf nicht mehr eingehen, sondern auf [Knauf, p. 70-72] und [Forster I, p. 122] verweisen.

Beispiel 1.29 *Als eine Art Unter-Beispiel zum vohergehenden Beispiel möchte ich das Newton-Verfahren für die Funktion $f(x) = x^2 - 2$ mit der Nullstelle $u = \sqrt{2}$ analysieren. Hier ist*

$$f'(x) = 2x, \quad f''(x) = 2.$$

Für $x > 0$ sind beide Ableitungen > 0. Weiter ist

$$F(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - 2}{2x} = \frac{x}{2} + \frac{1}{x}.$$

(Kam die damit rekursiv definierte Folge nicht schon mal in einer Übungsaufgabe vor?) Um nachzuweisen, dass F eine kontrahierende Abbildung ist, müssen wir nach Beispiel 1.28 noch abschätzen

$$f(x) \cdot \frac{f''(x)}{f'(x)^2} = (x^2 - 2) \cdot \frac{2}{4x^2} = \frac{x^2 - 2}{2x^2} = \frac{1}{2} - \frac{1}{x^2}.$$

Für $1 \leq x \leq 2$ hat dieser Ausdruck einen Wert zwischen $-1/2$ und $1/4$ und mit $L = 1/2$ läuft das Iterationsverfahren.

Beispiel 1.30 (Vereinfachtes Newton-Verfahren) *Wieder suchen wir eine Nullstelle der differenzierbaren Funktion f. Ausgehend von einem Startwert $(x_0, f(x_0))$ iterieren wir jetzt mit der Abbildung*

$$F(x) := x - \frac{f(x)}{f'(x_0)}.$$

Die Abbleitung im Nenner ist jetzt nicht mehr von x abhängig, sondern konstant. Natürlich müssen wir

$$f'(x_0) \neq 0$$

voraussetzen. Weiter setzen wir voraus:

- *Die Funktion f ist stetig differenzierbar auf einem offenen Intervall, das ein Intervall $I : [x_0 - r, x_0 + r]$ mit $r > 0$ enthält.*
- *Der Wert $f(x_0)$sei so klein, dass*
$$\left| \frac{f(x_0)}{f'(x_0)} \right| \leq \frac{r}{2}.$$
- *Die Zahl $r > 0$ sei so klein, dass für $x \in I$ gilt*
$$\left| \frac{f'(x)}{f'(x_0)} - 1 \right| \leq \frac{1}{2}.$$

Behauptung: Unter den genannten Voraussetzungen ist $F : I \to I$ kontrahierend.

Beweis. Für alle $x_1, x_2 \in I$ ist

$$|F(x_1) - F(x_2)| = \left|x_1 - x_2 - \frac{f(x_1) - f(x_2)}{f'(x_0)}\right| = \left|x_1 - x_2 - \frac{f'(\xi)}{f'(x_0)} \cdot (x_1 - x_2)\right| = |x_1 - x_2| \cdot \left|1 - \frac{f'(\xi)}{f'(x_0)}\right|$$

mit einem Zwischenwert $\xi \in I$ zwischen x_1 und x_2. Nach Voraussetzung ist dann der letzte Faktor $\leq 1/2$ und F ist Lipschitz-stetig mit $L = 1/2$.

Wir müssen noch zeigen, dass $F : I \to I$ abbildet. Sei also $x \in I$. Wir schätzen ab:

$$|F(x) - x_0| \leq |F(x) - F(x_0)| + |F(x_0) - x_0| \leq \frac{1}{2}|x - x_0| + \left|\frac{f(x_0)}{f'(x_0)}\right| \leq \frac{r}{2} + \frac{r}{2} \leq r.$$ □

Wegen der dritten Voraussetzung ist $f'(x) \neq 0$ auf ganz I, und die Nullstelle von f, gegen die das Iterationsverfahren konvergiert, ist eindeutig bestimmt. Es ist leicht nachzurechnen, dass unsere schon oft herangezogene Funktion $f(x) = x^2 - 2$ alle Voraussetzungen mit $x_0 = 2$ und $r = 1$ erfüllt.

Schließlich möchte ich noch zeigen, wie man mit dem Banachschen Fixpunktsatz die Existenz von Umkehrfunktionen beweisen kann. Dass stetig differenzierbare Funktionen mit $f'(x) \neq 0$ auf ihrem ganzen Definitionsintervall eine Umkehrfunktion besitzen, das wissen wir schon. Außerdem wird das gleich beschrieben Verfahren nur lokal funktionieren. Aber in der Theorie mehrerer Veränderlicher ist es eben auch anwendbar.

Sei also eine differenzierbare Funktion $f(x)$ gegeben mit $f(x_0) = y_0$. Wir wollen ein Iterationsverfahren beschreiben, das in der Nähe von (x_0, y_0) gegen die Umkehrfunktion $g(y) = f^{-1}(y)$ konvergiert. Die Idee ist, das vereinfachte Newtonverfahren anzuwenden auf die Funktion $f(x) - y$, die von einem Parameter y abhängt. Die Iterationsformel ist

$$x_0(y) :\equiv x_0, \quad x_{k+1}(y) := x_k(y) - \frac{f(x_k(y)) - y}{f'(x_0)}.$$

Zunächst ist klar: Wenn $x(y)$ ein Fixpunkt von des Verfahrens ist, dann muss $f(x(y)) = y$ gelten. Wir müssen die Voraussetzungen des vereinfachten Newton-Verfahrens für $f(x) - y$ fixieren:

- Auf einem offenen Intervall, welches das Intervall $I = [x_0 - r, x_0 + r]$ enthält, sei $f(x)$ und damit auch $f(x) - y$ stetig differenzierbar.

- Die Zahl r sei so klein, dass auf dem Intervall I gilt

  $$\left|\frac{(f(x) - y)'}{f'(x_0)} - 1\right| = \left|\frac{f'(x)}{f'(x_0)} - 1\right| < \frac{1}{2}.$$

  Das ist unabhängig von y dieselbe Bedingung wie oben.

- Die Voraussetzung

  $$\left|\frac{f(x_0) - y}{f'(x_0)}\right| = \left|\frac{y_0 - y}{f'(x_0)}\right| \leq \frac{r}{2}$$

  schränkt y ein auf das Intervall $J := [y_0 - s, y_0 + s]$ mit

  $$|y - y_0| \leq \frac{r}{2} \cdot |f'(x_0)| =: s.$$

Für jedes y im Intervall J konvergiert das Verfahren gegen ein $x(y)$ mit $f(x(y)) = y$. Damit ist $y \mapsto x(y)$ auf dem Intervall J eine Umkehrfunktion von f. Aber unser Verfahren liefert auch noch, dass diese Umkehrfunktion stetig ist. Dazu definieren wir es im Funktionenraum $C^0(J, I)$ durch

$$g_0(y) := x_0, \quad g_{k+1}(y) := g_k(y) - \frac{f(g_k(y)) - y}{f'(x_0)}.$$

Jede rekusiv definierte Funktion g_k ist stetig. Und die Abbildung

$$F : C^0(J, I) \to C^0(J, I), \quad g \mapsto F(g) := g - \frac{f(g) - y}{f'(x_0)}$$

ist kontrahierend, weil ihre Wirkung für jedes feste y kontrahierend ist. Nach Satz 1.16 ist $C^0(J, I)$ vollständig. Der Banachsche Fixpunktsatz ist anwendbar und liefert als Fixpunkt eine stetige Funktion $g \in C^0(J, I)$.

Aufgabe 1.25 *Für welche $c \in \mathbb{R}$ ist die Abbildung*

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \frac{c}{1 + x^2}$$

kontrahierend?

Aufgabe 1.26 *a) Bestimmen Sie maximale Teil-Intervalle $I \subset \mathbb{R}$, auf denen die Funktion*

$$f(x) := 1 + \frac{1}{x}$$

eine kontrahierende Abbildung definiert.
b) In welchem dieser Intervalle besitzt f welchen Fixpunkt?

Aufgabe 1.27 *Zeigen Sie, dass die Abbildung*

$$F : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \frac{1}{2} \begin{pmatrix} 1 - y \\ 2 + x \end{pmatrix}$$

kontrahierend ist und berechnen Sie ihren Fixpunkt.

Aufgabe 1.28 *a) Zeigen Sie, dass die Funktion*

$$f(x) = x^3 + x$$

mit $x_0 = y_0 = 0$ auf den Intervallen

$$I := \left[-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right] \quad \text{und} \quad J := \left[-\frac{2}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right]$$

die Voraussetzungen des verallgemeinerten Newton-Verfahrens erfüllt.
b) Berechnen Sie mit der Startfunktion $g_0(y) := 0$ die iterierten Funktionen $g_1(y), g_2(y), g_3(y)$.

Aufgabe 1.29 *a) Zeigen Sie, dass die Funktion*

$$f(x) := x^3 - 1$$

mit $x_0 = y_0 = 1$ *auf den Intervallen* $I := [\frac{4}{5}, \frac{6}{5}]$ *und* $J := [\frac{7}{10}, \frac{13}{10}]$ *die Voraussetzungen des vereinfachten Newton-Verfahrens erfüllt.*
b) Berechnen Sie (nicht unbedingt mit der Hand) ausgehend von der Startfunktion $g_0(y) := 1$ *die iterierten Funktionen* g_1, g_2, g_3 *des vereinfachten Newton-Verfahrens.*
c) Entwickeln Sie

$$g(y) := (1 + y)^{1/3}$$

in eine Taylor-Reihe um $y = 0$.
d) Bis zu welcher Potenz von y *stimmen* g_3 *und* g *überein?*

2 Differentialrechnung mit mehreren Veränderlichen

Hier wollen wir uns also um die Ableitung von Funktionen $f(x_1, ..., x_n)$ in mehr als einer Variablen, oder allgemeiner von Abbildungen $F : \mathbb{R}^m \to \mathbb{R}^n$ kümmern. Vieles geht ganz genauso, wie bei Funktionen einer Veränderlichen. Manches geht prinzipiell genauso, ist aber wegen der vielen Indizes komplizierter hinzuschreiben. Aber ein wesentlich neuer Gesichtspunkt ist, dass es verschiedene Differentiationsbegriffe gibt. Fangen wir damit an.

2.1 Partielle Ableitungen

Eine Funktion von zwei Variablen, etwa

$$f(x, y) = x \cdot y, \quad \text{oder } = x^2 + y^2, \quad \text{oder } = x^y$$

kann man nach jeder der beiden Variablen differenzieren. Dabei muss man dann die andere Variable als Konstante auffassen. Diese Art von Ableitung bezeichnet man als partielle Ableitung und schreibt sie mit einem gerundeten d als $\partial/\partial x$ oder $\partial/\partial y$.

Beispiel 2.1 *So ist z.B.*

$$\frac{\partial(xy)}{\partial x} = y, \quad \frac{\partial(x^2+y^2)}{\partial x} = 2x, \quad \frac{\partial(x^y)}{\partial x} = y \cdot x^{y-1}$$

$$\frac{\partial(xy)}{\partial y} = x, \quad \frac{\partial(x^2+y^2)}{\partial y} = 2y, \quad \frac{\partial(x^y)}{\partial y} = ln(x) \cdot x^y.$$

Definition 2.1 *Gegeben sei eine Funktion $f : U \to \mathbb{R}$ mit $U \subset \mathbb{R}^n$ und ein Punkt $\mathbf{x}^{(0)} = (x_1^{(0)}, ..., x_n^{(0)}) \in U$. Diese Funktion $f(x_1, ..., x_n)$ heißt* partiell differenzierbar *nach x_ν in dem gegebenen Punkt, wenn die Funktion einer Variablen*

$$t \mapsto f(x_1^{(0)}, ..., x_{\nu-1}^{(0)}, t, x_{\nu+1}^{(0)}, ..., x_n^{(0)})$$

in $t = x_\nu^{(0)}$ differenzierbar ist. Ihre Ableitung

$$\frac{\partial f}{\partial x_\nu}(\mathbf{x}^{(0)}) := \frac{d}{dt} f(x_0^{(0)}, ..., x_{\nu-1}^{(0)}, t, x_{\nu+1}^{(0)}, ..., x_n^{(0)})\Big|_{t=x_\nu^{(0)}}$$

heißt dann die partielle Ableitung *von f nach x_ν in dem festen Punkt.*

Die Funktion f heißt partiell differenzierbar *in $\mathbf{x}^{(0)}$, wenn sie dort nach allen n Variablen $x_1, ..., x_n$ partiell differenzierbar ist. Dann heißt der Vektor*

$$\nabla f(\mathbf{x}^{(0)}) := \left(\frac{\partial f}{\partial x_1}(\mathbf{x}^{(0)}), ..., \frac{\partial f}{\partial x_n}(\mathbf{x}^{(0)})\right)$$

der Gradient *von f im Punkt $\mathbf{x}_0$.*

Die Funktion f heißt partiell differenzierbar auf der offenen Menge $U \subset \mathbb{R}^n$, wenn sie in jedem Punkt $\mathbf{x} \in U$ partiell differenzierbar ist. Ihre partiellen Ableitungen $\partial f/\partial x_\nu$ sind dann Funktionen auf U, und ihr Gradient ∇f ist ein Vektorfeld auf U.

Beispiel 2.2 *Noch ein Beispiel:* $f(x,y) := \sqrt{x^2+y^2}$, *die Normfunktion. Für festes* $y^{(0)}$ *ist*

$$f(x, y^{(0)}) = \sqrt{x^2 + (y^{(0)})^2}$$

differenzierbar nach x*, falls* $(x, y^{(0)}) \neq (0,0)$, *mit Ableitung*

$$\frac{\partial f}{\partial x} = \frac{2x}{2\sqrt{x^2+(y^{(0)})^2}} = \frac{x}{\sqrt{x^2+(y^{(0)})^2}}.$$

Ebenso ist in $(x^{(0)}, y^{(0)}) \neq (0,0)$

$$\frac{\partial f}{\partial y} = \frac{y}{\sqrt{(x^{(0)})^2+(y^{(0)})^2}}.$$

Und der Gradient dieser Funktion ist

$$\nabla f = \frac{(x,y)}{\sqrt{x^2+y^2}} = \frac{1}{\parallel (x,y) \parallel} \cdot (x,y).$$

Manches geht für die partiellen Ableitungen genauso, wie für die Ableitung in einer Variablen. Dazu gehören insbesondere die Rechenregeln

$$\frac{\partial(a \cdot f + b \cdot g)}{\partial x_\nu} = a \cdot \frac{\partial f}{\partial x_\nu} + b \cdot \frac{\partial g}{\partial x_\nu}, \quad a, b \in \mathbb{R},$$

$$\frac{\partial f \cdot g}{\partial x_\nu} = \frac{\partial f}{\partial x_\nu} \cdot g + f \cdot \frac{\partial g}{\partial x_\nu},$$

$$\frac{\partial(1/f)}{\partial x_\nu} = -\frac{1}{f^2} \cdot \frac{\partial f}{\partial x_\nu}.$$

Auch die die notwendige Bedingung für lokale Extrema ist die gleiche: Ist $f : U \to \mathbb{R}$ eine Funktion, so sagt man, sie nimmt in einem Punkt $\mathbf{x}^{(0)} \in U$ ein Maximum (Minimum) an, wenn

$$f(\mathbf{x}) \leq f(\mathbf{x}^{(0)}) \text{ für alle } \mathbf{x} \in U,$$

(bzw. $\geq$). Sie nimmt ein lokales Maximum (Minimum) an, wenn es eine Kugel K um $\mathbf{x}^{(0)}$ gibt, so dass die Ungleichung für alle $\mathbf{x} \in K$ gilt. Ein (lokales) Extremum ist ein (lokales) Maximum oder Minimum.

Satz 2.1 (Notwendige Bedingung für lokale Extrema) *Die Funktion* $f : U \to \mathbb{R}$ *sei partiell differenzierbar auf der* <u>*offenen*</u> *Menge* $U \subset \mathbb{R}^n$*. Ist* $\mathbf{x}^{(0)}$ *ein lokales Extremum für* f*, so verschwinden dort alle partiellen Ableitungen:*

$$\frac{\partial f}{\partial x_\nu}(\mathbf{x}^{(0)}) = 0 \textit{ für } \nu = 1, ..., n, \quad \textit{bzw.} \quad \nabla f(\mathbf{x}^{(0)}) = \mathbf{0}.$$

Beweis. Weil U offen ist, gibt es eine ganze Kugel mit einem Radius $r > 0$, die noch ganz in U liegt. Die n Funktionen einer Variablen

$$t \mapsto f(x_1^{(0)}, ..., x_{\nu-1}^{(0)}, t, x_{\nu+1}^{(0)}, ..., x_n^{(0)})$$

sind dann differenzierbar auf den Intervallen $(x_\nu^{(0)} - r, x_\nu^{(0)} + r)$ und nehmen in $t = x_\nu^{(0)}$ alle ein lokales Extremum an. Wegen des notwendigen Kriteriums für lokale Extrema aus der Analysis einer Variablen verschwinden ihre Ableitungen in den betreffenden Punkten. Das sind aber die partiellen Ableitungen von f im Punkt $\mathbf{x}^{(0)}$. □

Beispiel 2.3 *Betrachten wir die wohlbekannte Funktion $f(x, y) := x^2 + y^2$. Wenn wir ihr (globales) Minimum im Nullpunkt noch nicht kennen würden, könnten wir es mit dem notwendigen Kriterium für Extrema suchen. Dazu müssten wir die beiden partiellen Ableitungen*

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

$= 0$ *setzen. Das ergibt die Bedingung $x = y = 0$, den Nullpunkt.*

Beispiel 2.4 *Ändern wir die Funktion etwas ab: $f(x, y) = x^2 - y^2$. Die hat jetzt im Nullpunkt kein Extremum mehr nicht einmal ein lokales, denn*

$$f(x, 0) = x^2 > 0 \text{ für } x \neq 0, \quad f(0, y) = -y^2 < 0 \text{ für } y \neq 0.$$

Wenn wir lokale Extrema mit dem notwendigen Kriterium suchen, müssen wir genauso

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = -2y$$

$= 0$ *setzen, und erhalten wieder den Nullpunkt. Das ist aber kein lokales Extremum. Immerhin können wir noch schließen: Außerhalb des Nullpunkts ist das notwendige Kriterium sicher nicht erfüllt, da gibt es keinerlei lokales Extremum.*

Aus Beispiel 2.4 lernen wir, dass das notwendige Kriterium für lokale Extrema keineswegs hinreichend ist. Das war ja auch nicht zu erwarten, weil es in einer Variablen auch nicht hinreichend war. Nur: in einer Variablen war das eine subtile Sache, und hing mit den höheren Ableitungen zusammen. In mehreren Variablen geht das Kriterium schon auf grandiose Weise schief: Es kann für x erfüllt sein, weil da ein Minimum vorliegt, und auch für y, weil da ein Maximum vorliegt. Und das Minimum in x-Richtung, zusammen mit dem Maximum in y-Richtung ergibt nie und nimmer mehr ein Extremum (zumindest, wenn das echte Maxima/Minima waren). Man nennt das dann einen *Sattelpunkt.*

Definition 2.2 *Ein Punkt $\mathbf{x}^{(0)}$ wo $\nabla f = 0$ ist, heißt* kritischer *oder* stationärer *Punkt von f, egal ob dort ein Extremum vorliegt oder nicht.*

Manches ist also ganz anders als in einer Variablen. Dazu gehört auch die Sache mit der linearen Approximation: Eine differenzierbare Funktion $f(x)$ in einer Variablen kann man linear approximieren:

$$f(x + h) = f(x) + f'(x) \cdot h + \varphi(h)$$

mit

$$\lim_{h \to 0} \frac{\varphi(h)}{h} = 0.$$

Die Gerade
$$h \mapsto f(x) + f'(x) \cdot h$$
heißt die Tangente and den Graphen der Funktion f im Punkt $(x, f(x))$.

Für eine partiell differenzierbare Funktion $f(\mathbf{x})$ kann man ganz analog eine *Tangential-Hyperebene* an deren Graphen im Punkt $(\mathbf{x}, f(\mathbf{x}))$ definieren. Das ist die Hyperebene im $\mathbb{R}^{n+1}$, die durch die affin-lineare Abbildung
$$\mathbb{R}^n \ni \mathbf{h} \mapsto f(\mathbf{x}) + \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu = f(\mathbf{x}) + (\nabla f(\mathbf{x}).\mathbf{h})$$
parametrisiert wird. Die approximiert den Funktionsgraphen schön linear in jeder x_ν-Richtung, aber es gibt zwischen diesen n Richtungen noch viele andere Richtungen, gegeben durch Geraden $\mathbb{R}{\cdot}\mathbf{a}$, $\mathbf{a} \in \mathbb{R}^n$. Und in diesen Richtungen kann die Approximation leider sehr böse schiefgehen.

Beispiel 2.5 *Wie dramatisch das aussehen kann, zeigt die Funktion*
$$f(x,y) := \begin{cases} \dfrac{xy}{(x^2+y^2)^2} & \textit{für} \quad (x,y) \neq (0,0) \\ 0 & \textit{für} \quad (x,y) = (0,0). \end{cases}$$
Außerhalb des Nullpunkts ist der Nenner, wie sich das gehört, von 0 *verschieden, und die Funktion ist dort überall partiell differenzierbar. Auf beiden Koordinatenachsen ist* $f(x,y) \equiv 0$, *identisch* $= 0$. *Deswegen ist* f *auch im Nullpunkt partiell differenzierbar mit den partiellen Ableitungen*
$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$$
Aber wenn wir uns dem Nullpunkt in Richtung der Geraden $\mathbb{R} \cdot (a,b)$ *nähern, wird die Funktion*
$$f(ta, tb) = \frac{t^2 \cdot ab}{t^4 \cdot (a^2+b^2)^2} = \frac{1}{t^2} \cdot \frac{ab}{(a^2+b^2)^2}.$$
Falls hier $ab \neq 0$, *dann wächst* $|f(ta,tb)|$ *für* $t \to 0$ *über alle Berge, und* f *ist nie und nimmer linear zu approximieren, schon gar nicht durch die harmlose Tangentialebene*
$$(h,k) \mapsto \frac{\partial f}{\partial x}(0,0) \cdot h + \frac{\partial f}{\partial y}(0,0) \cdot k \equiv 0.$$

Natürlich ist der Grund für dieses absonderliche Verhalten, der dass die Funktion f ziemlich künstlich ist. Bei einigermaßen vernünftigen Funktionen passiert soetwas nicht:

Satz 2.2 (Stetig partiell differenzierbare Funktionen) *Die Funktion* $f : U \to \mathbb{R}$ *sei partiell differenzierbar auf der offenen Menge* $U \subset \mathbb{R}^n$ *und die partiellen Ableitungen* $\partial f/\partial x_\nu, \nu = 1,..,n$, *seien auf* U *stetig. Dann kann man* f *in jedem Punkt* $\mathbf{x} \in U$ *linear approximieren,*
$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + (\nabla f(\mathbf{x}).\mathbf{h}) + \varphi(\mathbf{h})$$
mit
$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\varphi(\mathbf{h})}{\| \mathbf{h} \|} = 0.$$

Beweis. Wir führen den Beweis durch Induktion nach n. Für $n = 1$ ist nichts zu zeigen, weil hier Differenzierbarkeit = lineare Approximierbarkeit und partielle Differenzierbarkeit übereinstimmen.

Sei also $n > 1$. Wir spalten die Vektoren $\mathbf{x} = (\mathbf{x}', x_n) \in \mathbb{R}^n$ auf in einen $(n-1)$-dimensionalen Vektor $\mathbf{x}' \in \mathbb{R}^{n-1}$ und die n-te Komponente x_n. Entsprechend haben wir

$$\mathbf{x} + \mathbf{h} = (\mathbf{x}' + \mathbf{h}', x_n + h_n).$$

Damit wird

$$\begin{aligned}
\varphi(\mathbf{h}) &= f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - (\nabla f(\mathbf{x}).\mathbf{h}) \\
&= f(\mathbf{x}' + \mathbf{h}', x_n + h_n) - f(\mathbf{x}', x_n) - \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu \\
&= \left(f(\mathbf{x}' + \mathbf{h}', x_n + h_n) - f(\mathbf{x}' + \mathbf{h}', x_n) - \frac{\partial f}{\partial x_n}(\mathbf{x}) \cdot h_n \right) + \\
&\quad + \left(f(\mathbf{x}' + \mathbf{h}', x_n) - f(\mathbf{x}', x_n) - \sum_{\nu=1}^{n-1} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu \right)
\end{aligned}$$

In der ersten Klammer sind die ersten $n-1$ Komponenten $\mathbf{x}' + \mathbf{h}'$ des Vektors $\mathbf{x} + \mathbf{h}$ fest, nur die n-te Komponente ändert sich. Wegen der partiellen Differenzierbarkeit von $f(\mathbf{x}', x_n)$ bezüglich x_n ist der MWS der Differentialrechnung anwendbar und ergibt für diese Klammer

$$\left(\frac{\partial f}{\partial x_n}(\mathbf{x}' + \mathbf{h}', x_n + \theta h_n) - \frac{\partial f}{\partial x_n}(\mathbf{x}) \right) \cdot h_n$$

mit $0 < \theta = \theta(h_n) < 1$. Wegen $|h_n| \leq \| \mathbf{h} \|$ bleibt

$$\frac{|h_n|}{\| \mathbf{h} \|} \leq 1$$

für $\mathbf{h} \to \mathbf{0}$ und aus der Stetigkeit der partiellen Ableitung nach x_n in $\mathbf{x}$ folgt

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{1}{\| \mathbf{h} \|} \cdot \left(\frac{\partial f}{\partial x_n}(\mathbf{x}' + \mathbf{h}', x_n + \theta h_n) - \frac{\partial f}{\partial x_n}(\mathbf{x}) \right) \cdot h_n = 0.$$

In der zweiten Klammer ist x_n fest, hier ist f nur eine Funktion der ersten $n-1$ Veränderlichen. Mit der Induktionsannahme folgt hier

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{1}{\| \mathbf{h} \|} \cdot \left(f(\mathbf{x}' + \mathbf{h}', x_n) - f(\mathbf{x}', x_n) - \sum_{\nu=1}^{n-1} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu \right) =$$

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\| \mathbf{h}' \|}{\| \mathbf{h} \|} \cdot \frac{1}{\| \mathbf{h}' \|} \cdot \left(f(\mathbf{x}' + \mathbf{h}', x_n) - f(\mathbf{x}', x_n) - \sum_{\nu=1}^{n-1} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu \right) = 0.$$ □

Die Formel in Satz 2.2 ist genau die analoge Formel zu der linearen Approximation in einer Variablen.

Definition 2.3 *Eine Funktion $f : U \to R$, $U \subset \mathbb{R}^n$ offen, heißt im Punkt $\mathbf{x} \in U$* total differenzierbar, *wenn sie linear approximierbar ist, wenn also die Formel in Satz 2.2 gilt.*

Dieses Wort „total differenzierbar“ hat sich eingebürgert, wahrscheinlich vor allem deswegen, weil es sprachlich so ein schöner Gegensatz zu „partiell differenzierbar“ ist, aber wohl auch deswegen, weil man aus der Eigenschaft „total differenzierbar“ schließen kann, dass die Funktion f in jeder Richtung - nicht nur in den beiden Koordinatenrichtungen - differenzierbar ist:

Satz 2.3 (Stetigkeit, Richtungsableitung) *Ist $f : U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ offen, in $\mathbf{x} \in U$ total differenzierbar, so gilt:*
a) f ist in $\mathbf{x}$ stetig.
b) Für jeden Richtungsvektor $\mathbf{v} \neq \mathbf{0}$ ist die Funktion einer Variablen

$$g(t) = f(\mathbf{x} + t \cdot \mathbf{v})$$

differenzierbar im Punkt $t = 0$ mit Ableitung

$$\left.\frac{dg}{dt}\right|_{t=0} = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot v_\nu = (\nabla f(\mathbf{x}).\mathbf{v}).$$

Beweis. a) Es ist

$$\lim_{\mathbf{h}\to\mathbf{0}} [f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})] = \lim_{\mathbf{h}\to\mathbf{0}} \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu + \lim_{\mathbf{h}\to\mathbf{0}} \varphi(\mathbf{h})$$

mit

$$\lim_{\mathbf{h}\to\mathbf{0}} \varphi(\mathbf{h}) = \lim_{\mathbf{h}\to\mathbf{0}} \frac{\varphi(\mathbf{h})}{\parallel \mathbf{h} \parallel} \cdot \lim_{\mathbf{h}\to\mathbf{0}} \parallel \mathbf{h} \parallel = 0.$$

Daraus folgt

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}),$$

also die Stetigkeit von f in $\mathbf{x}$.

b) Wir setzen ganz einfach $\mathbf{h} = t \cdot \mathbf{v}$ in die Approximationsformel ein:

$$\begin{aligned} g(t) &= f(\mathbf{x}) + \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot t \cdot v_\nu + \varphi(t \cdot \mathbf{v}) \\ &= f(\mathbf{x}) + \left(\sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot v_\nu \right) \cdot t + \varphi(t \cdot \mathbf{v}). \end{aligned}$$

Dabei ist

$$\lim_{t\to 0} \frac{\varphi(t \cdot \mathbf{v})}{|t|} = \lim_{(t\cdot\mathbf{v})\to\mathbf{0}} \frac{\varphi(t \cdot \mathbf{v})}{|t| \cdot \parallel \mathbf{v} \parallel} \cdot \parallel \mathbf{v} \parallel = 0 \cdot \parallel \mathbf{v} \parallel = 0.$$

Also ist $g(t)$ im Punkt $t = 0$ differenzierbar, und für seine Ableitung gilt die angegebene Formel. □

Fassen wir zusammen: Es gelten die Implikationen

$$\boxed{\text{stetig partiell differenzierbar} \quad \Rightarrow \quad \text{total differenzierbar} \quad \Rightarrow \quad \text{partiell differenzierbar.}}$$

Definition 2.4 *Die Ableitung in Satz 2.3*

$$\frac{d}{dt}f(\mathbf{x}+t\cdot\mathbf{v})\Big|_{t=0} = \sum_{\nu=1}^{n}\frac{\partial f}{\partial x_\nu}(\mathbf{x})\cdot v_\nu = (\nabla f(\mathbf{x}).\mathbf{v})$$

heißt Richtungsableitung *der Funktion f im Punkt $\mathbf{x}$ in Richtung des Vektors $\mathbf{v}$. Manchmal benutzt man für diese Richtungsableitung auch die Notation*

$$\partial_{\mathbf{v}}f(\mathbf{x}).$$

Mit der Cauchy-Schwarz-Ungleichung aus der Linearen Algebra sieht man

$$|\partial_{\mathbf{v}}f(\mathbf{x})| \leq \| \nabla f(\mathbf{x}) \| \cdot \| \mathbf{v} \| .$$

Insbesondere ist für jeden Einheitsvektor $\mathbf{v}$, $\| \mathbf{v} \| = 1$,

$$\partial_{\mathbf{v}}f(\mathbf{x}) \leq \| \nabla f(\mathbf{x}) \| .$$

Und hier gilt die Gleichheit genau dann, wenn $\mathbf{v}$ in die Richtung des Gradienten $\nabla f(\mathbf{x})$ zeigt. Man drückt dies auch so aus: Der Gradient $\nabla f(\mathbf{x})$ weist in die Richtung des stärksten Anstiegs der Funktion f.

Satz 2.4 (MWS der Differentialrechnung) *Die Funktion f sei stetig differenzierbar auf der offenen Menge $U \subset \mathbb{R}^n$. Es seien $\mathbf{a},\mathbf{b} \in U$ zwei Punkte, deren Verbindungsstrecke*

$$\{\mathbf{a}+t\cdot(\mathbf{b}-\mathbf{a}) :\ 0 \leq t \leq 1\}$$

ganz in U liegt. Dann gibt es ein θ, $0<\theta<1$, so, dass

$$f(\mathbf{b})-f(\mathbf{a}) = (\nabla f(\mathbf{x}).(\mathbf{b}-\mathbf{a}))$$

mit einem Zwischenpunkt $\mathbf{x} = \mathbf{a}+\theta\cdot(\mathbf{b}-\mathbf{a})$.

Beweis. Nach Satz 2.3 ist die Funktion

$$g(t) = f(\mathbf{a}+t\cdot(\mathbf{b}-\mathbf{a}))$$

stetig differenzierbar auf einem offenen Intervall, welches das Intervall $[0,1]$ enthält. Ihre Ableitung ist

$$g'(t) = \frac{d}{dt}f(\mathbf{a}+t\cdot(\mathbf{b}-\mathbf{a})) = \frac{d}{du}f(\mathbf{a}+(t+u)\cdot(\mathbf{b}-\mathbf{a}))\Big|_{u=0} =$$

$$= \frac{d}{du}f(\mathbf{a}+t\cdot(\mathbf{b}-\mathbf{a})+u\cdot(\mathbf{b}-\mathbf{a}))\Big|_{u=0} = (\nabla f(\mathbf{a}+t\cdot(\mathbf{b}-\mathbf{a})).(\mathbf{b}-\mathbf{a})).$$

Der MWS in einer Variablen t zeigt

$$f(\mathbf{b})-f(\mathbf{a}) = g(1)-g(0) = g'(\theta) \text{ mit } 0<\theta<1.$$

Und das ist die Behauptung. □

Definition 2.5 (konvex) *Die Menge $M \subset \mathbb{R}^n$ heißt* konvex, *wenn sie mit je zwei Punkten $\mathbf{a}, \mathbf{b} \in M$ auch die ganze Strecke zwischen diesen Punkten enthält.*

Beispiel 2.6 *Die Funktion $f(x)$ sei zweimal stetig differenzierbar mit $f'' > 0$ auf dem offenen Intervall $I =]u, v[\subset \mathbb{R}$. Dann ist die Menge oberhalb des Funktionsgraphen*

$$M := \{(x, y) \in \mathbb{R}^2 : u < x < v, y \geq f(x)\}$$

konvex.

Beweis. Es seien $\mathbf{a} = (x_1, y_1)$ und $\mathbf{b} = (x_2, y_2)$ zwei Punkte aus M und

$$\mathbf{x}(t) = (x(t), y(t)) = (x_1 + t \cdot (x_2 - x_1), y_1 + t \cdot (y_2 - y_1)), 0 \leq t \leq 1,$$

ein Punkt auf ihrer Verbindungsstrecke. Es ist zu zeigen $y(t) \geq f(x(t))$.

Falls $x_1 = x_2 =: x$ ist, dann ist

$$y_1 \geq f(x), \quad y_2 \geq f(x),$$

und

$$y(t) = y_1 + t(y_2 - y_1) = (1-t)y_1 + ty_2 \geq (1-t)f(x) + tf(x) = f(x),$$

also $\mathbf{x}(t) \in M$.

Falls $x_1 \neq x_2$ ist, betrachten wir die zweimal stetig differenzierbare Funktion

$$g(t) := y(t) - f(x(t)) = y_1 + t(y_2 - y_1) - f(x_1 + t(x_2 - x_1))$$

mit

$$g(0) = y_1 - f(x_1) \geq 0, \quad g(1) = y_2 - f(x_2) \geq 0$$

und

$$g''(t) = \frac{d}{dt}\left(y_2 - y_1 - f'(x_1 + t(x_2 - x_1)) \cdot (x_2 - x_1)\right) = -f''(x_1 + t(x_2 - x_1)) \cdot (x_2 - x_1)^2 < 0.$$

Wenn $g(t) \geq 0$ für alle t, dann gehören die Punkte $\mathbf{x}(t)$ zu M. Andernfalls existiert ein t, $0 < t < 1$ mit $g(t) < 0$. Dann nimmt die Funktion $g(t)$ im Inneren des Intervalls $]0, 1[$ ein Minimum an. Dort würde aber gelten

$$g'(t) = 0 \text{ und } g''(t) < 0.$$

Es kann sich nur um ein lokales Maximum handeln, Widerspruch! $\square$

Satz 2.5 *Die Menge $U \subset \mathbb{R}^n$ sei offen und konvex. Die Funktion $f : U \to \mathbb{R}$ sei stetig partiell differenzierbar mit*

$$\| \nabla f(\mathbf{x}) \| \leq L$$

für alle $\mathbf{x} \in U$. Dann ist f auf U Lipschitz-stetig mit Lipschitz-Konstante L.

Beweis. Es seien $\mathbf{x}, \mathbf{y} \in U$. Weil U konvex vorausgesetzt ist, gehört die ganze Verbindungsstrecke zwischen $\mathbf{x}$ und $\mathbf{y}$ zu U. Nach dem MWS gibt es ein $\theta, 0 < \theta < 1$, mit

$$|f(\mathbf{y}) - f(\mathbf{x})| = |(\nabla f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})).\mathbf{y} - \mathbf{x})| \leq \| \nabla f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x})) \| \cdot \| \mathbf{y} - \mathbf{x} \| \leq L \cdot \| \mathbf{y} - \mathbf{x} \| . \quad \square$$

Aufgabe 2.1 *Für welche* $(x, y) \in \mathbb{R}^2$ *ist*

$$\begin{array}{lcc} f : \mathbb{R}^2 & \to \mathbb{R} & \\ (x, y) & \mapsto & |x \cdot y| \end{array}$$

partiell nach x *differenzierbar?*

Aufgabe 2.2 *Gegeben sei die Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$ *durch*

$$f(x, y) := \begin{cases} sin(x) \cdot arctang\left(ln\left(\frac{1}{\sqrt{x^2+y^2}}\right)\right) & falls \quad (x, y) \neq (0, 0) \\ 0 & falls \quad (x, y) = (0, 0) \end{cases}$$

Zeigen Sie, dass f *im Punkt* $(0, 0)$ *partiell nach* x *differenzierbar ist und bestimmen Sie dort diese erste partielle Ableitung nach* x.

Aufgabe 2.3 *Bestimmen Sie die Tangentialebene im Punkt* $(1, -2, 2)$ *der Fläche im* $\mathbb{R}^3$, *die durch die Gleichung*

$$z = 3x^2y + 2xy^2$$

gegeben ist.

Aufgabe 2.4 *Bestimmen Sie die Tangentialebene an den Graphen der Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$ *mit* $f(x, y) = x^3 + y^3 - 3x - 12y$ *im Punkt* $(2/0/2)$.

Aufgabe 2.5 *Sei* $f(x, y) = 1 + x^4 + ay^4$ *mit* $a \in \mathbb{R}$. *Man diskutiere in Abhängigkeit von* a, *ob* f *in* $(0, 0)$ *ein lokales Extremum hat.*

Aufgabe 2.6 *Gegeben sei die Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) := x^2 + 2xy - 4x - 4y$ *und das Quadrat* $Q = [0, 2] \times [0, 2]$.
a) Man begründe, dass es $a, b \in Q$ *gibt mit* $f(a) = inf(f(I))$ *und* $f(b) = sup(f(I))$.
b) Man zeige $f(Q) = [-8, 0]$.

Aufgabe 2.7 *Gegeben sei die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x, y) := sin(x) \cdot cos(y).$$

Bestimmen Sie alle globalen Extrema von f *auf dem Quadrat*

$$\{(x, y) \in \mathbb{R}^2 : \ 0 \leq x \leq \pi, 0 \leq y \leq \pi\}.$$

Aufgabe 2.8 *Gegeben sei die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x,y) := e^{-(x^2+y^2)} \cdot (x^2 - 2y^2).$$

a) Bestimmen Sie $inf\, f(\mathbb{R}^2)$ *und* $sup\, f(\mathbb{R}^2)$.
b) Untersuchen Sie f *auf lokale Extrema.*

Aufgabe 2.9 *Zeigen Sie, dass die Funktion*

$$f : \mathbb{R}\times]0,\infty[\to \mathbb{R}, \quad f(x,y) = y^x,$$

kein lokales Extremum besitzt.

Aufgabe 2.10 *Es seien* $a = (a_1, a_2), b = (b_1, b_2)$ *und* $c = (c_1, c_2)$ *drei verschiedene Punkte der Ebene* $\mathbb{R}^2$. *Bestimmen Sie alle Punkte* $p \in \mathbb{R}^2$ *mit der Eigenschaft, dass*

$$\| p - a \|^2 + 2 \| p - b \|^2 + 3 \| p - c \|^2$$

minimal ist.

Aufgabe 2.11 *Bestimmen Sie auf der abgeschlossenen Kreisscheibe*

$$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

alle lokalen und globalen Extrema der Funktion
a) $f_1(x,y) := 3x^2 - 2xy + 3y^2$.
b) $f_2(x,y) := e^{x^2+y^2} \cdot (x-2y)^2$.
c) $f_3(x,y) := x^3 - 3xy^2$.

Aufgabe 2.12 *a) Bestimmen Sie drei positive reelle Zahlen, deren Summe gleich 1 und deren Produkt maximal ist.*
b) Bei der Post dürfen nur Pakete verschickt werden, bei denen die Summe aus Länge und Umfang (Umfang = 2×Breite+2× Höhe) nicht größer als d cm ist. Welches Maß für Länge, Breite und Höhe muss man wählen, damit das Paket größtmögliches Volumen hat und verschickt werden kann?

Aufgabe 2.13 *Zeigen Sie die Konvexität der Menge* M *in Beispiel 2.6 unter der schwächeren Voraussetzung* $f''(x) \geq 0$ *für alle* $u < x < v$.

Aufgabe 2.14 *Es sei* A *eine symmetrische* $n \times n$*-Matrix und* $q(\mathbf{x})$ *die quadratische Form* $\mathbf{x}^t \cdot A \cdot \mathbf{x}$. *Zeigen Sie*

$$\nabla q(\mathbf{x}) = 2 \cdot A \cdot \mathbf{x}.$$

2.2 Höhere partielle Ableitungen, Taylor-Formel

Wenn alle partiellen Ableitungen $\partial f/\partial x_\nu$ einer Funktion f wieder stetig sind, dann nennt man die Funktion *stetig partiell differenzierbar*. Und dann können diese partiellen Ableitungen auch wieder partiell differenzierbar sein. Die partiellen Ableitungen der partiellen Ableitungen heißen dann zweite partielle Ableitungen. Für

$$\frac{\partial}{\partial x_\mu}\left(\frac{\partial f}{\partial x_\nu}\right)$$

schreibt man etwas kürzer

$$\frac{\partial^2 f}{\partial x_\mu \partial x_\nu}.$$

Und für

$$\frac{\partial}{\partial x_\nu}\left(\frac{\partial f}{\partial x_\nu}\right)$$

schreibt man

$$\frac{\partial^2 f}{\partial x_\nu^2}.$$

Manchmal kürzt man noch weiter ab:

$$\frac{\partial^2 f}{\partial x_\mu \partial x_\nu} = \partial_\mu \partial_\nu f = f_{\mu,\nu}, \quad \frac{\partial^2 f}{\partial x_\nu^2} = \partial_\nu^2 f = f_{\nu,\nu}.$$

Eine Funktion $f(x,y)$ kann also die folgenden vier zweiten partiellen Ableitungen haben:

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}.$$

Die zweiten partiellen Ableitungen einer Funktion $f(x_1, ..., x_n)$ kann man zu einer $n \times n$-Matrix

$$H_f := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

anordnen. Diese Matrix heißt *Hesse-Matrix* der Funktion f.

Auch diese zweiten partiellen Ableitungen können wieder alle stetig sein. Dann heißt die Funktion *zweimal stetig differenzierbar*.

Satz 2.6 (Symmetrie der zweiten Ableitungen) *Die Funktion f sei zweimal <u>stetig</u> partiell differenzierbar auf der offenen Menge $U \subset \mathbb{R}^n$. Dann ist für $1 \leq i < j \leq n$*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Beweis. Wenn wir die unbeteiligten Koordinaten weglassen, können wir uns auf den Fall $n = 2$ und $x_i = x_1, x_j = x_2$ beschränken. Wir fixieren $\mathbf{x} \in U$ und wählen $\epsilon > 0$ so klein, dass das Quadrat Q mit den Eckpunkten

$$\mathbf{x},\ \mathbf{x} + \epsilon\mathbf{e}_1,\ \mathbf{x} + \epsilon\mathbf{e}_2,\ \mathbf{x} + \epsilon(\mathbf{e}_1 + \mathbf{e}_2)$$

ganz in U enthalten ist. Diesem Quadrat Q ordnen wir die Zahl

$$\Box f := f(\mathbf{x}) - f(\mathbf{x} + \epsilon\mathbf{e}_1) - f(\mathbf{x} + \epsilon\mathbf{e}_2) + f(\mathbf{x} + \epsilon(\mathbf{e}_1 + \mathbf{e}_2))$$

zu. Wir behaupten: In Q gibt es einen Punkt $\mathbf{y}$ mit

$$\Box f = \epsilon^2 \cdot \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{y}).$$

Um diese Behauptung zu beweisen schreiben wir

$$\Box f = g(x_2 + \epsilon) - g(x_2) \text{ mit } g(z) := f(x_1 + \epsilon, z) - f(x_1, z).$$

Nach dem eindimensionalen MWS gibt es ein $y_2 \in]x_2, x_2 + \epsilon[$ mit

$$g(x_2 + \epsilon) - g(x_2) = \epsilon \cdot g'(y_2).$$

Mit $h(z) := \partial f/\partial x_2(z, y_2)$ erhalten wir

$$\Box f = \epsilon \cdot g'(y_2) = \epsilon \cdot \left(\frac{\partial f}{\partial x_2}(x_1 + \epsilon, y_2) - \frac{\partial f}{\partial x_2}(x_1, y_2) \right) = \epsilon \cdot (h(x_1 + \epsilon) - h(x_1)).$$

Jetzt wenden wir den eindimensionalen MWS auf h an und erhalten ein $y_1 \in]x_1, x_1 + \epsilon[$ mit

$$h(x_1 + \epsilon) - h(x_1)) = \epsilon \cdot h'(y_1).$$

Dies bedeutet

$$\Box f = \epsilon^2 \cdot h'(y_1) = \epsilon^2 \cdot \frac{\partial^2 f}{\partial x_1 \partial x_2}(y_1, y_2).$$

Vertauschen wir die Reihenfolge von x_1 und x_2, so erhalten wir analog ein $\mathbf{w}$ in Q mit

$$\Box f = \epsilon^2 \cdot \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{w}).$$

Nach Division durch ϵ^2 finden wir

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{w}) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{y}).$$

Jetzt lassen wir ϵ gegen 0 gehen. Dabei zieht sich das Quadrat Q auf den Punkt $\mathbf{x}$ zusammen. Beide Vektoren $\mathbf{y}$ und $\mathbf{w}$ konvergieren gegen $\mathbf{x}$. Mit der Stetigkeit der zweiten Ableitungen folgt daraus

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}).$$

□

Satz 2.6 kann man so ausdrücken: Es kommt bei der Bildung der zweiten partiellen Ableitungen nicht auf die Reihenfolge an, in der man nach den Veränderlichen x_i differenziert.

Beispiel 2.7 *Wir betrachten die Funktion*

$$f(x,y) = x^y.$$

Ihre ersten partiellen Ableitungen haben wir schon einmal ausgerechnet:

$$\frac{\partial f}{\partial x} = y \cdot x^{y-1}, \qquad \frac{\partial f}{\partial y} = x^y \cdot ln(x).$$

Daraus erhalten wir dann die folgenden zweiten partiellen Ableitungen

$$\begin{array}{ll} \dfrac{\partial^2 f}{\partial x^2} = y(y-1) \cdot x^{y-2} & \dfrac{\partial^2 f}{\partial x \partial y} = y \cdot x^{y-1} \cdot ln(x) + x^y \cdot \dfrac{1}{x} \\ \dfrac{\partial^2 f}{\partial y \partial x} = x^{y-1} + y \cdot x^{y-1} \cdot ln(x) & \dfrac{\partial^2 f}{\partial y^2} = x^y \cdot ln(x)^2. \end{array}$$

Die beiden gemischten zweiten partiellen Ableitungen sind untereinander gleich, und zwar

$$= x^{y-1} \cdot (y \cdot ln(x) + 1),$$

ganz, wie Satz 2.6 das behauptet.

Wenn die zweiten partiellen Ableitungen wieder partiell differenzierbar sind, kann man nochmal partiell differenzieren, und erhält so die dritten partiellen Ableitungen

$$\frac{\partial^3 f}{\partial x_1 \partial x_2 \partial x_3}(\mathbf{x}).$$

Weil hier die zweiten partiellen Ableitungen stetig sind, kann man nach Satz 2.6 die Reihenfolge der Ableitungen nach x_2 und x_3 vertauschen. Und falls die dritten partiellen Ableitunge wieder stetig sind, kann man auch die Reihenfolge der Ableitungen nach x_1 und x_2 vertauschen. Durch Hintereinander-Ausführung solcher Vertauschungen kann man alle Permutationen von drei partiellen Ableitungen erhalten. Wir haben gesehen: Ist f dreimal stetig partiell differenzierbar, so kommt es bei der Bildung der dritten partiellen Ableitungen nicht auf die Reihenfolge an. Durch Induktion nach der Anzahl der partiellen Ableitungen zeigt man:

Satz 2.7 *Die Funktion f sei k-mal stetig partiell differenzierbar. Dann kann man bei der Bildung der k-ten partiellen Ableitungen die Reihenfolge, in der man nach den einzelnen Variablen differenziert, vertauschen ohne das Ergebnis zu ändern.*

Und wozu braucht man diese höheren partiellen Ableitungen? Antwort: für die Taylor-Reihe der Funktion f. Das kann ja lustig werden, bei so viel partiellen Ableitungen!

Die Taylor-Reihe in einer Veränderlichen

$$f(x+h) = \sum_{\mu=0}^{\infty} \frac{1}{\mu!} f^{(\mu)}(x) \cdot h^\mu$$

berechnet (wenn alles glatt geht) den Funktionswert im ausgelenkten Punkt $x+h$ durch die Ableitungen im Aufpunkt x. Bei einer Funktion von mehreren Veränderlichen ist der Aufpunkt $\mathbf{x}$, der ausgelenkte

Punkt $\mathbf{x}+\mathbf{h}$, und wir wollen $f(\mathbf{x}+\mathbf{h})$ berechnen. Dazu brauchen wir aber nicht die Funktion f in Abhängigkeit von allen n Veränderlichen, sondern nur die Funktion $f(\mathbf{x}+t\cdot\mathbf{h})$ der einen Veränderlichen t. Die Taylorformel ist dann (wir setzen $t=1$)

$$f(\mathbf{x}+\mathbf{h}) = \sum_{\mu=1}^{m} \frac{1}{\mu!} \frac{d^\mu}{dt^\mu} f(\mathbf{x}+t\cdot\mathbf{h})\Big|_{t=0} + R_m$$

mit dem Restglied

$$R_m := \frac{1}{(m+1)!} \frac{d^{m+1}}{dt^{m+1}} f(\mathbf{x}+t\cdot\mathbf{h})\Bigg|_{t=\theta} \quad \text{mit} \quad 0<\theta<1.$$

Das ist nichts neues, genau die eindimensionale Formel. Das Problem besteht darin, die höheren Ableitungen nach t in den partiellen Ableitungen der Funktion f auszudrücken. Für $\mu=1$ ist das die Richtungsableitung

$$\frac{d}{dt} f(\mathbf{x}+t\cdot\mathbf{h})\Big|_{t=0} = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu.$$

Für $m=2$ müssen wir die Funktion von t

$$\frac{d}{dt} f(\mathbf{x}+t\cdot\mathbf{h}) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}+t\cdot\mathbf{h}) \cdot h_\nu$$

nochmal nach t ableiten. Das geht mit den Richtungsableitungen der ersten partiellen Ableitungen $\partial f/\partial x_\nu$ und liefert

$$\frac{d^2 f}{dt^2}(\mathbf{x}+t\cdot\mathbf{h})\Big|_{t=0} = \sum_{\nu_2,\nu_1=1}^{n} \frac{\partial^2 f}{\partial x_{\nu_2}\partial x_{\nu_1}}(\mathbf{x}) \cdot h_{\nu_1} h_{\nu_2}.$$

Durch Iteration erhält man die μ-te Ableitung

$$\frac{d^\mu f}{dt^\mu}(\mathbf{x}+t\cdot\mathbf{h})\Big|_{t=0} = \sum_{\nu_\mu,\ldots,\nu_1=1}^{n} \frac{\partial^\mu f}{\partial x_{\nu_\mu}\ldots\partial x_{\nu_1}}(\mathbf{x}) \cdot h_{\nu_1}\cdot\ldots\cdot h_{\nu_\mu}.$$

Der Schreibaufwand ist beträchtlich. Man kann ihn durch trickreiche Notation etwas vereinfachen, s. Knauf p. 58. Ich möchte mich aber hier nicht damit befassen, sondern die Taylorfomel in der hergeleiteten Form hinschreiben.

Satz 2.8 (Taylor-Formel) *Die Funktion $f(x_1,\ldots,x_n)$ sei $m+1$-mal stetig partiell differenzierbar in einer Kugel um $\mathbf{x}\in\mathbb{R}^n$ vom Radius r. Für alle $\mathbf{h}\in\mathbb{R}^n$ mit $\|\mathbf{h}\|<r$ gilt dann*

$$f(\mathbf{x}+\mathbf{h}) = \sum_{\mu=1}^{m} \frac{1}{\mu!} \sum_{\nu_\mu,\ldots,\nu_1=1}^{n} \frac{\partial^\mu f}{\partial x_{\nu_\mu}\ldots\partial x_{\nu_1}}(\mathbf{x}) \cdot h_{\nu_1}\cdot\ldots\cdot h_{\nu_\mu} + R_m$$

mit einem Restglied

$$R_m = \frac{1}{(m+1)!} \sum_{\nu_{m+1},\ldots,\nu_1=1}^{n} \frac{\partial^{m+1} f}{\partial x_{\nu_{m+1}}\ldots\partial x_{\nu_1}}(\mathbf{x}+\theta\cdot\mathbf{h}) \cdot h_{\nu_1}\cdot\ldots\cdot h_{\nu_{m+1}}, \quad 0<\theta<1.$$

Und dabei haben wir uns noch gar nicht um die üblichen Komplikationen gekümmert, nämlich:

- Die Taylor-Reihe braucht nicht zu konvergieren,
- und wenn sie konvergiert, dann braucht sie nicht gegen den Wert $f(\mathbf{x}+\mathbf{h})$ zu konvergieren.

Ob diese Taylor-Reihe gegen die Funktion konvergiert, hängt - wie bei Funktionen einer Variablen auch - vom Restglied R_m ab. Das Restglied ist jetzt eine Funktion von n Variablen: $R_m(h_1,...,h_n)$. Schauen wir uns die ersten Fälle einmal an:

<u>$m=1$</u>: Die Taylorformel wird

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) \cdot h_\nu + R_1(\mathbf{h})$$

mit

$$R_1(\mathbf{h}) = \frac{1}{2} \sum_{\nu_1,\nu_2=1}^{n} \frac{\partial^2 f}{\partial_{\nu_1}\partial_{\nu_2}}(\mathbf{x}+\theta\cdot\mathbf{h}) \cdot h_{\nu_1} h_{\nu_2}.$$

Für $\mathbf{h} \to \mathbf{0}$ gehen hier die zweiten Ableitungen gegen dieselben Ableitungen in $\mathbf{x}$, wegen der Stetigkeit dieser Ableitungen. Und wegen

$$\frac{h_{\nu_1} h_{\nu_2}}{\parallel \mathbf{h} \parallel} = \frac{h_1}{\parallel \mathbf{h} \parallel} \cdot h_2 \leq h_2$$

ist

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_1(\mathbf{h})}{\parallel \mathbf{h} \parallel} = 0.$$

Das ist die lineare Approximation der Funktion f durch die Tangentialebene an ihren Graphen.

<u>m=2</u>: Wir bekommen

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + (\nabla f(\mathbf{x}).\mathbf{h}) + \frac{1}{2} \cdot \mathbf{h}^t \cdot H_f(\mathbf{x}) \cdot \mathbf{h} + R_2(\mathbf{h})$$

mit

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_2(\mathbf{h})}{\parallel \mathbf{h} \parallel^2} = 0.$$

In der Taylor-Entwicklung von f bis zur Ordnung m haben wir die Funktion $m+1$-mal stetig partiell differenzierbar vorausgesetzt. Aber die $m+1$-ten partiellen Ableitungen haben wir nur gebraucht, um die Form des Restglieds R_m explizit anzugeben. Manchmal braucht man diese explizite Form nicht, sondern nur

Satz 2.9 (Abschätzung für das Restglied) *Es seien die Voraussetzungen von Satz 2.8 erfüllt. Allerdings braucht f jetzt nur m-mal stetig partiell differenzierbar zu sein. Auch dann gilt schon*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_m(\mathbf{h})}{\parallel \mathbf{h} \parallel^m} = 0.$$

Beweis. Wir verwenden die Taylorformel bis zur Ordnung $m-1$

$$\begin{aligned} f(\mathbf{x}+\mathbf{h}) &= \sum_{\mu=0}^{m-1} \frac{1}{\mu!} \sum_{\nu_\mu,...,\nu_1=1}^{n} \frac{\partial^\mu f}{\partial_{\nu_\mu}...\partial_{\nu_1}}(\mathbf{x}) \cdot h_{\nu_\mu} \cdot ... \cdot h_{\nu_1} + R_{m-1}(\mathbf{h}) \\ &= \sum_{\mu=0}^{m} \frac{1}{\mu!} \sum_{\nu_\mu,...,\nu_1=1}^{n} \frac{\partial^\mu f}{\partial_{\nu_\mu}...\partial_{\nu_1}}(\mathbf{x}) \cdot h_{\nu_\mu} \cdot ... \cdot h_{\nu_1} + R_m(\mathbf{h}) \end{aligned}$$

mit

$$R_m(\mathbf{h}) = \frac{1}{m!} \sum_{\nu_m,\ldots,\nu_1=1}^{n} \left(\frac{\partial^m f}{\partial_{\nu_m}\ldots\partial_{\nu_1}}(\mathbf{x}+\theta\mathbf{h}) - \frac{\partial^m f}{\partial_{\nu_m}\ldots\partial_{\nu_1}}(\mathbf{x}) \right) \cdot h_{\nu_1} \cdot \ldots \cdot h_m.$$

Weil die partiellel Ableitungen bis zur Ordnung m stetig vorausgesetzt sind, geht für $\mathbf{h} \to \mathbf{0}$ der Ausdruck in der Klammer gegen 0. Außerdem bleibt

$$q(\mathbf{h}) := \frac{1}{\| \mathbf{h} \|^m} \cdot h_1 \cdot \ldots \cdot h_m = \frac{h_1}{\| \mathbf{h} \|} \cdot \ldots \cdot \frac{h_m}{\| \mathbf{h} \|} \leq 1$$

beschränkt. $\square$

Aufgabe 2.15 *Es seien $h : \mathbb{R} \to \mathbb{R}$ und $k : \mathbb{R} \to \mathbb{R}$ zweimal stetig differenzierbare Funktionen. Die Funktion $F : \mathbb{R}^2 \to \mathbb{R}$ sei durch $F(x,y) := h(x + k(y))$ definiert. Berechnen Sie die partiellen Ableitungen zweiter Ordnung von F.*

Aufgabe 2.16 *Warum gibt es keine partiell differenzierbare Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit*

$$(\nabla f)(x,y) = (arctan(xy), e^x sin\ y)$$

für jedes $(x,y) \in \mathbb{R}^2$?

Aufgabe 2.17 *Mit Hilfe der Taylor-Formel (Entwicklung bei $(x_0, y_0) = (1,0)$) beweise man*

$$\frac{1}{x} e^y \geq 2 - x + y$$

für $0 < x \leq 1$ und $y \geq 0$.

Aufgabe 2.18 *Zeigen Sie, dass die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R} \text{ mit } f(0,0) = 0 \text{ und } f(x,y) = \frac{x^3y - xy^3}{x^2 + y^2} \text{ für } (x,y) \neq (0,0)$$

im Punkt $(0,0)$ zweimal partiell differenzierbar ist, und die Hesse-Matrix von f im Punkt $(0,0)$ nicht symmetrisch ist. Ist f zweimal stetig partiell differenzierbar?

Aufgabe 2.19 *Gegeben sei die Funktion $f :]0,\infty[\times]0,\infty[\to \mathbb{R}$ mit $f(x,y) = xy + \frac{a}{x} + \frac{a}{y}$, $a > 0$. Bestimmen Sie das Taylor-Polynom zweiten Grades von f im Entwicklungspunkt (a,a).*

Aufgabe 2.20 *Bestimmen Sie mit Hilfe der Taylor-Formel alle zweimal stetig partiell differenzierbaren Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ mit konstanten zweiten partiellen Ableitungen.*

Aufgabe 2.21 *Es seien F und G zweimal stetig partiell differenzierbare Funktionen von $\mathbb{R}$ in $\mathbb{R}$, und $u : \mathbb{R}^2 \to \mathbb{R}$ sei durch $u(t,x) = F(x+t) + G(x-t)$ definiert. Man beweise*

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}.$$

Aufgabe 2.22 *Es sei $f :]0,\infty[\times]0,\infty[\to \mathbb{R}$ die durch $f(x,y) = x^y$ für $x > 0$ und $y > 0$ gegebene reellwertige Funktion. Man berechne das Taylor-Polynom 2. Ordnung von f im Entwicklungspunkt $(1,1)$.*

Aufgabe 2.23 *Bestimmen Sie die Taylor-Reihe der Funktion*
a) $r(x,y) := \sqrt{x^2+y^2}$ um $(x,y) = (1,0)$
b) $f(x,y) := \frac{2xy}{x+y}$ um $(x,y) = (1,1)$
bis einschließlich der Summanden zweiter Ordnung.

2.3 Ableitung von Abbildungen

Jetzt verallgmeinern wir die Differentiation von Funktionen $f(x_1, ..., x_m)$ mehrerer Veränderlicher auf Abbildungen. Wie in 1.1 schreiben wir eine Abbildung

$$F : U \to \mathbb{R}^n, \quad U \subset \mathbb{R}^m, \quad F(x_1, ..., x_m) = \begin{pmatrix} F_1(x_1, ..., x_m) \\ F_2(x_1, ..., x_m) \\ \vdots \\ F_n(x_1, ..., x_m) \end{pmatrix}.$$

Die Differenzierbarkeit der Abbildung F in einem Punkt $\mathbf{x}$ der offenen Menge $U \subset \mathbb{R}^m$ definieren wir durch lineare Approximierbarkeit. Allerdings wird die lineare Abbildung $F'(\mathbf{x})$, mit der wir F in der Nähe von $\mathbf{x}$ approximieren, eine lineare Abbildung

$$\mathbb{R}^m \ni \mathbf{h} \mapsto F'(\mathbf{x})(\mathbf{h}) \in \mathbb{R}^n.$$

Hier wollen wir die lineare Abbildung $F'(\mathbf{x})$ mit ihrer darstellenden Matrix identifizieren, das ist also eine $n \times m$-Matrix, und schreiben

$$F'(\mathbf{x})(\mathbf{h}) = F'(\mathbf{x}) \cdot \mathbf{h}.$$

Definition 2.6 *Es sei $U \subset \mathbb{R}^m$ offen und $\mathbf{x} \in U$. Eine Abbildung $F : U \to \mathbb{R}^n$ heißt (total) differenzierbar in $\mathbf{x}$, wenn es eine $n \times m$-Matrix A gibt, so, dass für alle $\mathbf{x} + \mathbf{h} \in U$ gilt*

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + A \cdot \mathbf{h} + \Phi(\mathbf{h})$$

mit

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{1}{\| \mathbf{h} \|} \cdot \Phi(\mathbf{h}) = \mathbf{0}.$$

Die Matrix A soll natürlich die Ableitung $F'(\mathbf{x})$ sein. Aber bevor wir das definieren können, müssen wir uns überlegen, dass diese Matrix eindeutig bestimmt ist.

Satz 2.10 *Es seien A, B zwei $n \times m$-Matrizen so, dass die Bedingung aus Definition 2.6 für beide Matrizen gilt. Dann ist $A = B$.*

Beweis. Wir haben also jetzt

$$F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) = A \cdot \mathbf{h} + \Phi(\mathbf{h}) = B \cdot \mathbf{h} + \Psi(\mathbf{h})$$

mit

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{1}{\| \mathbf{h} \|} \cdot \Phi(\mathbf{h}) = \lim_{\mathbf{h} \to \mathbf{0}} \frac{1}{\| \mathbf{h} \|} \cdot \Psi(\mathbf{h}) = \mathbf{0}.$$

Wenn wir die Bedingungen für A und B voneinander abziehen, finden wir

$$\lim_{\mathbf{h} \to \mathbf{0}} (A - B) \cdot \frac{\mathbf{h}}{\| \mathbf{h} \|} = \mathbf{0}.$$

Weil U offen ist, gibt es eine abgeschlossenen Kugel vom Radius $r > 0$ um $\mathbf{x}$, die ganz in U enthalten ist. Für jeden Vektor $\mathbf{h}$ mit $\| \mathbf{h} \| = r$ und alle $0 < t \leq 1$ gilt also

$$(A - B) \cdot \mathbf{h} = (A - B) \cdot \frac{rt \cdot \mathbf{h}}{rt} = r \cdot (A - B) \cdot \frac{t \cdot \mathbf{h}}{\| t \cdot \mathbf{h} \|} = r \cdot \lim_{t \to 0} (A - B) \cdot \frac{t \cdot \mathbf{h}}{\| t \cdot \mathbf{h} \|} = \mathbf{0}.$$

Jeder Vektor $\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^m$ schreibt sich

$$\mathbf{x} = \frac{\| \mathbf{x} \|}{r} \cdot \mathbf{h} \quad \text{mit} \quad \mathbf{h} = \frac{r}{\| \mathbf{x} \|} \cdot \mathbf{x} \text{ und } \| \mathbf{h} \| = r.$$

Daraus folgt $(A - B) \cdot \mathbf{x} = \mathbf{0}$ für alle $\mathbf{x} \in \mathbb{R}^m$ und $A - B = 0$. □

Definition 2.7 *Die durch Satz 2.10 eindeutig bestimmte Matrix in Definition 2.6 heißt die* totale Ableitung $F'(\mathbf{x})$ *der Abbildung F im Punkt $\mathbf{x}$.*

Diese Definition sieht imposant aus. Wir lassen etwas Luft raus:

Satz 2.11 *a) Eine Abbildung F ist total differenzierbar im Punkt $\mathbf{x}$, wenn alle ihre Komponentenfunktionen $F_1, ..., F_n$ es sind.*
b) Die Ableitung $F'(\mathbf{x})$ ist dann die Funktionalmatrix *oder* Jacobi-Matrix

$$\left(\frac{\partial F_\nu}{\partial x_\mu}(\mathbf{x}) \right) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_1}{\partial x_m}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_n}{\partial x_m}(\mathbf{x}) \end{pmatrix}.$$

Beweis. Wir schreiben

$$F(\mathbf{x}+\mathbf{h}) = F(\mathbf{x}) + F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{x})$$

als Spaltenvektor

$$\begin{pmatrix} F_1(\mathbf{x}+\mathbf{h}) \\ \vdots \\ F_n(\mathbf{x}+\mathbf{h}) \end{pmatrix} = \begin{pmatrix} F_1(\mathbf{x}) \\ \vdots \\ F_n(\mathbf{x}) \end{pmatrix} + \begin{pmatrix} F_1'(\mathbf{x}) \cdot \mathbf{h} \\ \vdots \\ F_n'(\mathbf{x}) \cdot \mathbf{h} \end{pmatrix} + \begin{pmatrix} \Phi_1(\mathbf{h}) \\ \vdots \\ \Phi_n(\mathbf{h}) \end{pmatrix}.$$

Hier sind $F_1', ..., F_n'$ die Zeilen der Matrix F' und $\Phi_1, ..., \Phi_n$ die Komponentenfunktionen der Abbildung Φ. Nach Satz 1.1 ist

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\Phi(\mathbf{h})}{\parallel \mathbf{h} \parallel} = \mathbf{0}$$

äquivalent zu

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\Phi_\nu(\mathbf{h})}{\parallel \mathbf{h} \parallel} = 0$$

für $\nu = 1, ..., n$. Deswegen ist die totale Differenzierbarkeit der Abbildung F äquivalent zur totalen Differenzierbarkeit aller ihrer Komponentenfunktionen F_ν. Und in den Zeilen der Matrix $F'(\mathbf{x})$ stehen genau die Gradienten $\nabla F_\nu(\mathbf{x})$ dieser Komponentenfunktionen. $\square$

Satz 2.12 (Korollar) *Sind die Komponentenfunktionen F_ν der Abbildung F stetig partiell differenzierbar in $\mathbf{x}$, so ist die Abbildung F dort total differenzierbar.*

Ist F stetig partiell differenzierbar, so ist F total differenzierbar, und die Einträge $\partial F_\nu / \partial x_\mu$ der Funktionalmatrix $F'(\mathbf{x})$ sind stetig. Also ist die Abbildung $\mathbf{x} \mapsto F'(\mathbf{x})$ stetig. Und umgekehrt, ist diese Matrix $F'(\mathbf{x})$ stetig in $\mathbf{x}$, so ist F stetig partiell differenzierbar. Solche Abbildungen nennen wir künftig *stetig differenzierbar*.

Als Beispiele berechnen wir die Funktionalmatrizen der Transformation in Polar- und Kugelkoordinaten.

Beispiel 2.8 (Polarkoordinaten) *Die Transformation in Polarkoordinaten ist*

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cdot cos(\varphi) \\ r \cdot sin(\varphi) \end{pmatrix}.$$

Die Funktionalmatrix dieser Abbildung ist

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} cos(\varphi) & -r \cdot sin(\varphi) \\ sin(\varphi) & r \cdot cos(\varphi) \end{pmatrix}.$$

Beispiel 2.9 (Kugelkoordinaten) *Die Transformation in Kugelkoordinaten ist*

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \cdot cos(\varphi) \cdot cos(\theta) \\ r \cdot sin(\varphi) \cdot cos(\theta) \\ r \cdot sin(\theta) \end{pmatrix}$$

mit der Funktionalmatrix

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \theta} \end{pmatrix} = \begin{pmatrix} cos(\varphi) \cdot cos(\theta) & -r \cdot sin(\varphi) \cdot cos(\theta) & -r \cdot cos(\varphi) \cdot sin(\theta) \\ sin(\varphi) \cdot cos(\theta) & r \cdot cos(\varphi) \cdot cos(\theta) & -r \cdot sin(\varphi) \cdot sin(\theta) \\ sin(\theta) & 0 & t \cdot cos(\theta) \end{pmatrix}.$$

Beispiel 2.10 *Ein ziemlich formales Beispiel ist die Ableitung der Abbildung* $id : \mathbb{R}^n \to \mathbb{R}^n$*. Sie hat die Komponentenschreibweise*

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

Damit wird ihre Funktionalmatrix

$$id'(\mathbf{x}) = \left(\frac{\partial x_\mu}{\partial x_\nu}\right)_{\mu,\nu=1,\ldots,n} = (\delta_{\mu,\nu})_{\mu,\nu=1,\ldots,n} = \mathbb{1}_n,$$

die Einheitsmatrix unabhängig von $\mathbf{x}$*.*

Beispiel 2.11 (Kurve) *Eine stetig differenzierbare Kurve im* $\mathbb{R}^n$ *ist eine stetig differenzierbare Abbildung eines Intervalls* $]a,b[\subset \mathbb{R}$ *in den* $\mathbb{R}^n$*. Sie schreibt sich*

$$\gamma :]a,b[\to \mathbb{R}^n, \quad \gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{pmatrix}$$

mit stetig differenzierbaren Funktionen $\gamma_1, ..., \gamma_n$*. Ihre Ableitung ist jetzt die* $n \times 1$*-Matrix, d.h., der Vektor*

$$\gamma'(t) = \begin{pmatrix} \gamma_1'(t) \\ \vdots \\ \gamma_n'(t) \end{pmatrix}.$$

Dieser Vektor heißt der Geschwindigkeitsvektor *der Kurve. Für den Kreis vom Radius* r *haben wir*

$$\gamma(t) = r \cdot \begin{pmatrix} cos(t) \\ sin(t) \end{pmatrix} \quad mit \quad \gamma'(t) = r \cdot \begin{pmatrix} -sin(t) \\ cos(t) \end{pmatrix}$$

und für die Helix

$$\gamma(t) = \begin{pmatrix} r \cdot cos(t) \\ r \cdot sin(t) \\ c \cdot t \end{pmatrix} \quad mit \quad \gamma'(t) = \begin{pmatrix} -r \cdot sin(t) \\ r \cdot cos(t) \\ c \end{pmatrix}.$$

In 2.1 haben wir die aus der Differentialrechnung einer Veränderlichen bekannten Formeln auf Funktionen mehrerer Veränderlicher verallgemeinert, bis auf die Kettenregel. Das folgt jetzt.

Satz 2.13 (Kettenregel) *Es seien* $U \subset \mathbb{R}^m$ *und* $V \subset \mathbb{R}^n$ *offen,* $F : U \to V$ *und* $G : V \to \mathbb{R}^p$ *Abbildungen. Sind* F *in* $\mathbf{x} \in U$ *und* G *in* $F(\mathbf{x}) \in V$ *differenzierbar, so ist auch die zusammengesetzte Abbildung* $G \circ F : U \to \mathbb{R}^p$ *differenzierbar in* $\mathbf{x}$ *mit der Ableitung*

$$\underbrace{(G \circ F)'(\mathbf{x})}_{p\times m} = \underbrace{G'(F(\mathbf{x}))}_{p\times n} \cdot \underbrace{F'(\mathbf{x})}_{n\times m}.$$

Der Mal-Punkt bedeutet hier die Matrizenmultiplikation. Es lohnt sich, wenn man sich davon überzeugt, dass die vorkommenden Matrizen so dimensioniert sind, dass diese Multiplikation definiert ist.

Beweis des Satzes. Wir setzen $\mathbf{y} := F(\mathbf{x})$. Dann ist

$$F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) = F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{h}) \quad \text{mit} \quad \lim_{\mathbf{h}\to\mathbf{0}} \frac{\Phi(\mathbf{h})}{\| \mathbf{h} \|} = \mathbf{0}$$

und

$$G(\mathbf{y}+\mathbf{k}) - G(\mathbf{y}) = G'(\mathbf{y}) \cdot \mathbf{k} + \Psi(\mathbf{k}) \quad \text{mit} \quad \lim_{\mathbf{k}\to\mathbf{0}} \frac{\Psi(\mathbf{k})}{\| \mathbf{k} \|} = \mathbf{0}.$$

Weiter ist

$$(G \circ F)(\mathbf{x}+\mathbf{h}) - (G \circ F)(\mathbf{x}) = G(\mathbf{y} + F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x})) - G(\mathbf{y}).$$

Hier setzen wir

$$\mathbf{k} := F(\mathbf{x}+\mathbf{h}) - F(\mathbf{x}) = F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{h})$$

und haben

$$\begin{aligned}(G \circ F)(\mathbf{x}+\mathbf{h}) - (G \circ F)(\mathbf{x}) &= G'(\mathbf{y}) \cdot (F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{h})) + \Psi(\mathbf{k}) \\ &= G'(\mathbf{y}) \cdot F'(\mathbf{x}) \cdot \mathbf{h} + G'(\mathbf{y}) \cdot \Phi(\mathbf{h}) + \Psi(\mathbf{k}).\end{aligned}$$

Hier ist

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{1}{\| \mathbf{h} \|} G'(\mathbf{y}) \cdot \Phi(\mathbf{h}) = G'(\mathbf{y}) \cdot \lim_{\mathbf{h}\to\mathbf{0}} \frac{\Phi(\mathbf{h})}{\| \mathbf{h} \|} = \mathbf{0}.$$

Es bleibt zu zeigen

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\Psi(\mathbf{k})}{\| \mathbf{k} \|} = 0,$$

oder äquivalent dazu: Zu jedem $\epsilon > 0$ gibt es ein $\delta_\Psi(\epsilon) > 0$ mit

$$\| \Psi(\mathbf{k}) \| = \| \Psi(F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{h})) \leq \epsilon \cdot \| \mathbf{h} \| \quad \text{für} \quad \| \mathbf{h} \| < \delta_\Psi(\epsilon).$$

Nun sei $C := \| F'(\mathbf{x}) \|$. Nach Beispiel 1.2 ist

$$\| F'(\mathbf{x}) \cdot \mathbf{h} \| \leq C \cdot \| \mathbf{h} \| \quad \text{für alle} \quad \mathbf{h} \in \mathbb{R}^m.$$

Und es gibt ein $\delta_\Phi(C)$ mit

$$\| \Phi(\mathbf{h}) \| \leq C \cdot \| \mathbf{h} \| \quad \text{für} \quad \| \mathbf{h} \| < \delta_\Phi(C).$$

Die beiden letzten Ungleichungen zusammen zeigen für $\| \mathbf{h} \| < \delta_\Phi(C)$

$$\| \mathbf{k} \| = \| F'(\mathbf{x}) \cdot \mathbf{h} + \Phi(\mathbf{h}) \| \leq 2C \cdot \| \mathbf{h} \| .$$

Wenn jetzt auch noch

$$\| \mathbf{h} \| < \frac{1}{2C} \delta_\Psi(\epsilon)$$

gewählt wird, dann ist

$$\| \mathbf{k} \| \leq \delta_\Psi(\epsilon)$$

und

$$\| \Psi(\mathbf{k}) \| \leq \epsilon \cdot \| \mathbf{k} \| \leq \epsilon \cdot 2C \cdot \| \mathbf{h} \| .$$

Das reicht. $\square$

Beispiel 2.12 *In 2.1 berechneten wir die Richtungsableitung*

$$\frac{d}{dt} f(\mathbf{x} + t \cdot \mathbf{v}) \Big|_{t=0} = (\nabla f(\mathbf{x}).\mathbf{v})$$

mit der linearen Approximierbarkeit der Funktion F. Man kann dies aber auffassen als die Ableitung der hinterereinander geschalteten Abbildungen

$$F : \mathbb{R} \to \mathbb{R}^n, \quad t \mapsto \mathbf{x} + t \cdot \mathbf{v} \quad \textit{mit} \quad F'(0) = \mathbf{v}$$

und

$$f : \mathbb{R}^n \to \mathbb{R} \quad \textit{mit Ableitung} \quad f'(\mathbf{y}) = \nabla f(\mathbf{y}).$$

Die Kettenregel liefert

$$(f \circ F)'(0) = f'(F(0)) \cdot F'(0) = \underbrace{\nabla f(\mathbf{x})}_{1 \times n} \cdot \underbrace{\mathbf{v}}_{n \times 1}.$$

Dieses Matrizenprodukt liefert dasselbe Ergebnis wie früher.

Beispiel 2.13 *Beispiel 2.12 ist ein Spezialfall folgender Situation: Es sei $\gamma :]a, b[\to \mathbb{R}^n$ eine stetig differenzierbare Kurve und $f : U \to \mathbb{R}$ eine stetig differenzierbare Funktion auf einer offenen Menge $U \subset \mathbb{R}^n$, welche das Bild der Kurve enthält. Dann ist mit der Kettenregel*

$$\frac{d}{dt} f(\gamma(t)) = (\nabla f(\gamma(t)).\gamma'(t)).$$

Beispiel 2.14 *Es seien $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ offen, $F : U \to V$ differenzierbar und bijektiv mit differenzierbarer Umkehrabbildung $G = F^{-1} : V \to U$. Dann ist also*

$$G \circ F = id_U, \quad F \circ G = id_V.$$

Mit der Kettenregel und Beispiel 2.10 sehen wir

$$G'(F(\mathbf{x})) \cdot F'(\mathbf{x}) = \mathbb{1}_m \quad \textit{und} \quad F'(G(\mathbf{y})) \cdot G'(\mathbf{y}) = \mathbb{1}_n$$

für alle $\mathbf{x} \in U$ und $\mathbf{y} \in V$. Ist $\mathbf{y} = F(\mathbf{x})$, so sehen wir: Die Funktionalmatrizen

$$F'(\mathbf{x}) \quad \textit{und} \quad G'(\mathbf{y}) = (F^{-1})'(\mathbf{y})$$

sind invers zueinander. Insbesondere folgt:

- *die Dimensionen $m = n$ sind gleich,*
- *für alle $\mathbf{x} \in U$ ist die Funktionalmatrix $F'(\mathbf{x})$ invertierbar.*

Aufgabe 2.24 *Differenzieren Sie für $x > 0$ die Funktionen*

$$x^x \quad \textit{und} \quad x^{(x^x)}$$

mit Hilfe der Kettenregel in zwei Variablen.

Aufgabe 2.25 (Eulersche Formel) *Eine Funktion* $f : \mathbb{R}^n \to \mathbb{R}$ *heißt homogen vom Grad* p, $0 < p \in \mathbb{N}$, *wenn für alle* $\mathbf{x} \in \mathbb{R}^n$ *und* $t \in \mathbb{R}$ *gilt*

$$f(t \cdot \mathbf{x}) = t^p \cdot f(\mathbf{x}).$$

Zeigen Sie für jede stetig partiell differenzierbare, vom Grad p *homogene Funktion*

$$(\mathbf{x}.\nabla f(\mathbf{x})) = p \cdot f(\mathbf{x}).$$

Aufgabe 2.26 *Es sei* $f : \mathbb{R}^2 \to \mathbb{R}$ *eine stetig partiell differenzierbare Funktion. In etwas nachlässiger Notation schreibt man oft*

$$f(r, \varphi) = f(r \cdot cos(\varphi), r \cdot sin(\varphi)).$$

Es seien f_x *bzw.* f_y *die partiellen Ableitungen von* f *nach* x *bzw.* y *und analog* f_r *bzw.* f_φ *die partiellen Ableitungen von* $f(r, \varphi)$ *nach* r *bzw.* φ. *Zeigen Sie:*

$$\begin{array}{rcccc} f_r & = & f_x cos(\varphi) + f_y sin(\varphi) & = & \frac{1}{r}(x f_x + y f_y) \\ f_\varphi & = & r(-f_x sin(\varphi) + f_y cos(\varphi)) & = & -y f_x + x f_y \\ f_x & = & f_r cos(\varphi) - \frac{1}{r} f_\varphi sin(\varphi) & = & \frac{1}{r^2}(r f_r x - f_\varphi y) \\ f_y & = & f_r sin(\varphi) + \frac{1}{r} f_\varphi cos(\varphi) & = & \frac{1}{r^2}(r f_r y + f_\varphi x). \end{array}$$

Aufgabe 2.27 *Es sei* $A \in M(n \times n, \mathbb{R})$ *eine reelle* $n \times n$*-Matrix. Zeigen Sie: Die Abbildung*

$$\mathbb{R} \to M(n \times n, \mathbb{R}), \quad t \mapsto e^{t \cdot A}$$

ist differenzierbar mit Ableitung

$$A \cdot e^{t \cdot A}.$$

Aufgabe 2.28 *Es sei* $f(x, y)$ *eine zweimal stetig partiell differenzierbare Funktion auf* $\mathbb{R}^2$ *und*

$$f(r, \varphi) = f(x(r, \varphi), y(r, \varphi))$$

dieselbe Funktion in Polarkoordinaten. Zeigen Sie

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2}.$$

2.4 Umkehrung differenzierbarer Abbildungen, Implizite Funktionen

In Beispiel 2.14 sahen wir: Wenn eine differenzierbare Abbildung $F(\mathbf{x})$ eine differenzierbare Umkehrabbildung besitzt, dann ist die Funktionalmatrix $F'(\mathbf{x})$ invertierbar. Diese Aussage kehren wir jetzt um. Allerdings geht dies anders als in der Theorie einer Veränderlichen nur lokal.

Satz 2.14 *Es sei $M \subset \mathbb{R}^n$ offen, $F : M \to \mathbb{R}^n$ stetig differenzierbar und $\mathbf{x}_0 \in M$ so, dass $F'(\mathbf{x}_0)$ invertierbar ist. Dann gibt es offene Kugeln $U \subset M \subset \mathbb{R}^n$ um $\mathbf{x}_0$, $V \subset \mathbb{R}^n$ um $\mathbf{y}_0 := F(\mathbf{x}_0)$ und eine stetig differenzierbare Abbildung $G : V \to U$ mit $F(G(\mathbf{y})) = \mathbf{y}$ für alle $\mathbf{y} \in V$.*

Beweis. a) Zuerst zeigen wir die lokale Existenz einer *stetigen* Umkehrabbildung G. Dazu verwenden wir das vereinfachte Newton-Verfahren und den Banachschen Fixpunktsatz. Allerdings können wir die kontrahierende Abbildung nicht mehr F nennen, weil dieser Buchstabe schon verbraucht ist. Nennen wir sie C. Präzisieren wir zunächst die Definitionsbereiche unserer Abbildungen.

1) Wir wählen einen Radius $r > 0$ so klein, dass die abgeschlossene Kugel

$$U := \{\mathbf{x} \in \mathbb{R}^n : \parallel \mathbf{x} - \mathbf{x}_0 \parallel < r\}$$

ganz in M enthalten ist, und außerdem

$$\parallel \mathbb{1}_n - F'(\mathbf{x}_0)^{-1} \cdot F'(\mathbf{x}) \parallel < \frac{1}{2}$$

für alle $\mathbf{x} \in U$. Weil für $\mathbf{x} = \mathbf{x}_0$ gilt $\mathbb{1}_n = F'(\mathbf{x}_0)^{-1} \cdot F'(\mathbf{x})$, und $F'(\mathbf{x})$ stetig ist, ist das möglich. Weil außerdem die Determinante $det(F'(\mathbf{x}))$ stetig ist mit $det(F'(\mathbf{x}_0)) \neq 0$, können wir außerdem r noch so klein wählen, dass $F'(\mathbf{x})$ invertierbar ist für alle $\mathbf{x} \in U$.

2) Jetzt setzen wir

$$s := \frac{r}{2} \cdot \frac{1}{\parallel F'(\mathbf{x}_0)^{-1} \parallel}$$

und

$$V := \{\mathbf{y} \in \mathbb{R}^n : \parallel \mathbf{y} - \mathbf{y}_0 \parallel \leq s\}.$$

Nach Satz 1.16 ist der metrische Raum

$$C^0(V, U) := \{G : V \to U : G \text{ stetig}\}$$

vollständig. Wir definieren die Abbildung $C : G \mapsto C(G)$ durch

$$C(G)(\mathbf{y}) := G(\mathbf{y}) - F'(\mathbf{x}_0)^{-1} \cdot (F(G(\mathbf{y})) - \mathbf{y}).$$

Als erstes zeigen wir, dass $C : C^0(V, U) \to C^0(V, \mathbb{R}^n)$ kontrahierend ist. Dazu seien $G_1, G_2 \in C^0(V, U)$. Für alle $\mathbf{y} \in V$ haben wir

$$C(G_1)(\mathbf{y}) - C(G_2)(\mathbf{y}) = G_1(\mathbf{y}) - G_2(\mathbf{y}) - F'(\mathbf{x}_0)^{-1} \cdot (F(G_1(\mathbf{y}))) - F(G_2(\mathbf{y})).$$

Weil $G_1(\mathbf{y})$ und $G_2(\mathbf{y}) \in U \subset M$ liegen, gibt es nach dem MWS (Satz 2.4) ein $\xi \in U$ mit

$$F(G_1(\mathbf{y})) - F(G_2(\mathbf{y})) = F'(\xi) \cdot (G_1(\mathbf{y}) - G_2(\mathbf{y})).$$

Mit unserer Wahl von r sehen wir für alle $\mathbf{y} \in V$

$$\begin{aligned} \| C(G_1)(\mathbf{y}) - C(G_2)(\mathbf{y}) \| &= \| (\mathbb{1}_n - F'(\mathbf{x}_0)^{-1} \cdot F'(\xi)) \cdot (G_1(\mathbf{y}) - G_2(\mathbf{y})) \| \\ &\leq \| \mathbb{1}_n - F'(\mathbf{x}_0)^{-1} \cdot F'(\xi) \| \cdot \| G_1(\mathbf{y}) - G_2(\mathbf{y}) \| \\ &\leq \frac{1}{2} \cdot \| G_1(\mathbf{y}) - G_2(\mathbf{y}) \| . \end{aligned}$$

Daraus folgt

$$\| C(G_1) - C(G_2) \|_V \leq \frac{1}{2} \cdot \| G_1 - G_2 \|_V .$$

Unsere Abbildung C ist kontrahierend mit Lipschitz-Konstante 1/2.

Wir müssen noch zeigen, dass C eine Abbildung $C^0(V, U) \to C^0(V, U)$ definiert. Sei also $G \in C^0(V, U)$, d.h. $G : V \to U$ ist stetig mit $\| G(\mathbf{y}) - \mathbf{x}_0 \| \leq r$ für alle $y \in V$. Für die folgende Abschätzung fassen wir $\mathbf{x}_0$ auf als eine konstante Abbildung in $C^0(V, U)$. Dann folgt für alle $\mathbf{y} \in V$

$$\| C(G)(\mathbf{y}) - \mathbf{x}_0 \| \leq \| C(G)(\mathbf{y}) - C(\mathbf{x}_0) \| + \| C(\mathbf{x}_0) - \mathbf{x}_0 \| .$$

Mit der Kontraktionseigenschaft können wir den ersten Faktor abschätzen

$$\| C(G)(\mathbf{y}) - C(\mathbf{x}_0) \| \leq \frac{1}{2} \cdot \| G(\mathbf{y}) - \mathbf{x}_0 \| \leq \frac{r}{2}.$$

Und im zweiten Summanden haben wir

$$C(\mathbf{x}_0) - \mathbf{x}_0 = \mathbf{x}_0 - F'(\mathbf{x}_0)^{-1} \cdot (F(\mathbf{x}_0) - \mathbf{y}) - \mathbf{x}_0 = F'(\mathbf{x}_0)^{-1} \cdot (\mathbf{y}_0 - \mathbf{y}).$$

Daraus folgt

$$\| C(\mathbf{x}_0) - \mathbf{x}_0 \| \leq \| F'(\mathbf{x}_0)^{-1} \| \cdot s = \frac{r}{2}.$$

Insgesamt haben wir

$$\| C(G)(\mathbf{y}) - \mathbf{x}_0 \| \leq \frac{r}{2} + \frac{r}{2} = r,$$

also gehört $C(G)$ wieder zu $C^0(V, U)$.

Der Banachsche Fixpunktsatz liefert uns eine stetige Abbildung $G : V \to U$ mit $F(G(\mathbf{y})) = \mathbf{y}$ für alle $\mathbf{y} \in V$.

Wir müssen noch zeigen: G ist im Inneren von V differenzierbar. Dazu fixieren wir hier einen festen Punkt $\mathbf{y}_1$ (den können wir jetzt nicht mehr $\mathbf{y}_0$ nennen) und lassen $\mathbf{y} \in V$ gegen $\mathbf{y}_1$ gehen. Wir setzen $\mathbf{x}_1 = G(\mathbf{y}_1), \mathbf{x} = G(\mathbf{y}) \in U$ und haben nach dem MWS

$$\mathbf{y} - \mathbf{y}_1 = F'(\xi) \cdot (\mathbf{x} - \mathbf{x}_1)$$

mit einem Zwischenpunkt $\xi \in U$ auf der Strecke zwischen $\mathbf{x}$ und $\mathbf{x}_1$. Diese Gleichung schreiben wir um in

$$\begin{aligned} G(\mathbf{y}) - G(\mathbf{y}_1) &= \mathbf{x} - \mathbf{x}_1 \\ &= F'(\xi)^{-1} \cdot (\mathbf{y} - \mathbf{y}_1) \\ &= F'(\mathbf{x}_1)^{-1} \cdot (\mathbf{y} - \mathbf{y}_1) + (F'(\xi)^{-1} - F'(\mathbf{x}_1)^{-1}) \cdot (\mathbf{y} - \mathbf{y}_1) \\ &= F'(\mathbf{x}_1)^{-1} \cdot (\mathbf{y} - \mathbf{y}_1) + \Psi(\mathbf{y}). \end{aligned}$$

Weil F stetig differenzierbar ist, sind alle Matrix-Einträge in $F'(\mathbf{x})$ stetig und damit auch alle Einträge in der Matrix $F'(\mathbf{x})^{-1}$. Für $\mathbf{y} \to \mathbf{y}_1$ geht, wegen der Stetigkeit von G, auch $\mathbf{x} \to \mathbf{x}_1$ und damit $\xi \to \mathbf{x}_1$. Aus der erwähnten Stetigkeit folgt

$$\lim_{\mathbf{y}\to\mathbf{y}_1} \| F'(\xi)^{-1} - F'(\mathbf{x}_1)^{-1} \| = 0.$$

Damit wird

$$\lim_{\mathbf{y}\to\mathbf{y}_1} \frac{\| \Psi(\mathbf{y}) \|}{\| \mathbf{y} - \mathbf{y}_1 \|} \leq \frac{\| F'(\xi)^{-1} - F'(\mathbf{x}_1)^{-1} \| \cdot \| \mathbf{y} - \mathbf{y}_1 \|}{\| \mathbf{y} - \mathbf{y}_1 \|} = 0.$$

Wir haben gezeigt: G ist differenzierbar in $\mathbf{y}_1$ mit Funktionalmatrix $F'(\mathbf{x}_1)^{-1}$. □

Satz 2.15 (Zusatz) *Ist F k-mal stetig differenzierbar, dann ist es auch $G = F^{-1}$.*

Beweis. Es ist

$$G'(\mathbf{y}) = F'(G(\mathbf{y}))^{-1}.$$

Wenn F stetig differenzierbar ist, dann sind die Matrix-Einträge von F' und $(F')^{-1}$ stetig in $\mathbf{x}$, und weil G stetig ist, sind auch die Matrix-Einträge von G' stetig in $\mathbf{y}$. Durch Induktion nach k zeigt man, dass G genauso oft stetig differenzierbar ist wie F. □

Damit haben wir die Umkehrabbildung $G = F^{-1}$ bis auf einen kleinen Schönheitsfehler: Es ist zwar $F(G(\mathbf{y})) = \mathbf{y}$ für alle $\mathbf{y} \in V$, aber $G(F(\mathbf{x}))$ braucht nicht für alle $\mathbf{x} \in U$ definiert zu sein, weil $F(\mathbf{x})$ nicht in V zu liegen braucht. Das reparieren wir jetzt:

Satz 2.16 (Umkehrsatz) *Es sei $M \subset \mathbb{R}^n$ offen, $F : M \to \mathbb{R}^n$ stetig differenzierbar und $\mathbf{x}_0 \in M$ mit det$(F'(\mathbf{x}_0)) \neq 0$. Dann gibt es offene Mengen*

$$U \subset M \text{ mit } \mathbf{x}_0 \in U, \quad V \subset \mathbb{R}^n \text{ mit } \mathbf{y}_0 := F(\mathbf{x}_0) \in V$$

und eine steig differenzierbare Abbildung $G : V \to U$ so, dass $F : U \to V$ und $G : V \to U$ Umkehrabbildungen voneinander sind.

Beweis. Die abgeschlossenen Kugeln U, V aus Satz 2.14 bezeichnen wir jetzt mit U_0 und V_0. Es sei $V \subset V_0 \subset \mathbb{R}^n$ die offene Kugel um $\mathbf{y}_0$ vom gleichen Radius s wie V_0. Dann ist

$$U := F^{-1}(V) = \{\mathbf{x} \in M : F(\mathbf{x}) \in V\}$$

offen in $\mathbb{R}^n$ nach Satz 1.11 b). Mit diesen offenen Mengen U und V gilt die Behauptung. □

Ein Spezialfall des Umkehrsatzes ist ein lineares Gleichungssystem

$$A \cdot \mathbf{x} = \mathbf{b}$$

mit invertierbarer $n \times n$-Koeffizientenmatrix A. Zu jedem $\mathbf{b} \in \mathbb{R}^n$ hat dieses System die eindeutig bestimmte Lösung $\mathbf{x} = A^{-1} \cdot \mathbf{b}$. Die Abbildung $F\mathbf{x} \mapsto A \cdot \mathbf{x}$ hat die Umkehrabbildung $G : \mathbf{b} \mapsto A^{-1} \cdot \mathbf{b}$. Hier ist allerdings, im Unterschied zu Satz 2.16 die Umkehrabbildung global definiert.

Wesentlich allgemeiner sind Gleichungssysteme mit einer Koeffizientenmatrix A die nicht quadratisch ist. Betrachten wir etwa eine $p \times (n + p)$-Matrix A. Das System löst man, indem man A auf eine Zeilenstufenform bringt. In der Vorlesung Lineare Algebra versuche ich immer die Stufen möglichst weit links zu erzeugen. Aber das ist Vereinbarungssache. Wenn A den maximalen Rang $= p$ hat, kann man die Stufen innerhalb einer jeden $p \times p$-Teilmatrix von A erzeugen, die Rang p hat. Normieren wir die Situation folgendermaßen:

Es sei $A = (A_1, A_2)$ mit einer $p \times n$-Matrix A_1 und einer invertierbaren $p \times p$-Matrix A_2. Den Vektor der Unbekannten schreiben wir

$$(\mathbf{x}, \mathbf{y}) \quad \text{mit} \quad \mathbf{x} \in \mathbb{R}^n, \quad \mathbf{y} \in \mathbb{R}^p.$$

Das Gleichungssystem wird

$$A_1 \cdot \mathbf{x} + A_2 \cdot \mathbf{y} = \mathbf{b}, \quad \text{bzw.} \quad A_2 \cdot \mathbf{y} = \mathbf{b} - A_1 \cdot \mathbf{x}.$$

Dieses System wird gelöst durch

$$\mathbf{y} = A_2^{-1} \cdot (\mathbf{b} - A_1 \cdot \mathbf{x}),$$

wo wir $\mathbf{x} \in \mathbb{R}^n$ beliebig wählen können. Diese Tatsache können wir auch so formulieren:

Gegeben sei die Abbildung

$$F : \mathbb{R}^{n+p} \to \mathbb{R}^p, \quad (\mathbf{x}, \mathbf{y}) \mapsto A_1 \cdot \mathbf{x} + A_2 \cdot \mathbf{y} - \mathbf{b}$$

mit invertierbare Matrix

$$A_2 = \left(\frac{\partial F_\mu}{\partial y_\nu} \right)_{\mu,\nu=1,\ldots,p}.$$

Dann gibt es eine Abbildung

$$G : \mathbb{R}^n \to \mathbb{R}^p, \quad \mathbf{x} \mapsto \mathbf{y} = A_2^{-1} \cdot (\mathbf{b} - A_1 \cdot \mathbf{x}),$$

so, dass die Lösungen des Systems genau die Vektoren $(\mathbf{x}, G(\mathbf{x}))$, $\mathbf{x} \in \mathbb{R}^n$ sind. Mit der Abbildung G haben wir das Gleichungssystem nach $\mathbf{y}$ aufgelöst. Der folgende Satz verallgemeinert diese Tatsache auf nicht-lineare Gleichungssysteme. Dabei ist es aber wesentlich, dass die Auflösung nur lokal funktioniert.

Satz 2.17 (Implizite Funktionen) *Es sei $M \subset \mathbb{R}^{n+p}$ offen und $F : M \to \mathbb{R}^p$ stetig differenzierbar. Die Vektoren in $\mathbb{R}^{n+p}$ schreiben wie als $(\mathbf{x}, \mathbf{y})$ mit $\mathbf{x} \in \mathbb{R}^n$ und $\mathbf{y} \in \mathbb{R}^p$. Gegeben sei ein Punkt $(\mathbf{x}_0, \mathbf{y}_0) \in M$ mit*
1) $F(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$,
2) die $p \times p$-Matrix $(\partial F_i / \partial y_j(\mathbf{x}_0, \mathbf{y}_0))_{i,j=1,\ldots,p}$ habe maximalen Rang $= p$.
Dann gibt es eine offenen Menge $U \subset M$, die $(\mathbf{x}_0, \mathbf{y}_0)$ enthält, eine offenen Menge $V \subset \mathbb{R}^n$, die $\mathbf{x}_0$ enthält, eine stetig differenzierbare Abbildung $G : V \to \mathbb{R}^n$ so, dass gilt:

$$(\mathbf{x}, \mathbf{y}) \in U \text{ mit } F(\mathbf{x}, \mathbf{y}) = 0 \quad \Leftrightarrow \quad \mathbf{x} \in V \text{ und } \mathbf{y} = G(\mathbf{x}).$$

Schon für $n = p = 1$ ist dieser Satz nicht trivial. Bevor wir den Satz beweisen, wollen wir erst ein Beispiel (so ziemlich das einfachste und illustrativste) mit $n = p = 1$ diskutieren.

Beispiel 2.15 *Wir betrachten die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) := x^2 + y^2 - 1.$$

Diese Funktion ist stetig differenzierbar und ihre Nullstellenmenge

$$\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$$

ist der Einheitskreis. Die partielle Ableitung

$$\frac{\partial f}{\partial y} = 2 \cdot y$$

ist $\neq 0$ *für* $y \neq 0$. *Fixieren wir einen Punkt* (x_0, y_0) *auf dem Einheitskreis mit* $y_0 > 0$. *Dort kann man die Gleichung* $f(x, y) = 0$ *lokal nach* y *auflösen vermöge*

$$y = \sqrt{1 - x^2}.$$

In der Notation von Satz 2.17 bedeutet dies:

Die offenen Menge $V \subset \mathbb{R}$ *ist das Intervall* $]-1, 1[\subset \mathbb{R}$.

Die offenen Menge $U \subset \mathbb{R}^2$ *ist*

$$\{(x, y) \in \mathbb{R}^2 : x \in V, y > 0\}.$$

Die Abbildung $G : V \to \mathbb{R}$ *ist die Funktion* $g(x) = \sqrt{1 - x^2}$.

Im voranstehenden Beispiel ist die Gleichung

$$f(x, y) = x^2 + y^2 - 1 = 0$$

lokal nach y aufgelöst worden durch die Funktion $g(x) = \sqrt{1 - x^2}$. In der Situation des Satzes wird das Gleichungssystem $F(\mathbf{x}) = \mathbf{0}$, d.h.

$$F_1(\mathbf{x}, \mathbf{y}) = 0, ..., F_p(\mathbf{x}, \mathbf{y}) = 0$$

durch die Abbildung $\mathbf{y} = G(\mathbf{x})$, d.h., durch die Funktionen

$$y_1 = G_1(\mathbf{x}), ..., y_p = G_p(\mathbf{x})$$

lokal nach $y_1, ..., y_p$ aufgelöst. Man sagt: Die Funktionen $G_1(\mathbf{x}), ..., G_p(\mathbf{x})$ sind implizit durch das Gleichungssystem $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ definiert.

Beweis des Satzes. Wir definieren eine Abbildung $\Phi : U \to \mathbb{R}^{n+p}$ durch

$$\Phi(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \mathbf{x} \\ F(\mathbf{x}, \mathbf{y}) \end{pmatrix}.$$

Diese Abbildung ist stetig differenzierbar mit der Funktionalmatrix

$$\Phi' = \begin{pmatrix} \mathbb{1}_n & 0 \\ \dfrac{\partial F}{\partial \mathbf{x}} & \dfrac{\partial F}{\partial \mathbf{y}} \end{pmatrix}.$$

Nach Vorausssetzung 2) hat hier die rechte untere $p \times p$-Matrix in $(\mathbf{x}_0, \mathbf{y}_0)$ den maximalen Rang p. Dann hat die ganze Funktionalmatrix $\Phi'(\mathbf{x}_0, \mathbf{y}_0)$ im Punkt $(\mathbf{x}_0, \mathbf{y}_0)$ den Rang n. Der Umkehrsatz ist anwendbar. Es gibt also offenen Mengen $U \subset \mathbb{R}^{n+p}$ mit $(\mathbf{x}_0, \mathbf{y}_0) \in U$, sowie $V' \subset \mathbb{R}^{n+p}$ mit

$$\Phi(\mathbf{x}_0, \mathbf{y}_0) = (\mathbf{x}_0, F(\mathbf{x}_0, \mathbf{y}_0)) \in V'$$

und eine stetig differenzierbare Umkehrabbildung $\Psi : V' \to U$ für $\Phi : U \to V'$.

Dass Ψ die Abbildung Φ umkehrt, bedeutet explizit

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ y_1 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} \Psi_1(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) \\ \vdots \\ \Psi_n(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) \\ \Psi_{n+1}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) \\ \vdots \\ \Psi_{n+p}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) \end{pmatrix}.$$

Das bedeutet

$$\Psi_1(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = x_1, ..., \Psi_n(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = x_n$$

und

$$\Psi_{n+1}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = y_1, ..., \Psi_{n+p}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = y_p.$$

Unter Φ wird die Nullstellenmenge

$$\{(\mathbf{x}, \mathbf{y}) \in U : F_1(\mathbf{x}, \mathbf{y}) = ... = F_p(\mathbf{x}, \mathbf{y}) = 0\}$$

bijektiv auf die Menge

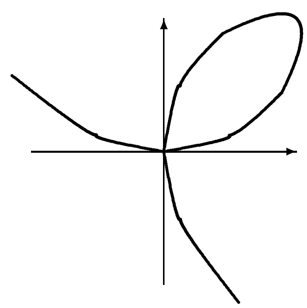
$$V := V' \cap \{(\mathbf{x}, \mathbf{0}) \in \mathbb{R}^{n+p}\}$$

abgebildet. Wir fassen V auf als offene Menge im $\mathbb{R}^n \subset \mathbb{R}^{n+p}$ der ersten n Koordinaten.

Die Abbildung $G : V \to U$ definieren wir durch

$$G(\mathbf{x}) = \begin{pmatrix} G_1(\mathbf{x}) \\ \vdots \\ G_p(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \Psi_{n+1}(\mathbf{x}, \mathbf{0}) \\ \vdots \\ \Psi_{n+p}(\mathbf{x}, \mathbf{0}) \end{pmatrix}.$$

Diese Abbildung ist stetig differenzierbar, weil Φ dies ist, und sie bildet V bijektiv auf die Menge $\{(\mathbf{x}, \mathbf{y}) \in U : F(\mathbf{x}, \mathbf{y}) = \mathbf{0})\}$ ab. □



Beispiel 2.16 (Cartesisches Blatt) *In der Zeichnung oben habe ich mal die Kurve mit der Gleichung*

$$x^3 + y^3 - xy = 0$$

geplottet. Wieder ist $p = n = 1$ und wir haben die Funktion

$$f(x, y) = x^3 + y^3 - xy.$$

Um Satz 2.17 anzuwenden berechnen wir

$$\frac{\partial f}{\partial y} = 3y^2 - x.$$

Falls (x_0, y_0) *ein Punkt der Kurve ist, wo diese Ableitung nicht verschwindet, können wir die Gleichung lokal nach* y *auflösen und als Graphen einer Funktion* $y = y(x)$ *darstellen. Nur, wo liegen die Punkte, wo das nicht geht, also wo*

$$x^3 + y^3 - xy = 0 \quad \textit{und} \quad x = 3y^2$$

gilt? Wir setzen $x = 3y^2$ *in die Gleichung der Kurve ein und erhalten*

$$27y^6 + y^3 - 3y^3 = y^3 \cdot (27y^3 - 2) = 0.$$

Dies führt auf zwei verbotene Werte für y*, nämlich*

1) $y = 0$ *und auch* $x = 0$. *Das ist der Nullpunkt, wo die Kurve sich selbst überkreutzt. Es ist ziemlich klar, dass man sie hier nicht als Graphen einer Funktion darstellen kann.*

2) $y = \frac{1}{3}\sqrt[3]{2}$ *und* $x = \frac{1}{3}\sqrt[3]{4}$. *Dies muss der Punkt im rechten oberen Quadranten sein, wo die Kurve eine senkrechte Tangente hat.*

Man kann die Parametrisierung $\mathbf{y} = G(\mathbf{x})$ der Nullstellenmenge

$$\{(\mathbf{x}, \mathbf{y}) : F(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$$

noch etwas weiter anlysieren. Es ist ja

$$F(\mathbf{x}, G(\mathbf{x})) \equiv \mathbf{0}.$$

Dann ist auch die Ableitung dieser geschachtelten Abbildung von $\mathbf{x}$ identisch $= 0$. Um das Resultat etwas übersichtlicher zu gestalten spalten wir F' auf in die beiden Matrizen aus den partiellen Ableitungen nach $\mathbf{x}$ und denen nach $\mathbf{y}$:

$$F' = \left(\underbrace{\frac{\partial F}{\partial \mathbf{x}}}_{p\times n}, \underbrace{\frac{\partial F}{\partial \mathbf{y}}}_{p\times p} \right).$$

Dann folgt für $(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, G(\mathbf{x}))$ mit der Kettenregel

$$\mathbf{0} = \frac{\partial F}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}) + \frac{\partial F}{\partial \mathbf{y}} \cdot G'(\mathbf{x}).$$

Weil die Matrix $\partial F/\partial \mathbf{y}$ invertierbar ist, können wir diese Gleichung umschreiben:

$$G'(\mathbf{x}) = -\left(\frac{\partial F}{\partial \mathbf{y}}\right)^{-1} \cdot \left(\frac{\partial F}{\partial \mathbf{x}}\right).$$

Wir haben bewiesen:

Satz 2.18 *Die Ableitung der durch die Gleichung $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ implizit definierten Abbildung $\mathbf{y} = \mathbf{y}(\mathbf{x})$ ist*

$$-\left(\frac{\partial F}{\partial \mathbf{y}}\right)^{-1} \cdot \left(\frac{\partial F}{\partial \mathbf{x}}\right).$$

Der Sinn dieser Aussage bsteht darin, dass man die Ableitung der implizit definierten Abbildung $\mathbf{y}(\mathbf{x})$ ausrechnen kann, ohne diese Abbildung wirklich anzugeben.

Schauen wir uns Beispiele an, wieder im einfachsten Fall $p = n = 1$.

Beispiel 2.17 *Betrachten wir $f(x, y) = x^2 - y^2 - 1 = 0$ mit der Parametrisierung*

$$y(x) = \sqrt{1 - x^2}$$

in der offenen Menge $|x| < 1, y > 0$. Die Ableitung der Parametrisierungs-Funktion ist

$$y'(x) = -\frac{x}{\sqrt{x^2 - 1}}.$$

Und die Formel aus Satz 2.18 liefert

$$y'(x) = -\frac{1}{f_y} \cdot f_x = -\frac{x}{y}.$$

Seitzt man hier $y = \sqrt{1 - x^2}$ ein, so erhält man das eben schon hergeleitete Ergebnis.

Beispiel 2.18 *Für $f(x, y) = x^3 + y^3 - xy = 0$ haben wir die Auflösung $y(x)$ nicht angegeben. Das geht im Prinzip mit der Formel für Polynomgleichungen vom Grad 3, führt aber zu weit vom Thema ab, und ist außerdem höllisch kompliziert. Aber mit*

$$f_x = 3x^2 - y, \quad f_y = 3y^2 - x$$

wird die Formel aus Satz 2.18

$$y'(x) = -\frac{1}{f_y} \cdot f_x = -\frac{3x^2 - y}{3y^2 - x}.$$

Insbesondere finden wir $y'(x) = 0$ genau für $3x^2 = y$. Das ist der Punkt im ersten Quadranten, wo die Kurve eine waagrechte Tangente hat. Wir können auch noch die zweite Ableitung von $y(x)$ ausrechnen. Mit der Quotientenregel erhalten wir

$$y''(x) = -\frac{1}{(3y^2 - x)^2} \cdot \Big((6x - y')(3y^2 - x) - (6yy' - 1)(3x^2 - y)\Big).$$

Das sieht schrecklich aus. Aber wenn wir

$$y' = 3x^2 - y = 0$$

einsetzen, schrumpft der Ausdruck zusammen auf

$$y' = -\frac{6x}{(3y^2 - x)^2}.$$

Weil $x > 0$ ist, haben wir $y''(x) < 0$. Damit haben wir verifiziert, dass die implizit definierte Funktion $y(x)$ im Punkt

$$(x, y) = \frac{1}{3}(\sqrt[3]{2}, \sqrt[3]{4})$$

ein lokales Maximum hat, und das, ohne die Funktion $y(x)$ überhaupt zu kennen.

Definition 2.8 (Niveaulinien) *Es sei $U \subset \mathbb{R}^2$ und $f: U \to \mathbb{R}$ gegeben. Für $c \in \mathbb{R}$ heißen die Teilmengen*

$$\{(x,y) \in U: \, f(x,y) = c\}$$

Niveaulinien *der Funktion f.*

Wenn U offen und f stetig differenzierbar ist, dann können wir den Satz über implizite Funktionen überall dort anwenden, wo der Gradient $\nabla f \neq \mathbf{0}$ ist. Dort ist entweder $\partial f/\partial y \neq 0$, und wir können die Gleichung $f(x,y) = c$ lokal nach y auflösen. Lokal ist die Niveaulinie der Graph einer Funktion $y = y(x)$. Oder es ist $\partial f/\partial x \neq 0$, dann können wir die Gleichung lokal nach x auflösen. Die Niveaulinie ist lokal der Graph einer Funktion $x = x(y)$.

Beispiel 2.19 *Es sei, wie schon so oft, $f(x,y) = x^2 + y^2$ mit*

$$\nabla f = (2x, 2y) = (0,0) \text{ nur für } x = y = 0.$$

Wenn $c > 0$ ist, liegt der Nullpunkt nicht auf der Niveaulinie, dem Kreis vom Radius $\sqrt{c}$. Und wir können auflösen, entweder mit $y = \sqrt{c - x^2}$ oder mit $x = \sqrt{c - y^2}$.

Und wenn $f(x,y) = x^3 + y^3 - xy$ ist, dann haben wir

$$\nabla f = (3x^2 - y, 3y^2 - x).$$

Wir können lokal auflösen, falls nicht

$$3x^2 = y \quad \text{und} \quad 3y^2 = x$$

gilt. Um diese vebotenen Punkte zu berechnen, substituieren wir $y = 3x^2$ in die zweite Gleichung und erhalten

$$3 \cdot (3x^2)^2 - x = x \cdot (27x^3 - 1) = 0.$$

Dies führt auf die beiden Punkte

1) *$x = y = 0$ und $c = 0$. Dies ist der Punkt, wo die Niveaulinie für $c = 0$ sich selbst überkreutzt.*

2) *$x = y = 1/3$ und $c = -1/27$. Was ist hier los? Auch dies ist ein kritischer Punkt der Funktion f. Wir berechnen die Hesse-Matrix*

$$H = \begin{pmatrix} 6x & -1 \\ -1 & 6y \end{pmatrix} \quad \text{mit der Determinante} \quad d = 36xy - 1.$$

   Im kritischen Punkt ist $d = 3 > 0$. Weil hier auch $x > 0$ ist, hat die Funktion f da ein lokales Minimum. Es gibt eine Kreisscheibe um den kritischen Punkt, in der alle Funktionswerte $f(x,y) > -1/27$ sind. Die Niveaulinie ist hier keine Linie, sondern besteht nur aus einem isolierten Punkt.

Eine Auflösung $y = y(x)$ oder $x = x(y)$ definiert eine lokale Parametrisierung

$$\gamma: \quad x \mapsto (x, y(x)) \quad \text{oder} \quad y \mapsto (x(y), y)$$

der Niveaulinie. Die Niveaulinie ist also eine parametrisierte Kurve mit Geschwindigkeitsvektor

$$\gamma' = \quad (1, y'(x)) \quad \text{bzw.} \quad (x'(y), 1) \quad \neq (0,0).$$

Die vom Geschwindigkeitsvektor aufgespannte Gerade heißt die *Tangente* an die Kurve im Punkt (x,y). Man kann sich diese Tangente im Punkt (x,y) angeheftet vorstellen, als affinen Unterraum des $\mathbb{R}^2$ oder abstrakt als linearen Unterraum des $T \subset \mathbb{R}^2$. Ihr orthogonale Komplement heißt die *Normale* an die Kurve im Punkt (x,y). Wegen $f(\gamma) = const$ ist nach der Kettenregel

$$\nabla f(x,y) \perp \gamma'.$$

Der Gradient von f spannt die Normale auf.

Wir müssen diese Situation auch in höheren Dimensionen analysieren.

Definition 2.9 (Untermannigfaltigkeit) *Es sei $M \subset \mathbb{R}^n$ offen und $F: M \to \mathbb{R}^p$ stetig differenzierbar. Weiter habe die Funktionalmatrix $F'(\mathbf{x})$ den (maximalen) Rang p für alle $\mathbf{x} \in M$. (Es ist also automatisch $p \leq n$.) Dann heißt die Menge*

$$X := \{\mathbf{x} \in M : F(\mathbf{x}) = \mathbf{0}\}$$

eine $n-p$-dimensionale Untermannigfaltigkeit *von M.*

Auf Untermannigfaltigkeiten werden wir im dritten Semester zurückkommen, weil wir dann über solche Untermannigfaltigkeiten integrieren werden. Hier wollen wir nur folgendes festhalten:

Zu jedem Punkt $\mathbf{x} \in X$ gibt es eine lokale Auflösung des Gleichungssystems $F(\mathbf{x}) = \mathbf{0}$ für X. D.h., unter den Koordinaten $x_1, ..., x_n$ gibt es p, nennen wir sie $\mathbf{t} = (t_1, ..., t_p)$ mit

$$Rang\left(\frac{\partial F_\nu}{\partial t_\mu}\right)(\mathbf{x}_0) = p,$$

und eine stetig differenzierbare Abbildung $\mathbf{t} \mapsto \mathbf{x}(\mathbf{t})$ so, dass lokal bei $\mathbf{x}_0$ gilt: $F(\mathbf{x}) = \mathbf{0}$ ist äquivalent zu $\mathbf{x} = \mathbf{x}(\mathbf{t})$. Ist $\mathbf{t}_0$ der Parameterpunkt für $\mathbf{x}_0$, so zeigt Satz 2.18, dass die Funktionalmatrix

$$\left(\frac{\partial x_\nu}{\partial t_\mu}\right)(\mathbf{t}_0)$$

den Rang p hat.

Definition 2.10 *Es sei $1 \leq m \leq p$. Wir variieren t_m, halten t_μ^0 fest für $\mu \neq m$ und betrachten die Kurve*

$$\gamma_m : t_m \mapsto \mathbf{x}(t_1^0, ..., t_{m-1}^0, t_m, t_{m+1}^0, ..., t_p^0).$$

Diese Kurve heißt Parameterkurve *auf X. Die Tangentialvektoren $\gamma_m'(t_m^0)$ an die p Parameterkurven γ_m in $\mathbf{x}_0$ sind linear unabhängig. Der von ihnen aufgespannte Untervektorraum*

$$T_X(\mathbf{x}_0) = \mathbb{R} \cdot \gamma_1'(t_1^0) + ... + \mathbb{R} \cdot \gamma_p'(t_p^0)$$

heißt der Tangentialraum *an die Untermannigfaltigkeit im Punkt $\mathbf{x}$. Sein orthogonales Komplement*

$$N_X(\mathbf{x}_0) = T_X(\mathbf{x}_0)^\perp \subset \mathbb{R}^n$$

heißt der Normalraum im Punkt $\mathbf{x}_0 \in X$.

Nach Definition sind T_X und N_X Untervektorräume des R^n. Stellt man sich diese Unterräume im Punkt $\mathbf{x}_0$ angeheftet vor, so erhält man die affinen Unterräume $\mathbf{x}_0 + T_X(\mathbf{x}_0)$ und $\mathbf{x}_0 + N_X(\mathbf{x}_0)$ des $\mathbb{R}^n$. Der Tangentialraum hat die Dimension $n-p$ der Untermannigfaltigkeit, der Normalraum hat als Dimension die Codimension dazu, nämlich p.

Beispiel 2.20 (Sphäre) *Bei der Transformation in Kugelkoordinaten halten wir den Radius r fest, etwa $r = 1$. Dann erhalten wir die Abbildung*

$$(\varphi, \theta) \mapsto \begin{pmatrix} cos(\varphi) \cdot cos(\theta) \\ sin(\varphi) \cdot cos(\theta) \\ sin(\theta) \end{pmatrix}.$$

Damit haben wir eine Parametrisierung der Einheitssphäre

$$\{\mathbf{x} \in \mathbb{R}^3 : \| \mathbf{x} \| = 1\} \subset \mathbb{R}^3.$$

Sie wird implizit beschrieben durch die Gleichung

$$f(\mathbf{x}) = 0 \quad mit \quad f(x, y, z) = x^2 + y^2 + z^2 - 1.$$

Die Funktionalmatrix der Parametrisierungsabbildung ist

$$\left(\frac{\partial \mathbf{x}}{\partial \varphi}, \frac{\partial \mathbf{x}}{\partial \theta}\right) = \begin{pmatrix} -sin(\varphi) \cdot cos(\theta) & -cos(\varphi) \cdot sin(\theta) \\ cos(\varphi) \cdot cos(\theta) & -sin(\varphi) \cdot sin(\theta) \\ 0 & cos(\theta) \end{pmatrix}.$$

Für $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$ hat diese Funktionalmatrix den Rang 2. Auf der Erdkugel nennt man die Parameterkurven zu φ die Breitenkreise und die Parameterkurven zu θ die Längenkreise. Die beiden Spalten der eben berechneten Funktionalmatrix spannen den Tangentialraum an die Einheitssphäre auf. Der Normalraum wird aufgespannt vom Ortsvektor $\mathbf{x} = \frac{1}{2}\nabla f$.

Aufgabe 2.29 *Die Abbildung $F : \mathbb{R}^2 \to \mathbb{R}^2$ sei definiert durch*

$$u(x, y) = x^2 - y^2, \quad v(x, y) = 2xy.$$

Wo ist F lokal umkehrbar? Bestimmen Sie eine lokale Umkehrung

$$F^{-1} : \quad x = x(u, v),\ y = y(u, v)$$

bei $(u, v) = (1, 0)$ derart, dass $F^{-1}(1, 0) = (1, 0)$.

Aufgabe 2.30 *Wo ist*

$$e^{x-y} - (x + y)^2 = 0$$

lokal nach y auflösbar? Berechnen Sie dort die Ableitung der Auflösung $y(x)$.

Aufgabe 2.31 *Für* $x, y \in \mathbb{R}$ *sei* $F(x,y) := e^x \cdot (cos(y), sin(y))$.
a) Berechnen Sie die Jacobi-Matrix $F'(x,y)$ *und begründen Sie, weshalb* F *auf* $\mathbb{R}^2$ *total differenzierbar ist.*
b) Es sei

$$A = \{(x,y) \in \mathbb{R}^2 : \quad F \text{ um } (x,y) \text{ lokal umkehrbar}\}.$$

Zeigen Sie:

$$A = \{(x,y) \in \mathbb{R}^2 : \quad y \notin \frac{\pi}{2} + \mathbb{Z}\pi\}.$$

c) Bestimmen Sie die Ableitung der lokalen Inversen von F *um* $(0,0)$ *im Punkt* $F(0,0)$.

Aufgabe 2.32 *Es sei*

$$F : \mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto ((x+y)^3, (x-y)^7).$$

a) Beweisen Sie, dass F *injektiv ist.*
b) Welche Punkte besitzen eine Umgebung U *so, dass die Umkehrabbildung* $(F|U)^{-1}$ *existiert und differenzierbar ist?*

2.5 Lokale Extrema

Hier wollen wir zunächst das hinreichende Kriterium für lokale Extrema differenzierbarer Funktionen von mehreren Variablen beweisen. Wir müssen dabei immer annehmen, dass wir einen inneren Punkt des Definitionsbereiches der Funktion f haben. Also ist die Voraussetzung vorläufig:

$$U \subset \mathbb{R}^n \text{ ist offen, } f : U \to \mathbb{R} \text{ ist zweimal stetig partiell differenzierbar.}$$

Wir können dann auch gleich annehmen, dass $U = K$ eine offene Kugel um den Punkt $\mathbf{x}$ ist, in dem wir die Extremumseigenschaft für f nachweisen wollen.

Natürlich muss in diesem Punkt das notwendige Kriterium für lokale Extrema (Satz 2.1) erfüllt sein: Alle ersten partiellen Ableitungen müssen dort verschwinden,

$$\nabla f(\mathbf{x}) = \mathbf{0}.$$

Information über die Funktionswerte von f beziehen wir aus der Taylor-Formel (Satz 2.9) für $m = 2$ um den Entwicklungspunkt ($\mathbf{x}$). Wenn die ersten partiellen Ableitungen $= 0$ sind, sieht die folgendermaßen aus:

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \frac{1}{2}\mathbf{h}^t \cdot H_f(\mathbf{x}) \cdot \mathbf{h} + R_2(\mathbf{h})$$

mit

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{R_2(\mathbf{h})}{\| \mathbf{h} \|^2} = 0.$$

Der homogene Anteil zweiter Ordnung ist das, was man in der linearen Algebra eine *quadratische Form* nennt. Und es wird darauf ankommen, ob dieser Anteil für alle $\mathbf{h} \neq \mathbf{0}$ größer als 0 oder kleiner als 0 ist. Und genau das untersucht man in der Linearen Algebra. Dort hat man sogar eine eigenen Definition.

Definition 2.11 *Eine reelle symmetrische $n \times n$-Matrix H heißt* positiv definit*, wenn für alle Vektoren $\mathbf{h} \in \mathbb{R}^n, \mathbf{h} \neq \mathbf{0}$ gilt*

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t > 0.$$

Sie heißt negativ definit*, wenn dieser Ausdruck immer < 0 ist, und sie heißt* indefinit*, wenn es Vektoren $\mathbf{h}$ gibt, für die das Resultat > 0 ist, und andere, für die es < 0 ist.*

Satz 2.19 *a) Die $n \times n$-Matrix H ist genau dann positiv definit, wenn es ein $m > 0$ gibt so, dass für alle Einheitsvektoren $\mathbf{h}_1 \in \mathbb{R}^n$ (mit $\| \mathbf{h}_1 \| = 1$) gilt*

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq m.$$

b) Die $n \times n$-Matrix H ist genau dann negativ definit, wenn es ein $m < 0$ gibt so, dass für alle Einheitsvektoren $\mathbf{h}_1 \in \mathbb{R}^n$ gilt

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \leq m.$$

c) Die $n \times n$-Matrix H ist genau dann indefinit, wenn es Einheitsvektoren $\mathbf{h}_+$ und $\mathbf{h}_-$ gibt mit

$$\mathbf{h}_+ \cdot H \cdot \mathbf{h}_+^t > 0, \quad \mathbf{h}_- \cdot H \cdot \mathbf{h}_-^t < 0.$$

Beweis. a) Sei H positiv definit vorausgesetzt. Dann ist für alle Einheitsvektoren $\mathbf{h}_1 \in \mathbb{R}^n$

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t > 0,$$

denn Einheitsvektoren sind $\neq \mathbf{0}$. Nun ist die Funktion

$$\mathbf{h} \mapsto \mathbf{h} \cdot H \cdot \mathbf{h}^t = \sum_{\mu,\nu=1}^{n} H_{\mu,\nu} h_\mu h_\nu$$

auf $\mathbb{R}^n$ ein Polynom und damit stetig. Weiter ist die Einheitssphäre

$$\{\mathbf{h}_1 \in \mathbb{R}^n : \| \mathbf{h}_1 \| = 1\}$$

beschränkt und abgeschlossen, also kompakt. Nach dem Satz vom Maximum gibt es einen Einheitsvektor, in dem sie ihr Minimum m annimmt. Auch dieses Minimum m ist > 0. Mit diesem Minimum m gilt somit

$$\mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq m$$

für alle Einheitsvektoren $\mathbf{h}_1$.

Sei nun umgekehrt diese Ungleichung vorausgesetzt. Jeden Vektor $\mathbf{h} \in \mathbb{R}^n, \quad \mathbf{h} \neq \mathbf{0}$, kann man schreiben

$$\mathbf{h} = \| \mathbf{h} \| \cdot \mathbf{h}_1 \text{ mit dem Einheitsvektor } \mathbf{h}_1 := \frac{1}{\| \mathbf{h} \|} \cdot \mathbf{h}.$$

Damit wird

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t = \| \mathbf{h} \|^2 \cdot \mathbf{h}_1 \cdot H \cdot \mathbf{h}_1^t \geq \| \mathbf{h} \|^2 \cdot m > 0.$$

b) Im negativ definiten Fall verläuft der Beweis genau so, man muss nur das $>$-Zeichen ersetzen durch $<$.

c) ist offensichtlich. □

Satz 2.20 (Lokale Extrema, hinreichende Bedingung) *Die Funktion $f : U \to \mathbb{R}$ sei zweimal stetig partiell differenzierbar auf der offenen Menge $U \subset \mathbb{R}^n$. In einem Punkt $\mathbf{x} \in U$ gelte*

$$\nabla f(\mathbf{x}) = \mathbf{0}.$$

Ist die Hesse-Matrix H_f von f im Punkt $\mathbf{x}$ definit, so liegt ein lokales Extremum vor, und zwar ein

$$\begin{array}{lcl} \textit{Maximum} & \Leftrightarrow & H_f \textit{ negativ definit} \\ \textit{Minimum} & \Leftrightarrow & H_f \textit{ positiv definit.} \end{array}$$

Ist die Hesse-Matrix indefinit, so liegt kein lokales Extremum vor.

Beweis. Wir benutzen die Taylor-Formel für $m = 2$

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \frac{1}{2} \cdot \mathbf{h} \cdot H \cdot \mathbf{h}^t + R_2(\mathbf{h}) \quad \text{mit} \quad \lim_{\mathbf{h}\to 0} \frac{R_2(\mathbf{h})}{\| \mathbf{h} \|^2} = 0.$$

a) Sei $H = H_f$ positiv definit und $m > 0$ wie in Satz 2.19 a) gewählt. Dann gilt also

$$\mathbf{h} \cdot H \cdot \mathbf{h}^t \geq m \cdot \| \mathbf{h} \|^2 .$$

Andererseits gibt es ein $r > 0$ mit

$$\frac{|R_2(\mathbf{h})|}{\| \mathbf{h} \|^2} < \frac{m}{4}, \quad \text{d.h.} \quad |R_2(\mathbf{h})| < \frac{m}{4} \cdot \| \mathbf{h} \|^2$$

für $\| \mathbf{h} \| < r$. Für diese $\mathbf{h}$ ist dann

$$\frac{1}{2} \cdot \mathbf{h} \cdot H \cdot \mathbf{h}^t + R_2(\mathbf{h}) \geq (\frac{m}{2} - \frac{m}{4}) \| \mathbf{h} \|^2 > 0$$

und

$$f(\mathbf{x}+\mathbf{h}) > f(\mathbf{x})$$

für $\mathbf{h} \neq \mathbf{0}$. In $\mathbf{x}$ hat die Funktion f ein lokales Minimum.

b) Ist H negativ definit, so sieht man mit Satz 2.19 b) ganz analog, dass $\mathbf{x}$ ein lokales Maximum für f ist.

c) Sei nun H indefinit. Nach Satz 2.19 c) gibt es ein $m > 0$ und Einheitsvektoren h_+, h_- mit

$$\mathbf{h}_+ \cdot H \cdot \mathbf{h}_+^t \geq m, \quad \mathbf{h}_- \cdot H \cdot \mathbf{h}_-^t < -m.$$

Wie im Beweis von a) findet man ein $r_+ > 0$ so, dass für $0 < |t| < r_+$ gilt

$$\frac{1}{2} \cdot (t\mathbf{h}_+) \cdot H \cdot (t\mathbf{h}_+^t) + R_2(t\mathbf{h}_+) > 0.$$

Ebenso gibt es ein $r_- > 0$ mit

$$\frac{1}{2} \cdot (t\mathbf{h}_+) \cdot H \cdot (t\mathbf{h}_+^t) + R_2(t\mathbf{h}_+) < 0$$

für $0 < |t| < r_-$. In jeder Kugel mit einem Radius

$$r < \min\{r_+, r_-\}$$

gibt es deswegen Punkte

$$\mathbf{x} + t \cdot h_+ \text{ mit } f(\mathbf{x} + t \cdot \mathbf{h}_+) > f(\mathbf{x})$$

und Punkte

$$\mathbf{x} + t \cdot h_- \text{ mit } f(\mathbf{x} + t \cdot \mathbf{h}_-) < f(\mathbf{x}).$$

In $\mathbf{x}$ hat die Funktion f weder ein lokales Maximum, noch ein lokales Minimum. $\square$

Natürlich bleibt jetzt noch das Problem, einer symmetrischen Matrix H anzusehen, ob sie positiv definit, negativ definit oder indefinit ist. (Dummerweise gibt es leider auch noch symmetrische Matrizen, die nichts von alledem sind, wo unsere wundervolle Theorie nicht greift.)

Einer Diagonalmatrix D sieht man sehr leicht an, ob sie eine dieser drei Eigenschaften besitzt: Sie ist

positiv definit	$\Leftrightarrow$	alle Diagonal-Einträge sind > 0
negativ definit	$\Leftrightarrow$	alle Diagonal-Einträge sind < 0
indefinit	$\Leftrightarrow$	es gibt Diagonal-Einträge > 0 und < 0.

Nicht jede symmetrische Matrix H ist eine Diagonalmatrix. Aber in der Linaren Algebra beweist man auch den Satz von der Hauptachsen-Transformation: Jede symmetrische Matrix H lässt sich mit einer orthogonalen Matrix T in eine Diagonalmatrix $D = T^t \cdot H \cdot T$ transformieren. Die Diagonaleinträge von D sind dabei die Eigenwerte von H. Daraus folgt: Die symmetrische Matrix H ist

positiv definit	$\Leftrightarrow$	alle ihre Eigenwerte sind > 0
negativ definit	$\Leftrightarrow$	alle ihre Eigenwerte sind < 0
indefinit	$\Leftrightarrow$	sie hat Eigenwerte > 0 und < 0.

Das ist sehr schön, vor allem, wenn man die Eigenwerte kennt. Bei 2×2-Matrizen kann man die immer ziemlich leicht ausrechnen. Bei größeren Matrizen ist es i.A. schwierig. Dafür gibt es dann das *Hurwitz-Kriterium*. Da muss man alle linken oberen Unterdeterminanten (Haupt-Minoren) der Matrix ausrechnen. Sind alle diese Haupt-Minoren > 0, so ist die Matrix positiv definit. In der linearen Algebra beweist man:

Satz 2.21 (Hurwitz-Kriterium) *Gegeben sei die symmetrische $n \times n$-Matrix*

$$H = \begin{pmatrix} h_{1,1} & ... & h_{1,n} \\ \vdots & & \vdots \\ h_{n,1} & ... & h_{n,n} \end{pmatrix}$$

Ihre n Hauptminoren *sind die linken oberen Unterdeterminanten*

$$d_k := det \begin{pmatrix} h_{1,1} & ... & h_{1,k} \\ \vdots & & \vdots \\ h_{k,1} & ... & h_{k,k} \end{pmatrix}, \quad k = 1, ..., n.$$

Die Matrix H ist positiv definit genau dann, wenn

$$h_k > 0 \quad \textit{für} \quad k = 1, ..., n.$$

Natürlich kann man dieses Kriterium auch für negativ-definite Matrizen formulieren: Es ist ja H negativ definit, genau dann, wenn die Matrix $-H$ positiv definit ist. Und die Haupt-Minoren der Matrix $-H$ sind die Determinanten

$$(-1)^k d_k.$$

Damit erhalten wir

Satz 2.22 *Die symmetrische Matrix H ist negativ definit, wenn für ihre Hauptminoren d_k gilt*

$$(-1)^k d_k > 0, \quad k = 1, ..., n,$$

d.h.

$$h_1 < 0,\ h_2 > 0,\ h_3 < 0, ...$$

Im Fall $n = 2$ kann man sogar noch ein einfaches Kriterium für Indefinitheit formulieren.

Satz 2.23 *Die symmetrische 2×2-Matrix*

$$H = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

mit der Determinante $d = ac - b^2$ ist

$$\begin{array}{lcc} \textit{positiv definit} & \Leftrightarrow & a > 0 \textit{ und } d > 0, \\ \textit{negativ definit} & \Leftrightarrow & a < 0 \textit{ und } d > 0, \\ \textit{indefinit} & \Leftrightarrow & d < 0. \end{array}$$

Beweis. Wir brauchen uns nur um die Indefinitheit kümmern. Dann gibt es zwei Eigenwerte λ_1, λ_2 mit $d = \lambda_1 \cdot \lambda_2$. Und genau dann, wenn $d < 0$ ist, haben diese verschiedene Vorzeichen. qed

Beispiel 2.21 *Es seien für die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad f(x, y) := x^2 + y^2 - e^{xy}$$

alle lokalen Extrema gesucht.

Wir differenzieren

$$\begin{array}{rcl} f_x & = & 2x - e^{xy} \cdot y \\ f_y & = & 2y - e^{xy} \cdot x \end{array}$$

und sicherheitshalber nochmal

$$\begin{array}{rclcl} f_{xx} & = & 2 - e^{xy} \cdot y^2 & & \\ f_{xy} & = & -e^{xy} \cdot xy - e^{xy} & = & -e^{xy} \cdot (xy + 1) \\ f_{yy} & = & 2 - e^{xy} \cdot x^2 & & \end{array}$$

Die notwendigen Bedingungen $f_x = f_y = 0$ schreiben wir als homogenes lineares Gleichungssystem

$$\begin{pmatrix} 2x & y \\ 2y & x \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -e^{xy} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

mit dem nicht-trivialen Lösungsvektor $(1, -e^{xy})$. *Wegen der nicht-trivialen Lösung muss die Determinante der Koeffizientenmatrix*

$$det \begin{pmatrix} 2x & y \\ 2y & x \end{pmatrix} = 2x^2 - 2y^2 = 0$$

sein. Wir finden also

$$x = \pm y.$$

Leider sind wir noch nicht fertig. Wenn $x = y$ *ist müssen wir für* $x = y \neq 0$ *die Bedingung*

$$2 - e^{xy} = 0, \quad 2 - e^{x^2} = 0, \quad x^2 = ln(2)$$

auswerten und finden

$$x = y = \sqrt{ln(2)}.$$

Natürlich ist auch noch

$$x = y = 0$$

eine Möglichkeit. Wenn $x = -y \neq 0$ *sein sollte, dann haben wir*

$$2 + e^{xy} = 0$$

auszuwerten. Aber wegen $e^{xy} > 0$ *gibt es da keine Lösung. Es gibt also nur die beiden folgenden stationären Punkte*

$$(x, y) = (\sqrt{\ln(2)}, \sqrt{\ln(2)}), \quad (x, y) = (0, 0).$$

Jetzt müssen wir die Hesse-Matrix

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$$

untersuchen. Mit den oben ausgerechneten zweiten Ableitungen finden wir

(x, y)	H	$det(H)$
$(0, 0)$	$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$	3
$(\sqrt{ln(2)}, \sqrt{ln(2)})$	$\begin{pmatrix} 2(1 - ln(2)) & -2(ln(2) + 1) \\ -2(ln(2) + 1) & 2(1 - ln(2)) \end{pmatrix}$	$4((1 - ln(2))^2 - (1 + ln(2))^2)$ $= -16ln(2)^2$

Im ersten Fall ist die Determinante positiv und das Hurwitz-Kriterium zeigt: H *ist positiv definit, der kritische Punkt ist ein lokales Minimum. Im zweiten Fall ist die Determinante negativ und das Hurwitz-Kriterium zeigt:* H *ist indefinit. Der kritische Punkt ist ein Sattelpunkt, weder lokales Maximum, noch lokales Minimum.*

Es gibt noch einen wichtigen Problemkreis in Zusammenhang mit der Ermittlung lokaler Extrema, der in der Theorie einer Veränderlichen nicht auftaucht, nämlich das Problem, lokale Extrema einer Funktion f unter *Nebenbedingungen* zu finden. Bevor ich das Problem präzise und allgemein formuliere, erst ein einfaches Beispiel.

Beispiel 2.22 *Es sind die globalen Extrema der Funktion*

$$f(x,y) := x^2 \cdot y$$

auf dem Einheitskreis $\{(x,y) : x^2+y^2 \leq 1\}$ *gesucht. Weil die Funktion stetig und der Einheitskreis kompakt ist, muss die Funktion diese Extrema irgenwo annehmen. Wenn sie das im Inneren des Einheitskreises tut, ist Satz 2.1 anwendbar und liefert*

$$f_x = 2xy = 0, \quad f_y = x^2 = 0.$$

Das sind die Punkte auf der y-Achse, wo $x = 0$ *ist. Weil hier* $f(x,y) = 0$ *ist, und weil* f *im Einheitskreis sowohl echt positive, als auch echt negative Werte annimmt, kann hier kein globales Extremum liegen. Die Extrema müssen auf dem Kreisrand liegen. Da gibt es mehrere Möglichkeiten, weiter zu kommen.*

Wir können den Kreisrand parametrisieren durch

$$x = cos(t),\ y = sin(t).$$

Dann wird

$$f(x,y) = f(t) = cos^2(t) \cdot sin(t).$$

In einem lokalen Extremum muss

$$f'(t) = -2cos(t) \cdot sin^2(t) + cos^3(t) = cos(t) \cdot (cos^2(t) - 2sin^2(t)) = 0$$

sein. Falls hier $cos(t) = 0$ *ist, sind wir in einem der Punkte auf der y-Achse. Sonst bleibt noch*

$$cos(t)^2(t) - 2sin^2(t) = 0.$$

Weil auf dem Kreisrand gilt $cos^2(t) = 1 - sin^2(t)$ *führt dies auf*

$$3sin^2(t) = 1, \quad y = sin(t) = \pm\frac{1}{\sqrt{3}}, \quad x = \pm\frac{\sqrt{2}}{\sqrt{3}}.$$

Der Funktionswert ist hier

$$x^2 y = \pm\frac{2}{3\sqrt{3}}.$$

Für positives Vorzeichen haben wir das Maximum, für negatives Vorzeichen das Minimum.

Wir können den Kreisrand auch lokal parametrisieren durch

$$x = \sqrt{1-y^2} \quad \text{für} \quad |y| < 1.$$

Dann wird

$$f(x,y) = (1-y^2) \cdot y = y - y^3.$$

In einem lokalen Extremum ist

$$(y - y^3)' = 1 - 3y^2 = 0.$$

Dies führt auf die gleichen Punkte, wie eben.

Bei beiden, eben durchgerechneten Verfahren, muss man die Funktion kennen, mit der die Kreislinie parametrisiert wird. Bei komplizierteren Nebenbedingungen versagt es meist. In diesen Situationen ist folgendes Verfahren anwendbar: Wir nehmen irgend eine Parametrisiereung

$$x = x(t),\ y = y(t),$$

der Kreislinie mit Geschwindigkeitsvektor

$$(x'(t), y'(t)) \neq (0,0).$$

In einem lokalen Extremum verschwindet

$$\frac{d}{dt} f(x(t), y(t)) = f_x \cdot x'(t) + f_y \cdot y'(t).$$

Das hilft noch nicht weiter. Aber auf der Kreislinie ist ja $x(t)^2 + y(t)^2 = 1$ *und*

$$2 \cdot (x(t) \cdot x'(t) + y(t) \cdot y'(t)) = 0.$$

In einem lokalen Extremum auf dem Kreisrand müssen die beiden Bedingungen

$$\begin{array}{rcrcl} f_x \cdot x'(t) & + & f_y \cdot y'(t) & = & 0 \\ x(t) \cdot x'(t) & + & y(t) \cdot y'(t) & = & 0 \end{array}$$

erfüllt sein. Dies ist ein homogenes lineares Gleichungssystem mit dem Lösungsvektor $(x'(t), y'(t)) \neq (0,0)$. *Dann muss die Determinante der Koeffizientenmatrix*

$$f_x \cdot y - f_y \cdot x = 2xy^2 - x^3 = x \cdot (2y^2 - x^2)$$

verschwinden. Das führt auf genau dieselben Werte wie vorher. Der entscheidende Unterschied ist, dass wir die parametrisierenden Funktionen $x(t)$ *und* $y(t)$ *überhaupt nicht zu kennen brauchen.*

Nach diesem Beispiel möchte ich das Kriterium für lokale Extrema unter Nebenbedingungen ganz allgemein formulieren.

Satz 2.24 (Extrema unter Nebenbedingungen) *Es sei* $M \subset \mathbb{R}^n$ *offen und* $f : M \to \mathbb{R}$ *stetig differenzierbar. Weiter sei die Abbildung* $F : M \to \mathbb{R}^p$ *stetig differenzierbar. Es sei* $\mathbf{x}_0 \in M$ *ein Punkt, in dem der Rang der Funktionalmatrix*

$$\left(\frac{\partial F_\mu}{\partial x_\nu}(\mathbf{x}_0) \right)_{\mu=1,\ldots,p,\,\nu=1,\ldots,n}$$

maximal $= p$ *ist. Nimmt* f *in* $\mathbf{x}_0 \in M$ *ein lokales Extremum unter den Nebenbedingungen*

$$F_1(\mathbf{x}) = F_1(\mathbf{x}_0), \ldots, F_p(\mathbf{x}) = F_p(\mathbf{x}_0)$$

an, dann gibt es Skalare $\lambda_1, \ldots \lambda_p \in \mathbb{R}$ *(* Lagrange-Multiplikatoren*) so, dass gilt*

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla F_1(\mathbf{x}_0) + \ldots + \lambda_p \nabla F_p(\mathbf{x}0).$$

Beweis. Weil die Funktionalmatrix von F im Punkt $\mathbf{x}_0$ maximalen Rang hat, können wir nach Satz 2.17 die Menge

$$X = \{\mathbf{x} \in M : F(\mathbf{x}) = F(\mathbf{x}_0)\}$$

lokal durch $n-p$ Parameter $\mathbf{t} = (t_1, ..., t_{n-p}) \in \mathbb{R}^{n-p}$ parametrisieren. Ist $\mathbf{t}_0$ der Parameterpunkt zu $\mathbf{x}_0$, so spannen die $n-p$ Tangentialvektoren

$$\frac{\partial \mathbf{x}}{\partial t_\mu}(\mathbf{t}_0), \quad m = 1, ..., n-p,$$

den Tangentialraum $T_{\mathbf{x}_0}(X)$ auf. Der Normalraum $N_{\mathbf{x}_0}(X)$ wir aufgespannt von den p Gradientenvektoren $\nabla F_\nu(\mathbf{x}_0)$, $\nu = 1, ..., p$. Die Behauptung ist äquivalent dazu, dass $\nabla f(\mathbf{x}_0)$ zu diesem Normalraum gehört.

Nun hat die Funktion $\mathbf{t} \mapsto f(\mathbf{x}(\mathbf{t}))$ bei $\mathbf{t} = \mathbf{t}_0$ ein lokales Extremum. Nach Satz 2.1 verschwinden für $\mu = 1, ..., n-p$ die partiellen Ableitungen

$$\frac{\partial f(\mathbf{x}(\mathbf{t}))}{\partial t_\mu}(\mathbf{t}_0) = \sum_{\nu=1}^{n} \frac{\partial f}{\partial x_\nu}(\mathbf{x}_0) \cdot \frac{\partial x_\nu}{\partial t_\mu}(\mathbf{t}_0).$$

Das heißt

$$(\nabla f(\mathbf{x}_0). \frac{\partial \mathbf{x}}{\partial t_\mu}(\mathbf{t}_0)) = 0.$$

Also steht $\nabla f(\mathbf{x}_0)$ senkrecht auf dem Tangentialraum und gehört zum Normalraum $N_{\mathbf{x}_0}(X)$. □

Beispiel 2.23 *Es sei S eine symmetrische $n \times n$-Matrix und q die quadratische Form $q(\mathbf{x}) = \mathbf{x}^t \cdot S \cdot \mathbf{x}$. Ihr Gradient ist (Aufgabe 2.14) $\nabla q(\mathbf{x}) = 2S \cdot \mathbf{x}$. Wir betrachten q auf der Einheitssphäre, d.h., unter der Nebenbedingung*

$$f(\mathbf{x}) = x_1^2 + ... + x_n^2 - 1 = 0.$$

Der Gradient von f ist

$$\nabla f(\mathbf{x}) = 2\mathbf{x} \neq 0 \quad \textit{für} \quad \| \mathbf{x} \| = 1.$$

Nun ist die Einheitssphäre kompakt und die Funktion q nimmt in einem Punkt $\mathbf{x}_0$ auf der Einheitssphäre ihr Maximum an. Aus Satz 2.24 mit $p = 1$ folgt

$$\nabla q(\mathbf{x}_0) = \lambda \nabla f(\mathbf{x}_0), \quad \textit{bzw.} \quad S \cdot \mathbf{x}_0 = \lambda \cdot \mathbf{x}_0.$$

Wir haben bewiesen: Die symmetrische Matrix S hat einen reellen Eigenvektor $\mathbf{x}_0$.

Aufgabe 2.33 *Sei*

$$f(x, y) := \frac{1}{x} - \frac{1}{y} + 4x - 9y \quad (x \neq 0, y \neq 0).$$

Man bestimme alle lokalen Extrema von f.

Aufgabe 2.34 *Bestimmen Sie das absolute Minimum und Maximum von*

$$f(x, y) := 2xy - x + y$$

auf $M := [-2, 2] \times [-2, 2]$.

Aufgabe 2.35 *Man bestimme Minimum und Maximum von*

$$f(x,y) = e^{-(x^2+y^2)}(2x^2+3y^2)$$

auf

$$M = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \leq 4\}.$$

Aufgabe 2.36 *Gegeben sei die Funktion*

$$f : \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto 2x^2 + 2y^2 - 4xy - x^4 - y^4.$$

a) Untersuchen Sie f auf relative Maxima und Minima.
b) Bestimmen Sie die Tangentialebene an die Fläche

$$G = \{(x,y,z) \in \mathbb{R}^3 \mid z = f(x,y)\} \text{ im Punkt } P(1,1,-2).$$

Aufgabe 2.37 *Gegeben sei die Funktion $f :]0,\infty[\times]0,\infty[\to \mathbb{R}$, mit*

$$f(x,y) = xy + \frac{a}{x} + \frac{a}{y},\ a > 0.$$

Untersuchen Sie f auf relative Maxima und Minima.

Aufgabe 2.38 *Für $f : \mathbb{R}^2 \to \mathbb{R}$, definiert durch $f(x,y) = y^2 + \cos x$, bestimme man die lokalen Extremstellen in $\mathbb{R}^2$ und die absoluten Extremstellen in*

$$M := \{(x,y) \in \mathbb{R}^2 \,:\, 0 \leq x \leq \pi/2,\ -1 \leq y \leq 1\}.$$

Aufgabe 2.39 *Gegeben sei die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = x^3 + y^3 - 3x - 12y$. Untersuchen Sie f auf relative Extrema; geben Sie gegebenenfalls die Art der Extrema an.*

Aufgabe 2.40 *Gegeben sei die Funktion*

$$f(x,y) = e^{-x^2-y^2}(x^2+y^2+xy+3).$$

Man bestimme Infimum und Supremum der Bildmenge $f(\mathbb{R}^2)$.

Aufgabe 2.41 *Für $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = 2x^3 - xy^2 + 5x^2 + y^2$ bestimme man die relativen (= lokalen) Extrema.*

Aufgabe 2.42 *Gegeben sei die Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$ *durch*

$$f(x, y) = (x^2 + 2y^2 - 1) \cdot (x^2 + y^2 - 1).$$

a) Skizzieren Sie die Menge $\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$ *und markieren Sie die Bereiche, wo* $f(x, y)$ *positiv, bzw. negativ ist.*
b) Berechnen Sie den Gradienten und die Hesse-Matrix von f.
c) Bestimmen Sie alle lokalen Extrema.

2.6 Die Differentialoperatoren *grad*, *rot*, *div* und Δ

Wir haben die partiellen Ableitungen einer stetig partiell differenzierbaren Funktion $f(x_1, ..., x_n)$ zum Gradientenfeld

$$\mathit{grad} f(\mathbf{x}) = \nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), ..., \frac{\partial f}{\partial x_n}(\mathbf{x})\right)$$

zusammengefasst. Es ist häufig wichtig zu entscheiden, ob ein gegebenes Vektorfeld

$$\mathbf{A}(\mathbf{x}) = (A_1(x_1, ..., x_n), ..., A_n(x_1, ..., x_n))$$

ein Gradientenfeld ist oder nicht:

$$\mathbf{A} = \mathit{grad}(f) \quad ??$$

Gradientenfelder tragen mehr oder weniger schöne Namen (z.B. *konservativ*), ebenso wie die Funktionen f deren Gradienten sie sind ($-f$ heißt *Potential* des Feldes $\mathbf{A}$.) Ist $\mathbf{A}$ stetig differenzierbar, so folgt aus der Symmetrie der zweiten Ableitungen (Satz 2.6) das folgende, wirklich enorm wichtige, Kriterium

Satz 2.25 (Notwendiges Kriterium für Gradientenfelder) *Das Vektorfeld* $\mathbf{A}$ *sei stetig differenzierbar. Ist* $\mathbf{A} = \mathit{grad}(f)$ *ein Gradientenfeld, so gilt*

$$\frac{\partial A_\mu}{\partial x_\nu} = \frac{\partial A_\nu}{\partial x_\mu} \qquad \text{(Symmetriebedingung)}$$

für alle $\mu, \nu = 1, .., n$.

Beweis. In der Tat:

$$\frac{\partial A_\mu}{\partial x_\nu} = \frac{\partial^2 f}{\partial x_\nu \partial x_\mu} = \frac{\partial^2 f}{\partial x_\mu \partial x_\nu} = \frac{\partial A_\nu}{\partial x_\mu}.$$

□

Beispiel 2.24 *Die Funktion* $f = x^2 + y^2$ *auf* $\mathbb{R}^2$ *hat den Gradienten*

$$\mathit{grad}(f) = (2x, 2y).$$

Natürlich ist deswegen die Symmetriebedingung

$$\frac{\partial(2x)}{\partial y} = 0 = \frac{\partial(2y)}{\partial x}$$

erfüllt. Die ist auch für das Feld

$$\mathbf{A} = (y, x)$$

erfüllt ($f = xy$)*, aber z.B. nicht für*

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

Dieses Feld kann also kein Potential besitzen.

In die Symmetriebedingung hat man $\mu, \nu = 1, .., n$ unabhängig voneinander einzusetzen. Natürlich braucht man $\mu = \nu$ nicht auszuprobieren, und wenn die Symmetriebedingung für $\mu \neq \nu$ erfüllt ist, dann gilt sie auch nach Vertauschen von μ und ν. Sie stellt also

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

skalare Bedingungen dar. I.A. ist diese Anzahl viel größer als n, aber in unserem 3-dimensionalen Anschauungsraum ist diese Anzahl genau gleich der Dimension:

$$\frac{1}{2}3 \cdot 2 = 3.$$

So banal das ist, so ist es doch eine der Grundtatsachen des Lebens. Man kann also die drei Gleichungen

$$\frac{\partial A_3}{\partial x_2} = \frac{\partial A_2}{\partial x_3}, \quad \frac{\partial A_1}{\partial x_3} = \frac{\partial A_3}{\partial x_1}, \quad \frac{\partial A_2}{\partial x_1} = \frac{\partial A_1}{\partial x_2}$$

zu einer Vektorgleichung zusammenfassen:

$$rot(\mathbf{A}) := \begin{pmatrix} \frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3} \\ \\ \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} \\ \\ \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \\ \frac{\partial}{\partial x_2} \\ \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} A_1 \\ \\ A_2 \\ \\ A_3 \end{pmatrix} = \mathbf{0}.$$

Jedem Vektorfeld $\mathbf{A}$ im *dreidimensionalen Raum* ordnet man so ein neues Vektorfeld zu, seine *Rotation* $rot(\mathbf{A})$. Und die Symmetriebedingung im Dreidimensionalen schreibt sich

$$\boxed{\mathrm{rot(grad(f))} = 0.}$$

Natürlich hat das Rotationsfeld etwas mit "Drehung" zu tun. Dazu betrachten wir als Beispiel das Geschwindigkeitsfeld bei Drehung um eine Achse mit Winkelgeschwindigkeit ω:

$$\mathbf{A}(x_1, x_2, x_3) = \omega \times \mathbf{x} = \begin{pmatrix} \omega_2 \cdot x_3 - \omega_3 \cdot x_2 \\ \omega_3 \cdot x_1 - \omega_1 \cdot x_3 \\ \omega_1 \cdot x_2 - \omega_2 \cdot x_1 \end{pmatrix},$$

$$rot(\mathbf{A}) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} \omega_2 \cdot x_3 - \omega_3 \cdot x_2 \\ \omega_3 \cdot x_1 - \omega_1 \cdot x_3 \\ \omega_1 \cdot x_2 - \omega_2 \cdot x_1 \end{pmatrix} = \begin{pmatrix} 2\omega_1 \\ 2\omega_2 \\ 2\omega_3 \end{pmatrix} = 2 \cdot \omega.$$

Bis jetzt haben wir nur die notwendige Bedingung dafür, dass $A(\mathbf{x})$ ein Gradientenfeld ist, nämlich die Symmetriebedingung, analysiert. Aber lokal, d.h., auf einer offenen Vollkugel im $\mathbb{R}^n$ ist diese Bedingung auch hinreichend.

Satz 2.26 *Das Vektorfeld $A(\mathbf{x})$ sei stetig differenzierbar auf der offfenen Kugel*

$$U := \{\mathbf{x} \in \mathbb{R}^n : \| \mathbf{x} \| < r\},$$

oder allgemeiner, auf einer offenen konvexen Menge $U \subset \mathbb{R}^n$. Dann gibt es eine differenzierbare Funktion $f : U \to \mathbb{R}$ mit

$$A = gradf.$$

Für den Beweis dieses Satzes brauchen wir eine Technik, die ich bis jetzt noch nicht behandelt habe: die Differentiation unter dem Integral.

Satz 2.27 *Die Funktion $f(x_1, ..., x_n)$ sei stetig auf dem Quader*

$$Q = \{\mathbf{x} \in \mathbb{R}^n : \, a_1 \leq x_1 \leq b_1, ..., a_n \leq x_n \leq b_n\} \subset \mathbb{R}^n.$$

Dann ist für

$$\mathbf{x}' = (x_2, .., x_n) \in Q' = \{\mathbf{x} \in \mathbb{R}^{n-1} : \, a_2 \leq x_2 \leq b_2, ..., a_n \leq x_n \leq b_n\} \subset \mathbb{R}^n$$

die Funktion

$$F(x_2, ..., x_n) := \int_{a_1}^{b_1} f(x_1, x_2, ..., x_n) dx_1$$

definiert.

a) (Stetige Abhängigkeit des Integrals von einem Parameter). Aus der Stetigkeit von f folgt die Stetigkeit von F.

b) (Differentiation unter dem Integral) Ist f auf der Menge

$$\{\mathbf{x} \in \mathbb{R}^n : \, a_1 \leq x_1 \leq b_1, a_2 < x_2 < b_2, ..., a_n < x_n < b_n\}$$

stetig partiell differenzierbar nach $x_2, ..., x_n$. Dann ist auch die Funktion $F(x_2, ..., x_n)$ partiell differenzierbar nach $x_2, ..., x_n$ mit der (nach a) stetigen) partiellen Ableitung

$$\frac{\partial F}{\partial x_\nu}(x_2, ..., x_n) = \int_{a_1}^{b_1} \frac{\partial f}{\partial x_\nu}(\mathbf{x}) dx_1.$$

Beweis a). Wir zeigen die $\epsilon - \delta$-Stetigkeit von $F(\mathbf{x}')$ in einem festen Punkt $\mathbf{x}'_0$. Dazu erinnern wir uns an Satz 1.12: Die Funktion f ist gleichmäßig stetig auf Q. Zu $\epsilon > 0$ gibt es also ein $\delta(\epsilon) > 0$ mit

$$\mathbf{x}, \mathbf{y} \in Q, \parallel \mathbf{x} - \mathbf{y} \parallel < \delta(\epsilon) \quad \Rightarrow \quad |f(\mathbf{x}) - f(\mathbf{y})| < \epsilon.$$

Falls $\parallel \mathbf{x}' - \mathbf{x}'_0 \parallel < \delta(\epsilon)$, dann folgt daraus für alle $x_1, a_1 \leq x_1 \leq b_1$

$$|f(x_1, \mathbf{x}') - f(x_1, \mathbf{x}'_0)| < \epsilon$$

und

$$\begin{aligned}
|F(\mathbf{x}') - F(\mathbf{x})| &= \left| \int_{a_1}^{b_1} f(x_1, \mathbf{x}') dx_1 - \int_{a_1}^{b_1} f(x_1, \mathbf{x}'_0) dx_1 \right| \\
&= \left| \int_{a_1}^{b_1} (f(x_1, \mathbf{x}') - f(x_1, \mathbf{x}'_0)) dx_1 \right| \\
&\leq | \int_{a_1}^{b_1} |f(x_1, \mathbf{x}') - f(x_1, \mathbf{x}'_0)| dx_1 | \\
&\leq \int_{a_1}^{b_1} \epsilon dx_1 \\
&= (b_1 - a_1) \cdot \epsilon.
\end{aligned}$$

Hatten wir hier $\delta(\epsilon/(b_1 - a_1))$ an Stelle von $\delta(\epsilon)$ genommen, dann wäre $|F(\mathbf{x}') - F(\mathbf{x}'_0)| < \epsilon$ herausgekommen.

b) Es genügt den Fall $n = 2$ und die stetige Differenzierbarkeit nach x_2 zu betrachten. Dazu fixieren wir ein x_2 und ein $r > 0$ mit

$$a_2 < x_2 - r < x_2 + r < b_2.$$

Auf der kompakten Menge

$$[a_1, b_1] \times [x_2 - r, x_2 + r]$$

ist $\partial f / \partial x_2$ stetig und nach Satz 1.12 sogar gleichmäßig stetig. Zu jedem $\epsilon > 0$ gibt es also ein $\delta(\epsilon) \leq r$ derart, dass gilt

$$a_1 \leq x_1 \leq b_1, \, |x'_2 - x_2| < \delta(\epsilon) \quad \Rightarrow \quad \left| \frac{\partial f}{\partial x_2}(x_1, x'_2) - \frac{\partial f}{\partial x_2}(x_1, x_2) \right| < \epsilon.$$

Wir wenden den MWS der Differentialrechnung an auf die in x_2 stetig differenzierbare Funktion f und finden für alle $a_1 \leq x_1 \leq b_1$, sowie $x'_2 \neq x_2$ mit $|x'_2 - x_2| < \delta(\epsilon)$

$$\left| \frac{f(x_1, x'_2) - f(x_1, x_2)}{x'_2 - x_2} - \frac{\partial f}{\partial x_2}(x_1, x_2) \right| = \left| \frac{\partial f}{\partial x_2}(x_1, x_2 + \theta(x'_2 - x_2)) - \frac{\partial f}{\partial x_2}(x_1, x_2) \right| < \epsilon.$$

Daraus folgt

$$\left| \frac{F(x'_2) - F(x_2)}{x'_2 - x_2} - \int_{a_1}^{b_1} \frac{\partial f}{\partial x_2}(x_1, x_2) dx_1 \right| \leq \int_{a_1}^{b_1} \left| \frac{f(x_1, x'_2) - f(x_1, x_2)}{x'_2 - x_2} - \frac{\partial f}{\partial x_2}(x_1, x_2) dx_1 \right| < \epsilon \cdot (b_1 - a_1).$$

Das bedeutet

$$\lim_{\substack{x'_2 \to x_2 \\ x'_2 \neq x_2}} \frac{F(x'_2) - F(x_2)}{x'_2 - x_2} = \int_{a_1}^{b_1} \frac{\partial f}{\partial x_2}(x_1, x_2) dx_1.$$

$\square$

Jetzt zum Beweis von Satz 2.26. Um Schreibaufwand zu sparen nehmen wir o.B.d.A. an, der Nullpunkt $\mathbf{0}$ gehöre zu U. Weil U konvex vorausgesetzt ist, gehört dann mit jedem festen Punkt $\mathbf{x} \in U$ auch die ganze Strecke

$$\{t \cdot \mathbf{x} : \ 0 \leq t \leq 1\}$$

zu U. Wenn f auf U stetig partiell differenzierbar ist, folgt für $0 \leq t \leq 1$ mit der Formel für die Richtungsableitung

$$\frac{d}{dt} f(t \cdot \mathbf{x}) = (\nabla f(t \cdot \mathbf{x}).\mathbf{x}).$$

Wenn außerdem noch $f(\mathbf{0}) = 0$ gilt, dann ist

$$f(\mathbf{x}) = \int_0^1 \frac{d}{dt} f(t \cdot \mathbf{x}) dt = \int_0^1 (\nabla f(t \cdot \mathbf{x}).\mathbf{x}) dt.$$

Wir suchen eine Funktion $f : U \to \mathbb{R}$ mit $\nabla f = A$. Wir können f durch $f - f(\mathbf{0})$ ersetzen und $f(\mathbf{0}) = 0$ annehmen. Wenn f existiert, muss es durch die Formel

$$f(\mathbf{x}) := \int_0^1 (A(t \cdot \mathbf{x}).\mathbf{x}) dt = \sum_{\nu=1}^n \int_0^1 A_\nu(t \cdot \mathbf{x}) \cdot x_\nu dt$$

gegeben sein. Umgekehrt definiert diese Formel eine Funktion f auf U, weil die A_ν stetig sind. Zu zeigen ist, dass die so definierte Funktion f partiell differenzierbar ist mit

$$\frac{\partial f}{\partial x_\nu}(\mathbf{x}) = A_\nu(\mathbf{x}) \quad \text{für} \quad \nu = 1, ..., n.$$

Es genügt, diesen Nachweis für $\nu = 1$ zu führen.

Weil U offen ist, gibt es ein $r > 0$ so, dass der ganze Würfel

$$Q := \{\mathbf{x}' \in \mathbb{R}^n : |x'_1 - x_1| < r, ..., |x'_n - x_n| < r\}$$

noch zu U gehört. Die Funktion $f(t \cdot \mathbf{x})$ ist stetig auf dem $n+1$-dimensionalen Quader

$$\{(t, \mathbf{x}) \in \mathbb{R}^{n+1} : \ 0 \leq t \leq 1, \mathbf{x} \in Q\}.$$

Nach Satz 2.27 b) sind die Funktionen

$$\int_0^1 A_\nu(t \cdot \mathbf{x}) dt$$

auf dieser Menge stetig differenzierbar bezüglich $x_1, ..., x_n$ mit

$$\frac{\partial}{\partial x_\mu} \int_0^1 A_\nu(t \cdot \mathbf{x}) dt = \int_0^1 \frac{\partial A_\nu}{\partial x_\mu}(t \cdot \mathbf{x}) dt.$$

Damit folgt

$$\begin{aligned} \frac{\partial f}{\partial x_1}(\mathbf{x}) &= \sum_{\nu=1}^n \int_0^1 \frac{\partial}{\partial x_1}(A_\nu(t \cdot \mathbf{x}) \cdot x_\nu) dt \\ &= \int_0^1 A_1(t \cdot \mathbf{x}) dt + \sum_{\nu=1}^n \int_0^1 \frac{\partial A_\nu}{\partial x_1}(t \cdot \mathbf{x}) \cdot t x_\nu dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 A_1(t\cdot\mathbf{x})dt + \sum_{\nu=1}^n \int_0^1 \frac{\partial A_1}{\partial x_\nu}(t\cdot\mathbf{x})\cdot tx_\nu dt \qquad \text{Symmetrie-Bedingung)}\\
&= \int_0^1 \frac{d}{dt}(t\cdot A_1(t\cdot\mathbf{x}))dt\\
&= t\cdot A_1(t\cdot\mathbf{x})|_{t=0}^{t=1}\\
&= A_1(\mathbf{x}).
\end{aligned}
$$

□

Ein Vektorfeld $\mathbf{A}$ auf dem $\mathbb{R}^n$ hat eine $n\times n$ Funktionalmatrix

$$\frac{\partial A_\mu}{\partial x_\nu}.$$

Deren *Spur*

$$div(\mathbf{A}) = \sum_{\nu=1}^n \frac{\partial A_\nu}{\partial x_\nu}$$

heißt *Divergenz* des Vektorfeldes. Im Unterschied zur Rotation ist die Divergenz

- ein Skalar, kein Vektor,
- in allen Dimensionen n sinnvoll zu definieren.

Beispiel 2.25 *Die Divergenz des Ortsvektorfeldes* $\mathbf{A}(\mathbf{x}) = \mathbf{x}$ *ist*

$$div(\mathbf{x}) = \sum_{\nu=1}^n \frac{\partial x_\nu}{\partial x_\nu} = n.$$

Und wenn man ein Vektorfeld $\mathbf{A}(\mathbf{x})$ *mit einer skalaren Funktion* $f(\mathbf{x})$ *multipliziert, bekommt man*

$$div(f\cdot\mathbf{A}) = \sum_{\nu=1}^n \frac{\partial(f\cdot A_\nu)}{\partial x_\nu} = \sum_{\nu=1}^n \frac{\partial f}{\partial x_\nu}A_\nu + \sum_{\nu=1}^n f\frac{\partial A_\nu}{\partial x_\nu} = (grad(f),\mathbf{A}) + f\cdot div(\mathbf{A}).$$

Man kann die Divergenz eines Rotationsfeldes bilden (nicht die Rotation einer Divergenz, da letztere ein Skalar ist):

$$
\begin{aligned}
div(rot(\mathbf{A})) &= \frac{\partial}{\partial x_1}(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}) + \frac{\partial}{\partial x_2}(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}) + \frac{\partial}{\partial x_3}(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2})\\
&= \frac{\partial^2 A_1}{\partial x_2\partial x_3} + \frac{\partial^2 A_2}{\partial x_3\partial x_1} + \frac{\partial^2 A_3}{\partial x_1\partial x_2} - \frac{\partial^2 A_1}{\partial x_3\partial x_2} - \frac{\partial^2 A_2}{\partial x_1\partial x_3} - \frac{\partial^2 A_3}{\partial x_2\partial x_1}\\
&= 0,
\end{aligned}
$$

wieder wegen der Symmetrie der zweiten Ableitungen! Wir haben also die Formel

$$\boxed{\text{div(rot}(\mathbf{A}))\text{=0}}$$

bewiesen. Ähnlich wie $rot(grad(f)) = 0$ kann man die Formel $div(rot(\mathbf{A})) = 0$ auch als notwendige Bedingung dafür auffassen, dass ein Vektorfeld $\mathbf{B}$ ein Rotationsfeld ist: Gibt es ein Feld $\mathbf{A}$ mit $\mathbf{B} = rot(\mathbf{A})$, so ist notwendigerweise $div(\mathbf{B}) = 0$.

Es gibt noch eine Möglichkeit, zwei der drei Differentialoperatoren *grad*, *rot* und *div* sinnvoll zu kombinieren: den *Laplace-Operator* angewandt auf eine Funktion f

$$\Delta(f) = div(grad(f)) = \sum_{\nu=1}^{n} \frac{\partial^2 f}{\partial x_\nu^2}.$$

Dies ist wieder in jeder beliebigen Dimension n möglich, liefert aber von der Dimension abhängige Ergebnisse. Betrachten wir etwa die Funktion

$$f(x_1, ..., x_n) = \| \mathbf{x} \|^m = (\sum_{\nu=1}^{n} x_\nu^2)^{m/2}.$$

Dann ist

$$\begin{array}{rcl} grad(\| \mathbf{x} \|) &=& \frac{1}{\| \mathbf{x} \|} \cdot \mathbf{x}, \\ grad(\| \mathbf{x} \|^m) &=& m \| \mathbf{x} \|^{m-1} \cdot grad(\| \mathbf{x} \|) \\ &=& m \| \mathbf{x} \|^{m-2} \cdot \mathbf{x}, \\ \Delta(f) &=& m \cdot div(\| \mathbf{x} \|^{m-2} \cdot \mathbf{x}) \\ &=& m \cdot [(grad(\| \mathbf{x} \|^{m-2}), \mathbf{x}) + \| \mathbf{x} \|^{m-2} \, div(\mathbf{x})] \\ &=& m \cdot [(m-2) \| \mathbf{x} \|^{m-4} \cdot (\mathbf{x}, \mathbf{x}) + n \cdot \| \mathbf{x} \|^{m-2}] \\ &=& m \cdot (n+m-2) \cdot \| \mathbf{x} \|^{m-2} \\ &=& 0 \\ &\Leftrightarrow& m = 2 - n. \end{array}$$

Funktionen f mit $\Delta(f) = 0$ heißen *harmonisch*. Wir haben ausgerechnet: Auf $\mathbb{R}^n$ (ohne Nullpunkt) ist

$$r^{2-n} = \| \mathbf{x} \|^{2-n}$$

harmonisch. Für $n = 1$ ist dies die kaum erwähnenswerte Tatsache, dass die zweite Ableitung der Funktion $|x|$ verschwindet ($x \neq 0$). Für $n = 2$ die noch uninteressantere, entsprechende Aussage für die konstanten Funktionen. Aber für $n = 3$ etwa ist $1/r$ harmonisch.

Allerdings ist die Dimension 2 dadurch nicht sehr benachteiligt: Für die Funktion $f(x, y) = ln(x^2 + y^2)$ ist

$$\begin{array}{rcl} \Delta(f) &=& \frac{\partial}{\partial x}\left(\frac{2x}{x^2+y^2}\right) + \frac{\partial}{\partial y}\left(\frac{2y}{x^2+y^2}\right) \\ &=& \frac{4}{x^2+y^2} - \frac{4x^2}{(x^2+y^2)^2} - \frac{4y^2}{(x^2+y^2)^2} \\ &=& 0, \end{array}$$

also ist $ln(r) = \frac{1}{2} ln(x^2 + y^2)$ in Dimension 2 harmonisch.

Die Produktregel der Differentiation führt nach dem Schema

$$grad(f \cdot g) = g \cdot grad(f) + f \cdot grad(g)$$

oder

$$div(f \cdot \mathbf{A}) = (grad(f), \mathbf{A}) + f \cdot div(\mathbf{A})$$

in Verbindung mit den drei Operationen *grad*, *rot* und *div* zu reichhaltigem Formelmaterial. Um die Merkbarkeit dieser Regeln zu erhöhen kann man dem Operator *Nabla*

$$\nabla = \begin{pmatrix} \dfrac{\partial}{\partial x} \\ \\ \dfrac{\partial}{\partial y} \\ \\ \dfrac{\partial}{\partial z} \end{pmatrix}$$

eine Art von vektoriellem Charakter zuerkennen. Einerseits ist er ein Vektorfeld auf dem $\mathbb{R}^3$, andererseits auch ein Differentialoperator. Mit etwas Fingerspitzengefühl wird

- $grad(f)$ das Produkt ∇f (von rechts) dieses Operators mit der Funktion f,
- $rot(\mathbf{A})$ das Kreuzprodukt $\nabla \times \mathbf{A}$,
- $div(\mathbf{A})$ das Skalarprodukt $(\nabla, \mathbf{A})$.

Die bekannten Eigenschaften des Skalar- und Kreuzproduktes aus der linearen Algebra erleichtern Herleitung und Erinnerung an die erwähnten Formeln.

Aufgabe 2.43 *Bestimmen Sie Divergenz und Rotation der Vektorfelder*

$$\mathbf{A} = \begin{pmatrix} y \\ z \\ x \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} z \\ x \\ y \end{pmatrix}.$$

Aufgabe 2.44 *Bestimmen Sie $rot(\mathbf{a})$ und $rot(rot(\mathbf{a}))$ für das Vektorfeld*

$$\mathbf{A} = \begin{pmatrix} yz^2 \\ zx^2 \\ xy^2 \end{pmatrix}.$$

Aufgabe 2.45 *Zeigen Sie für differenzierbare Funktionen f, g und Vektorfelder $\mathbf{A}, \mathbf{B}$ auf dem $\mathbb{R}^3$*

$$\begin{array}{lll}
1) & grad(f \cdot g) & = & g \cdot grad(f) + f \cdot grad(g) \\
2) & div(f \cdot \mathbf{A}) & = & (grad(f).\mathbf{A}) + f \cdot div(\mathbf{A}) \\
3) & \Delta(f \cdot g) & = & g \cdot \Delta(f) + 2(grad(f).grad(g)) + f \cdot \Delta(g) \\
4) & rot(f \cdot \mathbf{A}) & = & grad(f) \times \mathbf{A} + f \cdot rot(\mathbf{A}) \\
5) & rot(f \cdot grad(g)) & = & grad(f) \times grad(g) \\
6) & div(\mathbf{A} \times \mathbf{B}) & = & (\mathbf{B}.rot(\mathbf{A})) - (\mathbf{A}.rot(\mathbf{B})) \\
7) & div(grad(f) \times grad(g)) & = & 0
\end{array}$$

Aufgabe 2.46 *Es seien* $a \in \mathbb{R}$ *und* $\mathbf{x} \neq \mathbf{0}$.
a) Bestimmen Sie den Gradient der Funktion $(x^2 + y^2 + z^2)^a$.
b) Bestimmen Sie für $r := \sqrt{x^2 + y^2 + z^2}$ *den Gradient der Funktion* r^a.
c) Bestimmen Sie Divergenz und Rotation des Vektorfeldes $\hat{\mathbf{x}} := \frac{1}{r}\mathbf{x}$.

Aufgabe 2.47 *Es sei* $f(x, y)$ *eine genügend oft differenzierbare Funktion auf* $\mathbb{R}^2$ *und*

$$f(r, \varphi) = f(x(r, \varphi), y(r, \varphi))$$

dieselbe Funktion in Polarkoordinaten.
a) Zeigen Sie: $\Delta f = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2}$.
b) Bestimmen Sie alle harmonischen Funktionen f *auf* $\mathbb{R}^2 \backslash \mathbf{0}$, *die nur von* r *und nicht von* φ *abhängen.*

Aufgabe 2.48 *Es sei* $f : \mathbb{R}^2 \setminus \{\mathbf{0}\} \to \mathbb{R}$ *genügend oft differenzierbar. Zeigen Sie:* $f \cdot \mathbf{x}$ *ist genau dann ein Gradientenfeld, wenn* $f = f(r)$ *ist.*

Aufgabe 2.49 *Auf* $\mathbb{R}^2 \setminus \{(0, 0)\}$ *sei das Vektorfeld*

$$A(x, y) = \frac{1}{x^2 + y^2} \cdot (-y, x)$$

gegeben. Zeigen Sie:
a) A *die erfüllt Symmetrie-Bedingung 2.25.*
b) Ist U *eine Kreisscheibe, auf der eine Funktion* f *mit* $A = \nabla f$ *existiert, dann ist* $f = \varphi + const$, *mit* $x = rcos(\varphi), y = rsin(\varphi)$.
c) Es gibt keine differenzierbare Funktion $f : \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}$ *mit* $A = \nabla f$.

Aufgabe 2.50 *Bestimmen Sie eine differenzierbare Funktion* $f : \mathbb{R}^2 \to \mathbb{R}$ *mit*
a) $\nabla f = (3x^2 - y, 3y^2 - x)$,
b) $\nabla f = e^x \cdot (cos(y), -sin(y))$.

3 Gewöhnliche Differentialgleichungen

Als ich mit dem Computer die Bibliothek des Mathematischen Instituts nach Büchern über Differentialgleichungen durchsuchte, lieferte mir der über 200 Buchtitel. Die Mehrzahl waren Bücher über partielle Differentialgleichungen, haben mit dem Inhalt dieser Vorlesung wenig zu tun. Und von den Büchern über gewöhnliche Differentialgleichungen waren die allermeisten viel zu speziell. Hier möchte ich einige Bücher angeben, die vielleicht für Sie brauchbar sind.

Zunächst einmal steht fast alles, was ich hier behandeln werde, im zweiten Teil des Buches

O. Forster: Analysis 2, rororo vieweg, soundsovielste Auflage.

Ich selbst habe gewönhnliche Differentialgleichungen nach

F. Erwe: Gewöhnliche Differentialgleichungen, BI Mannheim, 1964

gelernt. Der Klassiker auf diesem Gebiet ist

F. Kamke: Differentialgleichungen reeller Funktionen, AVG Leipzig 1930 (und spätere Wiederauflagen mit ähnlichem Titel).

Schließlich ist auch

E.L. Ince: Die Integration gewöhnlicher Differentialgleichungen, BI Mannheim, 1956

ein sehr nützliches und kompaktes Buch.

3.1 Einführung

Differentialgleichungen sind die wichtigsten Anwendungen der Mathematik.

Beispiel 3.1 (Newton) *Die Mutter aller Differentialgleichungen ist die Newtonsche Gleichung*

$$\ddot{y} = -g$$

für den freien Fall. Dabei ist $y(t)$ die Höhe eines Massenpunktes zur Zeit t, der sich im freien Fall, nur unter Einfluss der Erdanziehungskraft, nach unten bewegt. Auf ihn wirkt die Erdanziehungskraft $K = -m \cdot g$, wo m seine Masse, und die Beschleunigung g ungefähr $10 m/sec^2$ ist. Newton hat herausgefunden, dass die Kraft, welche auf einen Körper wirkt, diesen Körper beschleunigt. *D.h., sie wirkt direkt auf die* zweite *Ableitung seiner Höhe $y(t)$. Quantitativ ist Newtons Formel*

$$K = m \cdot \ddot{y},$$

die Kaft ist Masse $\times$ Beschleunigung. Beim freien Fall kürzt sich die Masse m heraus, und man erhält dann die obige Gleichung.

Aus der Schule kennt man die Lösung

$$y(t) = y_0 + v_0 \cdot t - \frac{g}{2} \cdot t^2, \qquad t = \mathit{Zeit}.$$

Dabei ist y_0 die Anfangshöhe und v_0 die Anfangsgeschwindigkeit. Es ist leicht nachzurechnen, dass diese Funktion $y(t)$ eine Lösung der Gleichung ist. Und man überzeugt sich ebenso leicht davon, dass auch umgekehrt jede Lösung der Newtonschen Gleichung diese Form hat: Aus $\ddot{y} = -g$ folgt durch *einmalige Integration* $\dot{y} = -g \cdot t + v_0$ *mit einer Konstanten* v_0 *und durch eine zweite Integration* $y(t) = -(g/2) \cdot t^2 + v_0 \cdot t + y_0$ *mit einer weiteren Konstanten* y_0*.*

Die Newtonsche Differentialgleichung ist ganz besonders einfach: Außer der zweiten Ableitung kommt in der Gleichung von der gesuchten Funktion nichts vor. Deswegen behandelt man sie ja auch schon in der Schule. Das ist nicht besonders typisch für Differentialgleichungen: Im Allgemeinen verknüpft eine Differentialgleichung eine Funktion mit ihren Ableitungen. Das passiert z.B., wenn man in der Gleichung des fallenden Körpers den Luftwiderstand berücksichtigt. In erster Näherung ist der eine Art Reibung und damit direkt proportional zur Geschwindigkeit. Berücksichtigt man ihn, so nimmt die Newtonsche Gleichung die Form

$$\ddot{y} = -\mu \cdot \dot{y} - g$$

an. Da kommt zwar die erste Ableitung der Funktion $y(t)$*, aber immer noch nicht die Funktion* $y(t)$ *selbst vor. Das würde aber notwendig, wenn man einen Fall durch die gesamte Erdatmosphäre (etwa einer zurückkehrenden Rakete) verfolgt. In erster Näherung wäre die Luftdichte, und damit der Luftwiderstand proportional zur Nähe zu der Erde, etwa* $\mu = \mu(y) = \mu \cdot (y_0 - y(t))$*. Dann bekommen wir die schon recht komplizierte Gleichung*

$$\ddot{y} = -\mu \cdot (y_0 - y(t)) \cdot \dot{y} - g,$$

in der die Funktion $y(t)$ *und ihre beiden ersten Ableitungen vorkommen.*

Beispiel 3.2 (Oszillator) *Ein wesentlich einfacheres Beispiel ist die Gleichung*

$$\ddot{y} = -\omega^2 \cdot y(t)$$

des harmonischen Oszillators. In der Schule lernt man dafür die Lösungen

$$y(t) = c_1 \cdot sin(\omega t) + c_2 \cdot cos(\omega t), \quad c_1, c_2 \in \mathbb{R},$$

kennen. Allerdings ist es hier schon wesentlich schwerer, sich davon zu überzeugen, dass dies die einzigen Lösungen sind. Wenn man den Schwingungsvorgang auch noch gedämpft annimmt, kommt man auf eine Differentialgleichung der Form

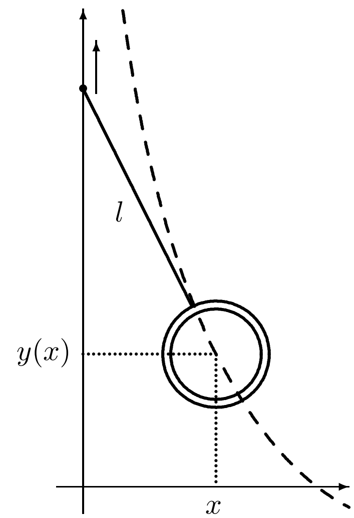
$$\ddot{y} - \mu \cdot \dot{y} + \omega^2 \cdot y = 0.$$

Hier kommt wieder die gesuchte Funktion $y(t)$ *mit ihren beiden ersten Ableitungen vor.*

Dise kurze Diskussion zeigt schon einige wesentliche Aspekte der Differentialgleichungen:

- Eine Differentialgleichung ist eine Gleichung, welche die Werte einer Funktion mit ihren Ableitungen verknüpft.
- Je realistischer die Differentialgleichung einen Vorgang beschreiben soll, desto komplizierter ist sie.
- Man kann den Ablauf eines Vorgangs (z.B. einer Bewegung) vorhersagen, wenn man die betreffende Differentialgleichung lösen kann.
- Kennt man alle Lösungen, so kennt man alle möglichen Bewegungsabläufe, welche durch diese Differentialgleichung beschrieben werden.

Newton schuf die Grundlagen der theoretischen Mechanik, mit der man die Bewegung von Körpern unter Einfluss von Kräften verstehen kann. Das war einer der beiden wichtigsten Impulse zur Entwicklung der Theorie der Differentialgleichungen. Er hing ganz deutlich mit den Anwendungen zusammen. Der andere, davon nicht ganz klar zu trennende Impuls kam aus der reinen Mathematik: Man wollte Kurven (Funktionsgraphen) beschreiben, über deren Richtung man etwas wusste. Das typische Beispiel dafür ist die Kurve, auf der sich eine Taschenuhr bewegt, wenn man sie an ihrer Kette über den Tisch zieht, und zwar immer in dieselbe Richtung. Diese Richtung soll natürlich nicht direkt von der Uhr wegzeigen, sondern etwa längs der y-Achse, von der die Uhr den Abstand x hat. Die Situation ist dann ungefähr so, wie nebenstehend gezeichnet.



Die Differentialgleichung, welche die Funktion $y(x)$ und damit die Bahn der Taschenuhr beschreibt, erhält man durch folgende Überlegung: Wir suchen erst einmal die Gleichung für die Gerade, auf der die (straff gespannte) Kette liegt. Sie ist die Tangente an die gesuchte Kurve. Um nicht mit den Koordinaten x, y durcheinander zu kommen, nehmen wir für die Tangentengleichung Koordinaten ξ, η. Der Anfangsvektor der Geraden ist $(x, y(x))$ und ihre Steigung $y'(x)$. Dann ist die Tangentengleichung also

$$\eta = y(x) + y'(x) \cdot (\xi - x).$$

Die Tangente schneidet die y-Achse dort, wo

$$\xi = 0, \qquad \eta = y(x) - x \cdot y'(x), \qquad \eta - y(x) = -x \cdot y'(x)$$

ist. Nach Pythagoras ist die Länge l der Uhrkette

$$l = \sqrt{x^2 + (\eta - y)^2} = \sqrt{x^2 + x^2 \cdot (y')^2} = \pm x \cdot \sqrt{1 + (y')^2},$$

und somit

$$1 + (y')^2 = \left(\frac{l}{x}\right)^2, \qquad y' = \pm\sqrt{\frac{l^2}{x^2} - 1} = \pm\frac{\sqrt{l^2 - x^2}}{x}.$$

Dies ist die gesuchte Differentialgleichung. Genauer gesagt, sind dies zwei Differentialgleichungen: Eine, wenn man in negativer y-Richtung zieht (Plus-zeichen), und eine andre (Minus-Zeichen), wenn man in positiver y-Richtung zieht. Wir wollen uns auf letzteres festlegen, sodass wir also

$$y' = -\frac{\sqrt{l^2 - x^2}}{x}$$

lösen müssen.

Man kann diese Gleichung lösen, wenn man die Funktion $\sqrt{l^2 - x^2}/x$ integrieren kann. Eine Stammfunktion ist

$$\sqrt{l^2 - x^2} - l \cdot ln\left|\frac{l + \sqrt{l^2 - x^2}}{x}\right|.$$

Als Lösungen findet man damit also die Funktionen

$$y = c - \sqrt{l^2 - x^2} + l \cdot ln \left| \frac{l + \sqrt{l^2 - x^2}}{x} \right|, \quad c \in \mathbb{R}.$$

Die Integrationskonstante c macht Sinn, man kann ja mit dem Ziehen der Taschenuhr bei einem beliebigen Wert von y anfangen, dieser y-Wert legt die Konstante c fest.

Diese Gleichung ist typisch für eine Reihe von Gleichungen, die mit der Tangente

$$T := \{(\xi, \eta) \in \mathbb{R}^2 : \ \eta = y + y' \cdot (\xi - x)\}$$

an den Funktionsgraphen $y = y(x)$ und der Normale

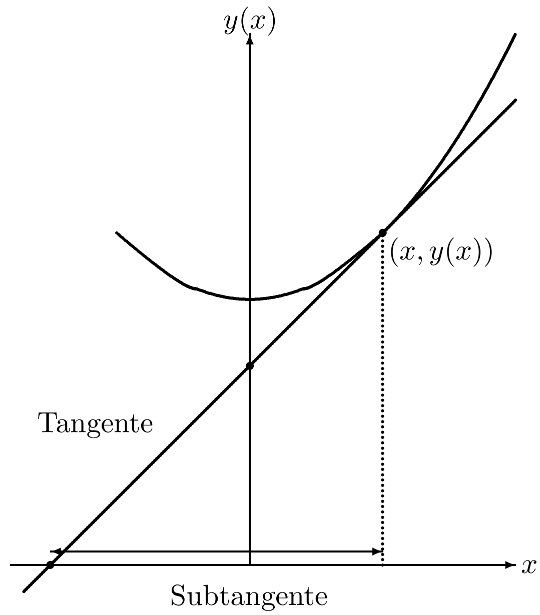
$$N := \{(\xi, \eta) \in \mathbb{R}^2 : \ \eta = y - \frac{1}{y'} \cdot (\xi - x)\}$$

zusammenhängen. Wir stellen einige davon systematisch zusammen. Dazu berechnen wir

- den *Schnittpunkt* $(\xi_T, 0)$ *der Tangente mit der x-Achse*, wobei $\xi_T = x - \frac{y}{y'}$,
- den *Schnittpunkt* $(0, \eta_T)$ *der Tangente mit der y-Achse*, wobei $\eta_T = y - xy'$,
- den *Schnittpunkt* $(\xi_N, 0)$ *der Normale mit der x-Achse*, wobei $\xi_N = x + yy'$,
- den *Schnittpunkt* $(0, \eta_N)$ *der Normale mit der y-Achse*, wobei $\eta_N = y + \frac{x}{y'}$.

Damit berechnen wir

	Tangente	Normale
	$\parallel (x,y) - (\xi_T, 0) \parallel = \left\lvert \frac{y}{y'} \sqrt{1 + (y')^2} \right\rvert$	$\parallel (x,y) - (\xi_N, 0) \parallel = \left\lvert y \cdot \sqrt{1 + (y')^2} \right\rvert$
	$\parallel (x,y) - (0, \eta_T) \parallel = \left\lvert x \cdot \sqrt{1 + (y')^2} \right\rvert$	$\parallel (x,y) - (0, \eta_N) \parallel = \left\lvert \frac{x}{y'} \sqrt{1 + (y')^2} \right\rvert$
-nabschnitt	$\parallel (\xi_T, 0) - (0, \eta_T) \parallel = \left\lvert \frac{xy' - y}{y'} \sqrt{1 + (y')^2} \right\rvert$	$\parallel (\xi_N, 0) - (0, \eta_N) \parallel = \left\lvert \frac{x + yy'}{y'} \sqrt{1 + (y')^2} \right\rvert$
Sub -	$\lvert x - \xi_T \rvert = \left\lvert \frac{y}{y'} \right\rvert$	$\lvert x - \xi_N \rvert = \lvert yy' \rvert$
Ursprungsabstand	$\left\lvert \frac{y - xy'}{\sqrt{1 + (y')^2}} \right\rvert$	$\left\lvert \frac{x + yy'}{\sqrt{1 + (y')^2}} \right\rvert$



Setzt man diese Größen konstant, oder schreibt für sie bestimmte, von x oder y abhängende Werte vor, so erhält man eine Differentialgleichung.

Beispiel 3.3 *Am einfachsten ist die Bedingung für konstante Subtangente (das Vorzeichen wollen wir einmal unterdrücken):*

$$\frac{y}{y'} = const, \qquad y' = c \cdot y.$$

Eine Lösung ist jede Funktion $k \cdot e^{cx}$, $0 \neq k \in \mathbb{R}$.

Jetzt möchte ich einige *Definitionen* zusammenstellen.

Definition 3.1 *Eine* gewöhnliche Differentialgleichung k-ter Ordnung *ist eine Gleichung*

$$F(x, y, y', y'', ..., y^{(k)}) = 0,$$

wo F eine stetige Funktion von $k+2$ Variablen ist. (Die meisten Differentialgleichungen der Physik sind Gleichungen zweiter Ordnung. Die Gleichungen, welche man erhält, wenn man eine der oben berechneten Größen gleich einer festen Konstanten setzt, sind Differentialgleichungen erster Ordnung.) Meistens setzt man die Funktion F stetig voraus. Zu F gehört dann natürlich auch ein Definitionsbereich $G \subset \mathbb{R}^{k+2}$, und wir haben $F : G \to \mathbb{R}$.

Eine Lösung *obiger Differentialgleichung auf einem Intervall* $]\alpha, \beta[\subset \mathbb{R}$ *ist eine k-mal differenzierbare Funktion $f :]\alpha, \beta[\to \mathbb{R}$ derart, dass für alle $x \in]\alpha, \beta[$ gilt:*

(i) der Vektor $(x, f(x), f'(x), ..., f^{(k)}(x))$ gehört zu G,

(ii) $F(x, f(x), f'(x), ..., f^{(k)}(x)) = 0$.

Definition 3.2 *Eine Differentialgleichung heißt* explizit*, wenn sie nach der höchsten vorkommenden Ableitung aufgelöst ist, also eine Form*

$$y^{(k)} = F(x, y, y', ..., y^{(k-1)})$$

hat.

Eine explizite Differentialgleichung erster Ordnung sieht also etwa so aus:

$$y' = F(x, y).$$

Definition 3.3 *Eine Differentialgleichung heißt* linear*, wenn F in $y, y', ..., y^{(k)}$ linear ist, also, wenn sie die Form*

$$a_k(x) \cdot y^{(k)} + ... + a_1(x) \cdot y' + a_0(x) \cdot y = b(x)$$

hat. Sie heißt homogen *linear, wenn $b(x) \equiv 0$ ist, sonst* inhomogen *linear.*

So ist z.B.

$$y' = a(x) \cdot y$$

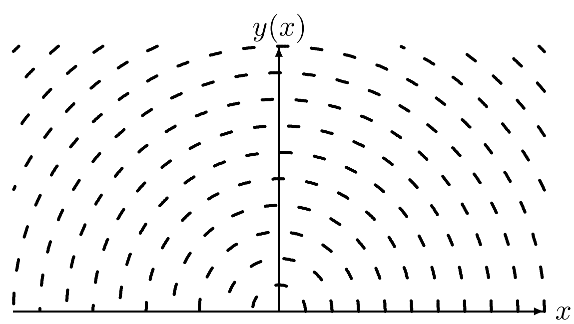
eine homogene lineare Differentialgleichung erster Ordnung, und

$$y' = a(x) \cdot y + b(x)$$

eine inhomogene lineare Differentialgleichung erster Ordnung.

Zu einer linearen Differentialgleichung gehört immer ein Definitions-Intervall $]\alpha, \beta[\subset \mathbb{R}$, auf dem alle Koeffizientenfunktionen $a_m(x)$, sowie die rechte Seite $b(x)$ definiert und stetig sind.

Explizite Differentialgleichungen erster Ordnung kann man folgendermaßen interpretieren: Sei eine Gleichung $y' = F(x, y)$ auf der Menge $G \subset \mathbb{R}^2$ gegeben. Dann legt sie in jedem Punkt $(x, y) \in G$ eine *Richtung* fest. Jede Lösungsfunktion $f(x)$ muss in jedem Punkt (x, y) die Ableitung $f' = F(x, y)$ besitzen. Dadurch ist also die Tangentialrichtung der Kurve in jedem Punkt festgelegt. Die Differentialgleichung definiert ein *Richtungsfeld* auf G:



Und die Aufgabe, die Differentialgleichung zu lösen, besteht darin, eine Funktion $f(x)$ zu suchen, deren Graph in jedem seiner Punkt die durch das Richtungsfeld vorgegebene Steigung hat.

Schließlich noch ein Wort zum Begriff *gewöhnliche* Differentialgleichung. Es gibt auch noch *partielle* Differentialgleichungen. Das sind Gleichungen, wo die gesuchte Funktion f nicht nur von einer, sondern von mehreren Variablen abhängt. Beispiele dafür sind die

$$\text{Wellengleichung} \qquad \frac{\partial^2 f}{\partial t^2} = c^2 \frac{\partial^2 f}{\partial x^2},$$

$$\text{Potentialgleichung} \qquad \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

$$\text{Wärmeleitungsgleichung} \qquad \frac{\partial f}{\partial t} = c^2 \frac{\partial^2 f}{\partial x^2}.$$

Partielle Differentialgleichungen sind viel schwerer zu lösen, als gewöhnliche, allerdings sind sie in den Anwendungen viel häufiger und viel wichtiger.

Aufgabe 3.1 *a) Sei $0 < c \in \mathbb{R}$ gegeben. Geben Sie eine explizite Differentialgleichung erster Ordnung an, die dazu äquivalent ist, dass für jede ihrer Lösungen f und für jeden Punkt x in deren Definitionsbereich gilt: Die Entfernung des Punktes $(x, f(x))$ vom Schnittpunkt seiner Tangente mit der y-Achse ist gleich c.*
b) Geben Sie eine Differentialgleichung zweiter Ordnung an, die dazu äquivalent ist, dass für jede ihrer Lösungen f und für jeden Punkt x in deren Definitionsbereich gilt: Die Entfernung des Punktes $(x, f(x))$ vom Schnittpunkt seiner Tangente mit der y-Achse ist konstant, unabhängig von x.

Aufgabe 3.2 *a) Sei $c \in \mathbb{R}$ gegeben. Geben Sie eine Differentialgleichung erster Ordnung an, die dazu äquivalent ist, dass für jede ihrer Lösungen f und für jeden Punkt x in deren Definitionsbereich gilt: Die Tangente in $(x, f(x))$ schneidet die y-Achse in $(0, c)$.*
b) Geben Sie eine Differentialgleichung zweiter Ordnung an, die dazu äquivalent ist, dass für jede ihrer Lösungen f und für jeden Punkt x in deren Definitionsbereich gilt: der Schnittpunkt der Tangente mit der y-Achse ist konstant, unabhängig von x. Lösen Sie diese Differentialgleichung.

Aufgabe 3.3 *Sei $g : \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare Funktion. Seien $a, b \in \mathbb{R}$ mit $0 < a < 1 < b$. Weiter sei $f : [a, b] \to \mathbb{R}$ eine stetig differenzierbare Funktion mit der Eigenschaft, dass für jedes $x \in [a, b]$ die Tangente an den Graphen von f die y-Achse im Punkt $(0, g(f(x)))$ schneidet. Man zeige: f ist in $[a, b]$ Lösung der Differentialgleichung*

$$y' = \frac{y - g(y)}{x}.$$

Aufgabe 3.4 *Sei $\mathbb{R}_+ := \{x \in \mathbb{R} : \ x > 0\}$. Gesucht ist eine differenzierbare Funktion*

$$f : \mathbb{R}_+ \to \mathbb{R}_+$$

mit der Eigenschaft: für jedes $x_0 \in \mathbb{R}_+$ habe die Tangente an den Graphen von f im Punkte $P_0 = (x_0, f(x_0))$ Schnittpunkte P_1, P_2 mit den positiven Koordinatenachsen, und für diese gilt

$$\| P_1 - P_0 \| = \| P_2 - P_0 \| .$$

Geben Sie eine Differentialgleichung für f an, aus welcher diese Eigenschaft folgt.

3.2 Elementare Lösungsmethoden

3.2.1 Homogene lineare Differentialgleichung erster Ordnung

Beginnen wir mit einem Beispiel, einem einfachen:

Beispiel 3.4 *Wir betrachten die lineare Differentialgleichung*

$$y' = y.$$

Wir sollen also eine Funktion $y(x)$ finden, die mit ihrer Ableitung übereinstimmt. Vielleicht erinnert sich der eine oder andere daran, dass die e-Funktion $y = e^x$ diese Eigenschaft hat. Damit haben wir eine Lösung durch Erinnern gefunden. Aber es muss doch auch systematischer gehen!

Dazu bringt man y auf die linke Seite

$$\frac{y'}{y} = 1$$

und erkennt auf der linken Seite die Ableitung

$$\frac{d}{dx} ln|y(x)| = \frac{1}{y} \cdot y'$$

der Funktion $ln|y(x)|$. Und die rechte Seite ist die Ableitung der Funktion x. Deswegen ist (für $y \neq 0$) die Differentialgleichung äquivalent mit

$$\frac{d}{dx} \left(ln|y(x)| - x \right) \equiv 0,$$

bzw. mit

$$ln|y(x)| = x + c, \qquad |y(x)| = e^{x+c} = c' \cdot e^x$$

mit der positiven Konstante $0 < c' = e^c \in \mathbb{R}$. Wollen wir die Absolutstriche weglassen, so müssen wir

$$y(x) = \pm c' \cdot e^x$$

schreiben. Und das $\pm$-Zeichen können wir uns schenken, wenn wir für c' positive oder negative Werte zulassen. Schließlich ist aber $y(x) \equiv 0$ auch eine Lösung (die Null-Lösung*). Diese Nullösung ist bei unserem Lösungs-Verfahren nicht mit herausgekommen. Aber wir können sie mit dazu nehmen, und dann haben wir die Lösungen*

$$y(x) = c \cdot e^x, \quad c \in \mathbb{R}$$

gefunden.

Dieses, relativ naive Lösungs-Verfahren ist sehr wichtig, man muss es auswendig können. Vor allem deswegen, weil man damit alle homogenen linearen Differentialgleichungen erster Ordnung behandeln kann:

Sei etwa die Differentialgleichung

$$y' = a(x) \cdot y$$

mit einer stetigen Koeffizientenfunktion $a :]\alpha, \beta[\to \mathbb{R}$ vorgelegt. Auch hier gibt es die Null-Lösung $y \equiv 0$. Und wo $y \neq 0$ ist, dividieren wir durch y, um die Gleichung auf die Form

$$\frac{y'}{y} = a(x)$$

zu bringen. Wenn wir jetzt eine Stammfunktion $A(x)$ für $a(x)$ kennen, können wir genau wie eben integrieren:

$$ln|y(x)| = A(x) + c, \qquad |y(x)| = e^c \cdot e^{A(x)}, \qquad y(x) = \pm c' \cdot e^{A(x)}.$$

Nehmen wir die Null-Lösung hinzu, so haben wir die Lösungen

$$y(x) = c \cdot e^{A(x)}, \quad c \in \mathbb{R},$$

ermittelt.

Auf diese Weise kann man *alle* homogenen Differentialgleichungen erster Ordnung $y' = a(x) \cdot y$ lösen. Alles was man braucht, ist eine Stammfunktion $A(x)$ für die Koeffizientenfunktion $a(x)$. Man kann es auch so sagen: Die Ermittlung von Lösungen ist auf die Integration einer Funktion $a(x)$ zurückgeführt. Die Ermittlung einer Stammfunktion $A(x)$ kann explizit möglich sein, oder auch nicht. Das ist jetzt nicht mehr unser Problem. Schlimmstenfalls schreiben wir halt

$$A(x) = \int_{x_0}^{x} a(\xi) d\xi.$$

Immer, wenn man eine Differentialgleichung auf die Ermittlung einer Stammfunktion zurückgeführt hat, ist man happy, und lehnt sich entspannt zurück. Daher kommt auch der Sprachgebrauch: 'Eine Differentialgleichung integrieren'. Damit meint man: 'Eine Differentialgleichung lösen'.

Beispiel 3.5 *Noch ein paar Beispiele:*

$$\begin{array}{rcl|rcl|rcl}
y' &=& a \cdot y, \quad a \in \mathbb{R}, & y' &=& x \cdot y & y' &=& y/x \\
y'/y &=& a & y'/y &=& x & y'/y &=& 1/x \\
ln|y| &=& a \cdot x + c & ln|y| &=& x^2/2 + c & ln|y| &=& ln|x| + c \\
y &=& \pm e^c \cdot e^{a \cdot x} & y &=& \pm e^c \cdot e^{x^2/2} & y &=& \pm e^c \cdot e^{ln|x|} \\
y &=& c' \cdot e^{a \cdot x} & y &=& c' \cdot e^{x^2/2} & y &=& c' \cdot x
\end{array}$$

So, das dürfte als Illustration dieser Methode genügen. Es ist eine Methode, um Lösungen für homogene, lineare Differentialgleichungen erster Ordnung zu finden. Unter dem *Lösen einer Differentialgleichung* versteht man aber das Auffinden *aller* ihrer Lösungen. Die verwendete Methode gibt in der Tat *alle Lösungen*. Das ist allerdings etwas subtiler (wegen der Probleme mit $y \neq 0$), das möchte ich deswegen nicht hier, sondern etwas allgemeiner in 3.3 diskutieren. Ich halte es eben für wichtiger, Beispielmaterial zusammenzustellen, und ein gewisses Gefühl für das Problem zu entwickeln, als gleich mit der allgemeinen Theorie ins Haus zu fallen.

3.2.2 Inhomogene lineare Differentialgleichung erster Ordnung

Deswegen wende ich mich jetzt einer Variation der homogenen linearen Differentialgleichung erster Ordnung zu, nämlich der *inhomogenen* linearen Differentialgleichung erster Ordnung

$$y' = a(x) \cdot y + b(x).$$

Ihre Beziehung zur homogenen Gleichung ist genau dieselbe wie die Beziehung eines inhomogenen linearen Gleichungs-Systems der Linearen Algebra zu seinem homogenen Gleichungssystem:

Satz 3.1 *Die allgemeine Lösung der inhomogenen Differentialgleichung*

$$y' = a(x) \cdot y + b(x)$$

erhält man, indem man zu einer speziellen Lösung $y_0 = f_0(x)$ *dieser Gleichung alle Lösungen* $y = f(x)$ *der zugehörigen homogenen Differentialgleichung*

$$y' = a(x) \cdot y$$

addiert.

Beweis (wie in der Linearen Algebra). a) Sei $y_0 = f_0(x)$ eine Lösung der inhomogenen Gleichung und $y = f(x)$ eine der homogenen. Dann gilt also

$$y_0' = a(x) \cdot y_0 + b(x), \qquad y' = a(x) \cdot y.$$

Daraus folgt

$$(y_0 + y)' = a(x) \cdot y_0 + b(x) + a(x) \cdot y = a(x) \cdot (y + y_0) + b(x).$$

Also ist $y_0 + y$ auch eine Lösung der inhomogenen Gleichung.

b) Sei y_0 eine spezielle Lösung der inhomogenen Gleichung und u eine weitere Lösung. Das bedeutet also

$$y_0' = a(x) \cdot y_0 + b(x) \quad \text{und} \quad u' = a(x) \cdot u + b(x).$$

Für $y := u - y_0$ folgt daraus

$$y' = [a(x) \cdot u + b(x)] - [a(x) \cdot y_0 + b(x)] = a(x) \cdot (u - y_0) = a(x) \cdot y.$$

Also ist y eine Lösung der homogenen Gleichung und $u = y_0 + y$. □

Somit kommt es darauf an, mit Gewalt eine, eine einzige Lösung der inhomogenen Gleichung zu finden. Auch dafür gibt es eine Methode, die man wieder auswendig wissen muss. Sie trägt den in sich widerspruchsvollen Namen: *Variation der Konstanten.* Eine Konstante kann ja nicht variieren, sie ist doch konstant, deswegen heißt sie ja auch so. Macht nichts, variieren wir die Konstante, und zwar die Konstante c bei der Lösung

$$y = c \cdot e^{A(x)}$$

der homogenen Gleichung. Wir machen also den Ansatz

$$y_0(x) = c(x) \cdot y(x), \qquad \text{wo } y \text{ Lösung der homogenen Gleichung.}$$

Wir differenzieren mit der Produktregel und vergleichen das Resultat mit der inhomogenen Differentialgleichung:

$$\begin{array}{rcll} y_0'(x) & = & c'(x) \cdot y(x) + c(x) \cdot y'(x) & \text{(Produktregel)} \\ & = & c'(x) \cdot y(x) + c(x) \cdot a(x) \cdot y(x) & \text{(homogene Dgl.)} \\ & = & c'(x) \cdot y(x) + a(x) \cdot y_0(x) & \\ y_0'(x) & = & b(x) + a(x) \cdot y_0(x) & \text{(inhomogene Dgl.)} \end{array}$$

Somit ist y_0 Lösung der inhomogenen Gleichung, genau dann wenn

$$c'(x) \cdot y(x) = b(x), \quad \text{d.h.} \quad c'(x) = \frac{b(x)}{y(x)}.$$

Damit ist das Auffinden von y_0, d.h., das Auffinden von $c(x)$, auf die Integration

$$c(x) = \int \frac{b(x)}{y(x)} dx = \int b(x) e^{-A(x)} dx$$

zurückgeführt. Wieder einmal eine Integration!

Rechnen wir mal zwei Beispiele, ein einfaches, damit man das Prinzip deutlich erkennt, und ein technisches, aus einer Klausuraufgabe.

Beispiel 3.6 (einfaches)

$$y' = y + 1.$$

Die zugehörige homogene Differentialgleichung ist $y' = y$. Wir kennen ihre Lösungen $y = c \cdot e^x$ und variieren die Konstante:

$$y_0 = c(x) \cdot e^x.$$

Wir müssen mit der Produktregel differenzieren, und das Resultat mit der Differentialgleichung vergleichen:

$$\begin{aligned} y_0' &= c'(x) \cdot e^x + c(x) \cdot e^x \\ &= c'(x) \cdot e^x + y_0, \\ y_0' &= y_0 + 1. \end{aligned}$$

Subtraktion beider Gleichungen liefert

$$c'(x) \cdot e^x - 1 = 0, \qquad c'(x) = e^{-x}, \qquad c(x) = \int e^{-x} dx = -e^{-x}.$$

(Die Integrationskonstante bei dieser Integration können wir vergessen, weil es uns nur auf eine spezielle, eine einzige spezielle Lösung y_0 ankommt.) Wir erhalten also

$$y_0 = -e^{-x} \cdot e^x = -1.$$

Überraschend einfach! In der Tat, y_0 ist konstant, also $y_0' = 0$ und $y_0 + 1 = 0 = y_0'$. Das hätte man auch erraten können. Mit der speziellen Lösung $y_0 = -1$ der inhomogenen Gleichung und der allgemeinen Lösung $y = c \cdot e^x$ der homogenen Gleichung erhalten wir die allgemeine Lösung der inhomogenen Gleichung

$$y_0 + y = -1 + c \cdot e^x.$$

Beispiel 3.7 (technisches)

$$y' = -cot(x) \cdot y + 5e^{cos(x)}.$$

Die homogene Gleichung

$$y' = -cot(x) \cdot y$$

führt uns auf

$$\frac{y'}{y} = -cot(x) = -\frac{cos(x)}{sin(x)} = -\frac{d}{dx} ln|sin(x)|,$$

$$ln|y| = -ln|sin(x)| + c, \qquad y = c \cdot \frac{1}{sin(x)}.$$

Wir variieren die Konstante mit dem Ansatz

$$y_0 = c(x) \cdot \frac{1}{sin(x)}.$$

Wieder wenden wir die Produktregel an und vergleichen mit der Differentialgleichung:

$$\begin{aligned} y_0' &= c'(x) \cdot \frac{1}{sin(x)} + c(x) \cdot \frac{-cos(x)}{sin^2(x)}, \\ y_0' &= -cot(x) y_0 + 5e^{cos(x)} \\ &= -c(x) \cdot \frac{cos(x)}{sin^2(x)} + 5e^{cos(x)}. \end{aligned}$$

Das liefert

$$\frac{c'(x)}{sin(x)} = 5e^{cos(x)}, \qquad c'(x) = 5sin(x)e^{cos(x)}, \qquad c(x) = 5e^{cos(x)}.$$

Und die allgemeine Lösung der inhomogenen Gleichung ist

$$y_0 + y = \frac{5e^{cos(x)}}{sin(x)} + \frac{c}{sin(x)}.$$

Wegen des Sinus im Nenner, mit seinen vielen Nullstellen, ist es eine ganz andere Frage, wo diese Lösung existiert, darauf kommen wir auch bei einer anderen Gelegenheit zurück.

3.2.3 Trennung der Variablen

Die inhomogene lineare Differentialgleichung erster Ordnung $y' = a(x) \cdot y + b(x)$ ist eine Verallgemeinerung der homogenen linearen Differentialgleichung $y' = a(x) \cdot y$. Eine andere Verallgemeinerung ist die Differentialgleichung

$$y' = a(x) \cdot b(y).$$

Hier kommt y auf der rechten Seite nicht mehr linear vor, sondern innerhalb einer neuen Funktion b. Das Wesentliche ist, dass die rechte Seite ein Produkt

$$\text{Funktion } a \text{ von } x \quad \times \quad \text{Funktion } b \text{ von } y$$

ist. Deswegen kann man *die Variablen x und y trennen:*

$$\frac{y'}{b(y)} = a(x).$$

Diese Trennung der Variablen treiben wir noch weiter, indem wir von der Gleichung

$$\frac{1}{b(y)} \cdot \frac{dy}{dx} = a(x)$$

zu der Gleichung

$$\frac{1}{b(y)} \cdot dy = a(x)dx$$

übergehen. Die ist natürlich sinnlos, deswegen integrieren wir bevor das jemand merkt ganz schnell darüber
$$\int_{y_0} \frac{dy}{b(y)} = \int_{x_0} a(x)dx + c.$$
Und diese Gleichung ist glücklicherweise wieder sinnvoll.

Dass dieses Verfahren legitim ist, steht mit allen Voraussetzungen liebevoll im Forster II. Wir wollen uns keine Gedanken darüber machen, sondern einige Beispiele rechnen, um zu dieser Methode Vertrauen zu fassen.

Beispiel 3.8 *Sei etwa die Gleichung*
$$y' = \frac{x}{y}$$
vorgelegt. Wir trennen die Variablen
$$y \cdot \frac{dy}{dx} = x, \qquad ydy = xdx$$
und integrieren
$$\int ydy = \int xdx + c', \qquad \frac{y^2}{2} = \frac{x^2}{2} + c', \qquad y^2 = x^2 + c.$$
Das Verfahren liefert die Funktionen
$$y = \pm\sqrt{x^2 + c}, \quad c \in \mathbb{R}.$$
Machen wir lieber noch die Probe:
$$y' = \pm\frac{x}{\sqrt{x^2 + c}} = \frac{x}{y},$$
es stimmt.

Beispiel 3.9 *Ein anderes Beispiel ist*
$$y' = -y^r, \quad 0 < r \in \mathbb{R}.$$
Auf der rechten Seite kommt jetzt gar kein x vor, macht nichts wir trennen es trotzdem vom y:
$$y^{-r} \cdot \frac{dy}{dx} = -1, \qquad y^{-r}dy = -dx$$
und integrieren
$$\frac{1}{1-r}y^{1-r} = c - x, \qquad y = ((1-r)(c-x))^{1/(1-r)}.$$
Das war ja ganz schön einfach. Wo die Lösung existiert, das ist wieder eine andere Frage, und den Fall $r = 1$ dürfen wir auch nicht vergessen:
$$\frac{dy}{y} = -dx, \qquad ln|y| = -x, \qquad y = c \cdot e^{-x}.$$

3.2.4 Substitutionen

Manche Integrale kann man durch Substitution ausrechenen, ähnlich kann man manche Differentialgleichungen durch Substitution integrieren. Es gibt im wesentlichen zwei Substitutionen, deren Konsequenzen man kennen muss:

$$z = y^k \quad \text{und} \quad z = y/x.$$

Schauen wir uns das Resultat der ersten Substitution an, wenn man sie auf inhomogene, lineare Differentialgleichungen erster Ordnung loslässt:

Betrachten wir die Gleichung

$$z' = a(x) \cdot z + b(x)$$

und substituieren hier $z = y^k$. Was passiert? Wir erhalten

$$k \cdot y^{k-1} \cdot y' = a(x) \cdot y^k + b(x),$$

bzw. nach Division durch $k \cdot y^{k-1}$ die Differentialgleichung

$$y' = \frac{a(x)}{k} \cdot y + \frac{b(x)}{k} \cdot y^{1-k}.$$

Wenn wir jetzt noch umbenennen $a/k \to a,\ b/k \to b,\ 1-k \to k$ wird daraus die Differentialgleichung

$$y' = a(x) \cdot y + b(x) \cdot y^k.$$

Sie heißt *Bernoullische Differentialgleichung* und sieht ziemlich anders aus, als die inhomogene lineare, aus der sie entstanden ist: Wo bei der inhomogenen linearen der von y freie Summand $b(x)$ steht, haben wir jetzt einen Summanden $b(x) \cdot y^k$ mit einer gefährlichen Potenz y^k.

Hier kann k eine beliebige reelle, nicht notwendig ganze Zahl sein. Notfalls muss man sich halt auf den Bereich $y > 0$ beschränken. Das muss man sich im Einzelfall genau ansehen. Wir wollen jetzt die Transformation nocheinmal rückwärts (d.h., in der für die Anwendung wichtigen Richtung) durchgehen:

Sei eine Bernoullische Differentialgleichung

$$y' = a(x) \cdot y + b(x) \cdot y^k$$

vorgelegt. Man dividiert durch das störende y^k

$$y^{-k} y' = a(x) \cdot y^{1-k} + b(x)$$

und macht mit der Substitution

$$z = y^{1-k}, \qquad z' = (1-k) y^{-k} y'$$

daraus die lineare Gleichung

$$\frac{1}{1-k} z' = a(x) \cdot z + b(x), \quad \text{bzw.} \quad z' = (1-k)a(x) \cdot z + (1-k)b(x).$$

Der Fall $k = 1$ muss dabei natürlich ausgeschlossen werden. Da wäre die ursprüngliche Bernoullische Gleichung aber auch eine ganz gewöhnliche homogene lineare Differentialgleichung gewesen.

Beispiel 3.10 *Schauen wir mal, ob das in der Praxis tatsächlich funktioniert. Die einfachste, nicht ganz triviale Bernoullische Gleichung ist wohl*

$$y' = y + y^2.$$

Was sagt unser Rezept? Durch y^2 dividieren

$$y^{-2}y' = y^{-1} + 1$$

und $z = y^{-1}$ substituieren

$$-z' = z + 1, \qquad z' = -z - 1.$$

Die homogene Gleichung $z' = -z$ hat die allgemeine Lösung $z = c \cdot e^{-x}$, und Variation der Konstanten, $z_0 = c(x) \cdot e^{-x}$, führt zu

$$z_0' = c'(x) \cdot e^{-x} - c(x) \cdot e^{-x} = c'(x) \cdot e^{-x} - z_0(x).$$

Mit der Differentialgleichung $z_0' = -z_0 - 1$ erhalten wir daraus

$$c'(x) \cdot e^{-x} = -1, \qquad c'(x) = -e^x, \qquad c(x) = -e^x$$

und die spezielle Lösung $z_0(x) = -1$. Auch das hätten wir wieder erraten können! Die allgemeine Lösung der inhomogenen linearen Gleichung ist also

$$z(x) = -1 + c \cdot e^{-x}, \quad c \in \mathbb{R},$$

und die Lösung der Bernoullischen Gleichung wird

$$y = \frac{1}{z} = \frac{1}{c \cdot e^{-x} - 1}.$$

Natürlich sollte man die Probe machen, ob man sich nicht vielleicht verrechnet hat.

Lassen wir jetzt die Substitution $z = y/x$ auf die Differentialgleichung

$$z' = \frac{f(z)}{x}$$

los, die man durch Trennung der Variablen behandeln kann. Mit

$$z' = \frac{xy' - y}{x^2}$$

erhalten wir daraus

$$\frac{xy' - y}{x^2} = \frac{f(y/x)}{x}, \qquad y' = \frac{y}{x} + f\left(\frac{y}{x}\right).$$

Wenn wir $g(z) := z + f(z)$ setzen, wird daraus

$$y' = g\left(\frac{y}{x}\right).$$

Die rechte Seite hängt nur vom Quotienten $z = y/x$ ab, soetwas nennt man homogen, und die Differentialgleichung nennt man eine *homogene Differentialgleichung*. Das ist eine etwas unglückliche Sprachregelung, weil sie nichts mit einer homogenen linearen Differentialgleichung zu tun hat.

Spielen wir diese Substitution jetzt nocheinmal rückwärts durch, so wie man sie anwendet. Vorgelegt sei also die homogenen Differentialgleichung

$$y' = g\left(\frac{y}{x}\right).$$

Wir müssen substituieren

$$z = \frac{y}{x}, \qquad y = z \cdot x, \qquad y' = z' \cdot x + z.$$

Dann wird die Gleichung

$$z' \cdot x + z = g(z), \qquad z' = \frac{g(z) - z}{x} = \frac{f(z)}{x}$$

mit $f(z) = g(z) - z$, eine Differentialgleichung mit trennbaren Variablen.

Beispiel 3.11 *So, und jetzt schauen wir uns noch an, wie das praktisch funktioniert. Nehmen wir die homogene Differentialgleichung*

$$y' = \frac{1}{2}\left(1 + \frac{y^2}{x^2}\right)$$

und substituieren $z = y/x,\ y = x \cdot z,\ y' = x \cdot z' + z$. *Wir finden*

$$x \cdot z' + z = \frac{1}{2}(1 + z^2), \qquad z' = \frac{1}{2} \cdot \frac{1 - 2z + z^2}{x}, \qquad \frac{z'}{(1-z)^2} = \frac{1}{2x}.$$

Nach einer Integration wird daraus

$$\frac{1}{1-z} = \frac{1}{2} ln|x| + c', \qquad 1 - z = \frac{2}{ln|x| + c}.$$

Also wird

$$z = 1 - \frac{2}{ln|x| + c}, \qquad y = x - \frac{2x}{ln|x| + c}.$$

Auch hier muss man unbedingt wieder die Probe machen. (Natürlich habe ich es mit viel Phantasie so hingetrickst, dass man die entstehende Differentialgleichung mit getrennten Variablen in geschlossener Form integrieren kann.)

Fassen wir die beiden Rezepte plakativ zusammen:

$$\begin{array}{rl}
\text{Bernoullische Gleichung:} & y' = a(x) \cdot y + b(x) \cdot y^k \\
\text{Division durch } y^k\text{:} & y^{-k} y' = a(x) \cdot y^{1-k} + b(x) \\
\text{Substitution } z = y^{1-k}\text{:} & \dfrac{z'}{1-k} = a(x) \cdot z + b(x)
\end{array}$$

$$\begin{array}{rl}
\text{Homogene Gleichung:} & y' = g\left(\dfrac{y}{x}\right) \\
\text{Substitution } y = z \cdot x\text{:} & z' \cdot x + z = g(z) \\
 & z' = \dfrac{g(z) - z}{x}
\end{array}$$

Jetzt noch zwei Fallstudien aus Staatsexamensaufgaben mit anderen Substitutionen, merkwürdigerweise beide vom Frühjahr 93:

Beispiel 3.12 *Damals war in T 1, A 6 die Differentialgleichung*

$$(4y + 2xy) \cdot y' = y^2 - 1$$

zu behandeln. Die linke Seite der Gleichung ist auf den ersten Blick unverdaulich, aber auf der rechten Seite bleibt der Blick am y^2 hängen. Das widersetzt sich jeder linearen Theorie. Sollte man es wegsubstituieren, etwa

$$z = y^2$$

setzen? Ja! Denn links kann man $2y$ ausklammern und die Gleichung

$$(2 + x) \cdot 2yy' = (2 + x)(y^2)' = (2 + x)z' = z - 1$$

schreiben. Dann wird daraus die inhomogene lineare Differentialgleichung

$$z' = \frac{z - 1}{2 + x}.$$

Hier bietet es sich an, statt der allgemeinen Theorie auf die zweite Substitution $u = z - 1$ zurückzugreifen, und die Gleichung in die Form

$$u' = \frac{u}{2 + x}$$

zu bringen. Wir trennen die Variablen

$$\frac{u'}{u} = \frac{1}{2 + x}, \qquad ln|u| = ln|2 + x| + c', \qquad u = c \cdot (2 + x).$$

Jetzt machen wir noch die Substitutionen rückgängig:

$$z = 1 + c(2 + x), \qquad y = \sqrt{1 + c(2 + x)}.$$

Wo die erhaltenen Lösungen existieren, das wird auf einem anderen Blatt stehen.

Beispiel 3.13 *In T 2, A 5 war damals die Gleichung*

$$y'(x) \cdot sin(x) + y(x) \cdot cos(x) = sin(x) + x \cdot cos(x)$$

vorgelegt. Scheußlich, was ist zu tun? Kein Hinweis, die Zeit verrinnt! Hätte ich doch damals beim Barth nicht soviel geschwätzt sondern mehr aufgepasst! Oder vielleicht weniger Aufgaben abgeschrieben! Solche an sich korrekten Gedanken helfen auch nicht weiter. Irgendwas ist aber doch komisch: soviele Winkelfunktionen, wo eine die Ableitung der anderen ist. Das ist der Schlüssel! Die linke Seite ist die Ableitung

$$(y(x) \cdot sin(x))' = y'(x) \cdot sin(x) + y(x) \cdot cos(x).$$

Was passiert, wenn wir

$$z(x) = y(x) \cdot sin(x)$$

substituieren? Etwas sehr erleichterndes: Die Differentialgleichung wird

$$z' = sin(x) + x \cdot cos(x) = (x \cdot sin(x))'.$$

Nach einer Integration erhalten wir

$$z(x) = x \cdot sin(x) + c, \qquad y(x) = x + \frac{c}{sin(x)}.$$

Aufgabe 3.5 *Lösen Sie die Differentialgleichungen*

a) $y'(x) = exp(2x - y(x))$, *b)* $y' = 2xy^2$, *c)* $xy' = (1-x)(1+y^2)$, *d)* $x^2y' + 2xy = ln(x)$, $(x > 0)$.

Aufgabe 3.6 *Man bestimme alle differenzierbaren Funktionen* $f :]0, \infty[\to \mathbb{R}$ *mit der Eigenschaft, dass für alle* $x \in]0, \infty[$ *die Tangente an den Graphen von* f *im Punkte* $(x, f(x))$ *die* y*-Achse im Punkt* $(0, \frac{1}{2}f(x))$ *schneidet.*

Aufgabe 3.7 *Bestimmen Sie für* $x \in]-\pi/2, \pi/2[$ *die allgemeine Lösung der Differentialgleichung* $y' \cdot cos(x) - 2y \cdot sin(x) = x$.

Aufgabe 3.8 *Seien* K, r, s *Konstanten. Bestimmen sie die allgemeine Lösung der Differentialgleichung*

$$y'(t) - s \cdot y(t) = K \cdot e^{rt}.$$

Aufgabe 3.9 *Lösen Sie die Differentialgleichungen*

$$a)\ y' = -4y + x^2, \quad b)\ y' = 2xy + x^4, \quad c)\ y' = cos(x) \cdot e^y, \quad d)\ y' = \frac{xy-1}{1-x^2}.$$

Aufgabe 3.10 *Lösen Sie die Differentialgleichungen*

$$a)\ y' = \frac{x}{y}, \quad b)\ y' = \frac{3x^2}{2y^3}, \quad c)\ y' = \frac{1+y+y^2}{x(x^2-4)}.$$

Aufgabe 3.11 *Lösen Sie die Bernoullischen Differentialgleichungen*

$$\mathbf{a)}\ y' = y^r \text{ für } r = 1/2,\ 1/3,\ 2/3, \quad \mathbf{b)}\ y' = x \cdot y^{2/3}, \quad \mathbf{c)}\ y' = x(x^2-1) \cdot y^{2/3}.$$

Aufgabe 3.12 *Lösen Sie die homogenen Differentialgleichungen*

$$a)\ y' = \frac{y}{x} - \frac{\sqrt{x^2+y^2}}{x}, \quad b)\ y' = \frac{x^3-y^3}{xy^2},$$

Aufgabe 3.13 *Verwandeln Sie mit einer Transfomation $x = u + a,\ y = v + b,\ a, b \in \mathbb{R}$ die Differentialgleichung*

$$y' = \frac{x - 2y + 1}{2x - y + 1}$$

in eine homogene Differentialgleichung.

3.3 Existenz und Eindeutigkeit

In 3.2 haben wir Methoden kennen gelernt, mit denen man explizite Differentialgleichungen erster Ordnung (manchmal) lösen kann. D. h., es waren Methoden, um Lösungen zu finden. In diesem Paragraphen wollen wir uns darum kümmern, ob es immer Lösungen gibt, und wieviele es gibt.

Satz 3.2 (Existenz-Satz von Peano) *Die Funktion $F : U \to \mathbb{R}^2$ sei stetig auf der offenen Menge $U \subset \mathbb{R}^2$. Dann gibt es zu jedem Punkt $(x_0, y_0) \in U$ eine lokale Lösung der Differentialgleichung*

$$y' = F(x, y).$$

D.h., es gibt ein $\epsilon > 0$ und eine differenzierbare Funktion

$$f :]x_0 - \epsilon, x_0 + \epsilon[\to \mathbb{R} \quad \text{mit} \quad f(x_0) = y_0$$

und so, dass für alle $x \in]x_0 - \epsilon, x_0 + \epsilon[$ gilt

$$(i) \quad (x, f(x)) \in U, \qquad (ii) \quad f'(x) = F(x, f(x)).$$

Diesen Satz wollen wir hier nicht beweisen. Er wird weder im Forster, noch im Erwe bewiesen. Einen Beweis findet man etwa im Kamke.

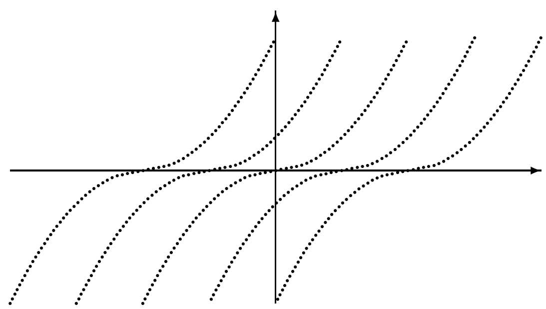
Dieser Satz sagt also: Wenn die rechte Seite $F(x, y)$ stetig ist, dann existieren (lokale) Lösungen der Differentialgleichung $y' = F(x, y)$. Aber es gilt keine Eindeutigkeit: durch einen Punkt $(x_0, y_0) \in G$ können mehrere Lösungen gehen. Ein Beispiel dafür ist die Differentialgleichung

$$y' = 3 \cdot y^{2/3}.$$

Die rechte Seite $F(x, y) = 3y^{2/3}$ ist für alle $(x, y) \in G := \mathbb{R}^2$ definiert und stetig. Wir lösen die Differentialgleichung nach dem Schema 'Trennung der Variablen':

$$\begin{aligned} \frac{1}{3} \cdot y^{-\frac{2}{3}} y' &= 1 \\ y^{\frac{1}{3}} &= x + c \\ y &= (x + c)^3. \end{aligned}$$

Als Lösungen erhalten wir kubische Parabeln $y = x^3$, auf der x-Achse etwas hin- und her-verschoben, mit der Translationskonstante $c \in \mathbb{R}$. Das sieht doch ganz ordentlich aus:



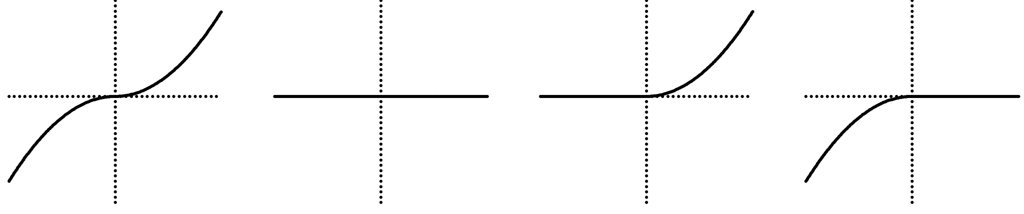
Aber! Durch den Nullpunkt gibt es die beiden verschiedenen Lösungen

$$y = x^3 \quad \text{und} \quad y \equiv 0$$

der Differentialgleichung. Noch schlimmer: Auch die beiden Funktionen

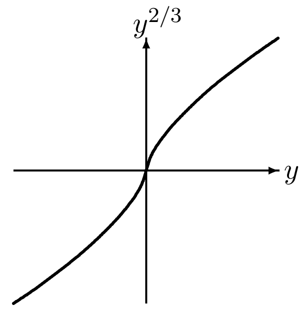
$$y = \begin{cases} 0 & (x \leq 0) \\ x^3 & (x \geq 0) \end{cases} \qquad \text{und} \qquad y = \begin{cases} x^3 & (x \leq 0) \\ 0 & (x \geq 0) \end{cases}$$

sind differenzierbar, und lösen die Differentialgleichung.



Und noch viel schlimmer: Man kann noch viel mehr Funktionen aus Stücken von verschobenen kubischen Parabeln $y = (x - c)^3$ und der x-Achse im Punkt $(c, 0)$ zusammensetzen, die differenzierbar sind, und die Differentialgleichung lösen. Daraus lernen wir: Wenn man für die rechte Seite $F(x, y)$ nur die Stetigkeit voraussetzt, gibt es zwar durch jeden Punkt der Definitionsmenge von F Lösungen, aber i.A. mehr als eine.

Woran liegt das im Fall $F(x, y) = y^{2/3}$? Beim kritischen Wert $y = 0$ ist die Funktion $y^{2/3}$ zwar stetig, aber sie hat die Steigung ∞. Oder anders ausgedrückt: Bei einer winzigen Veränderung von y ändert sich F in diesem Bereich sehr merklich. Die folgende Anforderung an die Funktion F schließt so etwas aus:



Definition 3.4 *Eine Funktion* $F : U \to \mathbb{R}$, $U \subset \mathbb{R}^2$, *genügt auf* U *einer* Lipschitz-Bedingung, *wenn es eine Konstante* $0 < L \in \mathbb{R}$ *gibt, so, dass für alle* (x, y_1) *und* $(x, y_2) \in U$ *gilt:*

$$|F(x, y_1) - F(x, y_2)| \leq L \cdot |y_1 - y_2|.$$

Eine Konstante L mit dieser Eigenschaft heißt Lipschitz-Konstante.

Das ist genau dieselbe Bedingung, wie in Definition 1.19. Allerdings bezieht sie sich hier nur auf die zweite Variable y, bei festem x.

Satz 3.3 (Lokale Lipschitz-Bedingung) *Die Funktion* $F : U \to \mathbb{R}$ *auf der offenen Menge* $U \subset \mathbb{R}^2$ *sei stetig partiell nach* y *differenzierbar. Dann genügt* F *auf* U <u>*lokal*</u> *einer Lipschitz-Bedingung. D.h., zu jedem Punkt* $(x_0, y_0) \in G$ *gibt es ein Quadrat*

$$Q = \{(x, y) \in \mathbb{R}^2 : \ |x - x_0| \leq \epsilon, \ |y - y_0| \leq \epsilon\} \subset U, \ \epsilon > 0,$$

und eine Konstante L *so, dass für alle* (x, y_1) *und* $(x, y_2) \in Q$ *gilt:*

$$|F(x, y_1) - F(x, y_2)| \leq L \cdot |y_1 - y_2|.$$

Beweis. Weil das Quadrat U kompakt ist, ist die stetige partielle Ableitung $\partial F / \partial y$ dort beschränkt. Wenn etwa

$$\left| \frac{\partial F}{\partial y}(x, y) \right| \leq L \quad \text{für alle } (x, y) \in U,$$

dann folgt die Behauptung aus dem MWS der Differentialrechnung, ganz so wie die Aussage von Satz 2.5. □

Satz 3.4 (Existenz- und Eindeutigkeitssatz von Picard - Lindelöf) *Die stetige Funktion* $F : U \to \mathbb{R}$ *genüge auf der offenen Menge* $U \subset \mathbb{R}^2$ <u>*lokal*</u> *einer* <u>*Lipschitz-Bedingung*</u>. *Dann* <u>*gibt es*</u> *durch jeden Punkt* $(x_0, y_0) \in U$ <u>*lokal*</u> *eine Lösung* $y(x)$ *der Differentialgleichung*

$$y' = F(x, y),$$

und diese ist durch (x_0, y_0) <u>*eindeutig bestimmt*</u>.

Die Behauptung bedeutet: Es gibt ein $\epsilon > 0$ und eine differenzierbare Funktion $f(x)$ auf einem Intervall $]x_0 - \epsilon, x_0 + \epsilon[$ mit $f(x_0) = y_0$ so, dass für alle x in diesem Intervall gilt:

$$(x, f(x)) \in U, \quad \text{und} \quad f'(x) = F(x, f(x)).$$

Und falls g eine andere Funktion mit diesen Eigenschaften ist, so gilt $f(x) = g(x)$ für alle $x \in]x_0 - \epsilon, x_0 + \epsilon[$.

Beweis des Satzes. Weil U offen ist, gibt es ein $r > 0$ so, dass das Quadrat

$$Q := \{(x, y) \in \mathbb{R}^2 : |x - x_0| \leq r, |y - y_0| \leq r$$

ganz zu U gehört. Um Schreibarbeit zu sparen, verwenden wir gelegentlich auch die Notation

$$I := [x_0 - r, x_0 + r], \quad J := [y_0 - r, y_0 + r].$$

Dann ist also $Q = I \times J$. Weil F lokal einer Lipschitz-Bedingung genügt, können wir außerdem $r > 0$ so klein wählen, dass

$$|F(x, y_1) - F(x, y_2)| \leq L \cdot |y_1 - y_2| \leq L \cdot r$$

für alle $(x, y_1), (x, y_2) \in Q$.

Die Beweisidee besteht darin, die Differentialgleichung in eine Integralgleichung umzuwandeln. Dazu betrachten wir eine stetige Funktion $f : I \to J$ und zeigen die Äquivalenz folgender beiden Eigenschaften:

a) f ist auf $]x_0 - r, x_0 + r[$ differenzierbar mit $f'(x) = F(x, f(x))$ für $|x - x_0| < r$ und $f(x_0) = y_0$.

b) F genügt auf I der Integralgleichung

$$f(x) = y_0 + \int_{x_0}^{x} F(t, f(t))dt.$$

Beweis von a) $\Rightarrow$ b): Für $|x - x_0| < r$ ist nach dem HDI

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt = y_0 + \int_{x_0}^{x} F(t, f(t))dt.$$

Weil die rechte und die linke Seite dieser Gleichung stetig in x sind, gilt die Gleichung auch in den Randpunkten $x_0 \pm r$.

Beweis von b) $\Rightarrow$ a): Im offenen Intervall ist das Integral nach seiner oberen Grenze differenzierbar mit Ableitung $F(x, f(x))$. Dann ist auch f differenzierbar mit dieser Ableitung. Und $f(x_0) = y_0$ ist offensichtlich.

Und eine Funktion f, welche der Integralgleichung genügt, wollen wir finden durch Anwendung des Banachschen Fixpunktsatzes auf die *Picard-Abbildung*

$$P : C^0(I, J) \to C^0(I, \mathbb{R}), \quad (P(f))(x) := f(x_0) + \int_{x_0}^{x} F(t, f(t))dt.$$

Zuerst untersuchen wir, ob $P(f)$ wieder in $C^0(I, J)$ liegt. Dazu schätzen wir ab mit $M := max_{(x,y) \in Q} |F(x, y)|$

$$|(P(f))(x) - f(x_0)| = \left| \int_{x_0}^{x} F(t, f(t))dt \right| \leq \left| \int_{x_0}^{x} |F(t, f(t))|dt \right| \leq M \cdot |x - x_0| \leq r$$

falls

$$|x - x_0| \leq r_1 := min\{r, \frac{r}{M}\}.$$

Wir ersetzen r durch r_1, also I durch $I_1 := [x_0 - r_1, x_0 + r_1]$ und haben für die Picard-Abbildung

$$P : C^0(I_1, J) \to C^0(I_1, J).$$

Jetzt müssen wir noch beweisen, dass die Picard-Abbildung kontrahierend ist. Wir schätzen ab:

$$\begin{aligned}
\| P(f) - P(g) \|_{I_1} &= \max_{x \in I_1} |(P(f))(x) - (P(g))(x)| \\
&= \max_{x \in I_1} \left| \int_{x_0}^{x} |F(t, f(t)) - F(t, g(t))| dt \right| \\
&\leq \max_{x \in I_1} \left| \int_{x_0}^{x} L \cdot |f(t) - g(t)| dt \right| \\
&\leq |x - x_0| \cdot L \cdot \| f - g \|_{I_1} .
\end{aligned}$$

Wir müssen r_1 noch weiter verkleinern zu

$$r_2 := min\{r_1, \frac{1}{2L}\}$$

und übergehen zum Intervall

$$I_2 := [x_0 - r_2, x_0 + r_2] \subset I_1.$$

Dann haben wir bewiesen:

$$P : C^0(I_2, J) \to C^0(I_2, J)$$

ist kontrahierend mit Lipschitz-Konstante 1/2. Der Banachsche Fixpunkt-Satz liefert die Existenz- und die Eindeutigkeitsaussage. □

Beispiel 3.14 (Forster II (10.2), Kamke Nr. 32) *Schauen wir uns mal die homogene lineare Differentialgleichung*

$$y' = 2x \cdot y$$

an. Bestimmen wir mit dem Näherungsverfahren von Picard-Lindelöf die Lösung $y(x)$ mit $y(0) = c$. Es ist also $y_0 \equiv c$. Und die Integralgleichung für die sukzessive Bestimmung der Näherungen y_k ist

$$y_k(x) = c + \int_0^x 2t \cdot y_{k-1}(t) dt.$$

Damit erhalten wir

$$\begin{aligned}
y_1(x) &= c + \int_0^x 2t \cdot c \, dt \\
&= c \cdot (1 + x^2), \\
y_2(x) &= c + c \cdot \int_0^x 2t \cdot (1 + t^2) dt \\
&= c \cdot (1 + x^2 + \frac{x^4}{2}), \\
y_3(x) &= c + c \cdot \int_0^x 2t \cdot (1 + t^2 + \frac{t^4}{2}) dt
\end{aligned}$$

$$
\begin{aligned}
&= c\cdot(1+x^2+\frac{x^4}{2}+\frac{x^6}{6}),\\
&\vdots\\
y_k(x) &= c\cdot(1+x^2+\frac{x^4}{2}+\frac{x^6}{6}+...+\frac{x^{2k}}{k!})\\
&= c\cdot\sum_{\nu=0}^{k}\frac{x^{2\nu}}{\nu!}.
\end{aligned}
$$

Und die Grenzfunktion, die Lösung der Differentialgleichung, ist

$$y(x)=\lim_{k\to\infty}y_k(x)=c\cdot\sum_{\nu=0}^{\infty}\frac{x^{2\nu}}{\nu!}=c\cdot e^{2x}.$$

Man kann also das Näherungsverfahren manchmal auch zur expliziten Lösung der Differentialgleichung brauchen.

Ich möchte versuchen, die bisher skizzierte Theorie etwas plakativ zusammenzufassen: Für eine *explizite Differentialgleichung erster Ordnung*

$$y'=F(x,y)$$

sagen

Peano:
F stetig $\Longrightarrow$ lokale Existenz
Picard-Lindelöf:
F lokal Lipschitz $\Longrightarrow$ lokale Existenz globale Eindeutigkeit

Was heißt eigentlich *lokale Existenz*? Heißt das, die Mathematiker bringen halt nichts Globales, oder liegt es in der Natur der Sache? Dumme Frage, letzteres natürlich.

Beispiel 3.15 *Sehen wir uns als Beispiel mal die Differentialgleichung*

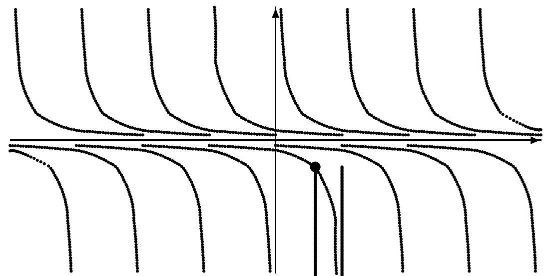
$$y'=-y^2$$

an. Sie ist nicht-linear, jedoch lokal Lipschitz, weil $F(x,y):=y^2$ nach y stetig partiell differenzierbar ist. Aber die partielle Ableitung $\partial F/\partial y=-2y$ wird für große $|y|$ selbst auch dem Betrage nach sehr groß. Dort wird dann die lokale Lipschitz-Konstante L auch immer größer. Und je größer die Lipschitz-Konstante L, desto weniger Kontrolle hat man über die Lösung.

Was sind die Lösungen? Etwa mit der Trennung der Variablen findet man

$$y(x)=\frac{1}{x+c},\quad c\in\mathbb{R}.$$

Die Lösungskurven sehen so aus:



In der Tat, je größer die y-Werte auf einer Lösungskurve, desto zielstrebiger strebt sie nach unendlich. Und wenn sie dort ist, dann existiert sie natürlich nicht mehr. Deswegen gibt es für jeden Punkt (x_0, y_0) mit $y_0 \neq 0$ eben nur ein Intervall, auf dem die Lösungsfunktion $y(x)$ mit $y(x_0) = y_0$ existiert. Dieses Intervall ist in unserem Beispiel auf der einen Seite zwar unendlich, auf der anderen aber begrenzt.

Und was heißt *lokale Eindeutigkeit*? Das heißt, wenn zwei Lösungen

$$y_1 :]a_1, b_1[\to \mathbb{R} \quad \text{und} \quad y_2 :]a_2, b_2[\to \mathbb{R}$$

in einem Punkt

$$x_0 \in]a_1, b_1[\cap]a_2, b_2[$$

denselben Wert haben, dann stimmen sie auf einem Intervall $]x_0 - r, x_0 + r[$ überein. Das ist der Eindeutigkeits-Teil des Banachschen Fixpunkt-Satzes auf dem Intervall, wo wir diesen Satz anwenden können. In Wirklichkeit gilt sogar die globale Eindeutigkeit:

Satz 3.5 (Globale Eindeutigkeit) *Mit den Voraussetzungen von Satz 3.4 seien $f_1, f_2 :]\alpha, \beta[\to \mathbb{R}$ zwei Lösungen der Differentialgleichung $f'(x) = F(x, f(x))$ mit $f_1(x_0) = f_2(x_0)$ für ein $x_0 \in]\alpha, \beta[$. Dann gilt $f_1(x) = f_2(x)$ für alle $x \in]\alpha, \beta[$.*

Beweis. Nach Satz 3.4 gibt es ein $r > 0$ so, dass $f_1(x) = f_2(x)$ für alle x mit $x_0 \leq x \leq r$. Wir betrachten die Menge

$$X := \{\xi \in [x_0, \beta[: \ f_1(x) = f_2(x) \text{ für alle } x \text{ mit } x_0 \leq x \leq \xi\}.$$

Wegen $\xi := x_0 + r \in X$ ist diese Menge nicht leer. Andererseits ist β eine obere Schranke. Deswegen existiert $x_1 := sup(X)$. Wenn $x_1 = \beta$ ist, dann stimmen die Funktionen f_1 und f_2 auf $[x_0, \beta[$ überein. Sei etwa $x_1 < \beta$. Mit Satz 3.4 gibt es ein $r_1 > 0$ so, dass f_1 und f_2 auch noch auf $[x_1, x_1 + r_1]$ übereinstimmen. Das ist ein Widerspruch zur Wahl von x_1. Also stimmen f_1 und f_2 doch auf dem ganzen Teilintervall $[x_0, \beta[$ überein.

Analog zeigt man die Aussage für das linke Teilintervall $]\alpha, x_0]$. □

Für das Angeben von Lösungen, die durch einen festen Punkt gehen, hat man noch eine eigene Redewendung geprägt, weil das halt in der Praxis so eine wichtige Frage ist:

Definition 3.5 *Eine Funktion $y(x)$ heißt Lösung des* Anfangs-Wert-Problems *(AWP)*

$$y' = F(x, y), \qquad y(x_0) = y_0,$$

wenn $y(x)$ eine Lösung der Differentialgleichung $y' = F(x, y)$ ist und im Punkt x_0 den Wert y_0 hat.

Aufgabe 3.14 *Es sei $f :]a, b[\to \mathbb{R}$ eine differenzierbare Funktion ohne Nullstelle. Zeigen Sie: Es gibt eine lineare Differentialgleichung erster Ordnung mit Lösungsraum $\{c \cdot f \, : \, c \in \mathbb{R}\}$. Welche Differentialgleichung ergibt sich für $f(x) = tan(x)$?*

Aufgabe 3.15 *Gegeben sei die Differentialgleichung*

$$(D) \qquad y'(x) = cos\left[\frac{\pi}{2}(y(x))^2\right], \quad x \in \mathbb{R}.$$

a) Man zeige, dass $y(x) := 1$ für alle $x \in \mathbb{R}$ eine Lösung von (D) mit $y(1) = 1$ ist.
b) Man zeige mit Hilfe des Existenz- und Eingeutigkeitssatzes (ohne explizite Berechnung der Lösung!), dass für jede Lösung $\varphi : \mathbb{R} \to \mathbb{R}$ der Differentialgleichung (D) mit $\varphi(0) = 0$ gilt: $\varphi(x) < 1$ für alle $x \in \mathbb{R}$. (Hinweis: Widerspruchsbeweis führen!)

Aufgabe 3.16 *Gegeben sei die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = |y|^{\frac{2}{3}}$ für jedes $(x, y) \in \mathbb{R}^2$.*
a) Beweisen Sie, dass die Funktion f für keine positive reelle Zahl b auf der Menge $\mathbb{R} \times [-b, b]$ einer Lipschitz-Bedingung in der zweiten Variablen genügt.

Überlegen Sie: Welche Konsequenzen hat dies hinsichtlich des Existenz- und Eindeutigkeitssatzes von Picard-Lindelöf für die lokale Untersuchung des AWP im Hinweis zur folgenden Teilaufgabe b)?
b) Bestimmen Sie eine Lösung $\varphi : \mathbb{R} \to \mathbb{R}$ der Differentialgleichung $y' = f(x, y)$ mit $\varphi(4) = 1$, $\varphi(0) = 0$, $\varphi(-1) = -\frac{1}{27}$.

(Hinweis: Es ist günstig, zunächst je eine lokale Lösung für die Anfangswertprobleme

$$y' = |y(x)|^{\frac{2}{3}}$$

mit den Anfangswerten

$$y(4) = 1, \quad bzw.\ y(0) = 0, \quad bzw.\ y(-1) = -\frac{1}{27}$$

zu bestimmen.)

Aufgabe 3.17 *Es sei $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ eine stetige Funktion, die lokal in $\mathbb{R}$ einer Lipschitzbedingung genüge (d.h.: Zu jeder kompakten Menge $K \subset \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ (z.B. sei K ein abgeschlossenes, beschränktes Rechteck) gibt es ein $L_K \geq 0$ mit $|f(x, y) - f(x, y')| \leq L_K |y - y'|$ für alle Punkte (x, y), $(x, y') \in K$). Weiter gelte*

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

für alle $(x, y) \in \mathbb{R} \times \mathbb{R}$.

Zeigen Sie: Ist φ eine beliebige Lösung der Differentialgleichung

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

auf einem Intervall $[-a, a]$ mit $a > 0$, d.h. $\varphi : [-a, a] \to \mathbb{R}$ differenzierbar und

$$\varphi'(x) = f(x, \varphi(x)) \qquad (x \in [-a, a]),$$

so gilt

$$\varphi(-x) = \varphi(x)$$

für alle $x \in [-a, a]$.

Hinweis: Zu einer Lösung φ betrachte man die Funktion

$$\psi(x) := \varphi(-x), \quad x \in [-a, a].$$

Aufgabe 3.18 *Bestimmen Sie die Lösung des AWP*
a) $y' = x \cdot y, \quad y(0) = 1$,
b) $y'x^2 + 2xy = \ln x \quad y(1) = 2, \quad$ auf $\quad \mathbb{R}_+ := \{x \in \mathbb{R} |\ x > 0\}$,
c) $xy' + (x-1)(1+y^2) = 0, \quad y(1) = 1, \quad$ und zeigen Sie, dass diese Lösung nicht auf das Intervall $]0, \infty[$ fortsetzbar ist,
d) $(4y + 2xy) \cdot y' = y^2 - 1, \quad y(1) = 2, \quad$ auf $]-2, \infty[$,
e) $y' = 2xy^2, \quad y(1) = a, \quad$ in den Fällen $a = 1/2$ und $a = -1/2$ und bestimmen Sie jeweils das maximale Definitionsintervall der Lösung.

Aufgabe 3.19 *Die 'logistische Gleichung'*

$$\frac{dp}{dt} = a \cdot p - b \cdot p^2, \quad 0 < a, b \in \mathbb{R},$$

beschreibt in Abhängigkeit von der Zeit t das Wachstum einer Population (z.B. Bakterien), bei der die Vermehrungsrate zur Anzahl $p(t)$ der Individuen proportional ist, aber durch einen Zusatzeffekt gedämpft wird, der zur Anzahl der Kontakte unter den Individuen proportional ist. Lösen Sie für diese Gleichung das AWP mit $p(0) = p_0 > 0$ und zeigen Sie, dass $\lim_{t\to\infty} p(t)$ von p_0 unabhängig ist.

Aufgabe 3.20 *Die Zuwachsrate einer Population (z.B. Hasen) sei proportional zur Zahl der Kontakte ihrer Individuen, die Sterberate proportional zur Zahl $p(t)$ der Individuen. Dann gilt*

$$\frac{dp}{dt} = b \cdot p^2 - a \cdot p, \quad 0 < a, b \in \mathbb{R}.$$

Zeigen Sie, dass die Population ausstirbt, wenn sie einmal auf einen Wert $p_0 < a/b$ absinkt.

3.4 Systeme von Differentialgleichungen

In der Linearen Algebra lernt man mit Systemen linearer Gleichungen umzugehen. Da hat man nicht nur eine, sondern mehrere Unbekannte. Genauso gibt es Systeme von Differentialgleichungen. Da hat man nicht mehr eine, sondern mehrere unbekannte Funktionen. Ein System von Differentialgleichungen erster Ordnung sieht so aus:

$$\begin{aligned} y_1' &= F_1(x, y_1, y_2, ..., y_n), \\ y_2' &= F_2(x, y_1, y_2, ..., y_n), \\ &\vdots \\ y_n' &= F_n(x, y_1, y_2, ..., y_n). \end{aligned}$$

Da sind also n differenzierbare Lösungsfunktionen $y_1(x), ..., y_n(x)$ zu suchen, die das Gleichungssystem erfüllen. Die Definitionsmenge G der rechten Seite ist dann eine Teilmenge $G \subset \mathbb{R}^{n+1}$, auf der alle Funktionen $F_1, ..., F_n$ stetig sind. Und eine Lösung des Systems ist ein n-tupel $(y_1, ..., y_n)$ differenzierbarer Funktionen, definiert auf einem gemeinsamen Intervall $]\alpha, \beta[\subset \mathbb{R}$, die zusammen den Gleichungen des Systems genügen.

Beispiel 3.16 *Eines der einfachsten Beispiele ist das System*

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -y_1. \end{aligned}$$

Die Definitionsmenge $G \subset \mathbb{R}^3$ ist der ganze $\mathbb{R}^3$, und eine Lösung ist z.B. das Paar

$$y_1(x) = sin(x), \quad y_2(x) = cos(x),$$

definiert auf der ganzen reellen Achse.

Die wesentlichsten Unterschiede zwischen Systemen von Differentialgleichungen und einer einzigen Differentialgleichung sind:

- Man kann sich die Lösungskurven viel schwerer vorstellen. Für $n = 2$ ist eine Lösungskurve eine Kurve im $\mathbb{R}^3$ mit den Koordinaten x, y_1, y_2. Die kann man sich noch vorstellen, aber nur schwer zeichnen. Schon für $n = 3$ hat man eine Kurve im $\mathbb{R}^4$, unvorstellbar.
- Es gibt keine allgemeinen Lösungsmethoden. Die wichtigsten Systeme sind Systeme linearer Differentialgleichungen. Die werden wir uns später anschauen. Eine Theorie, die man immer anwenden kann gibt es nur für sehr spezielle Systeme ('konstante Koeffizienten').

Es gibt aber auch ganz wesentliche Gemeinsamkeiten mit der Theorie einer gewöhnlichen Differentialgleichung. Die Existenz- und Eindeutigkeitssätze lauten - richtig interpretiert - ganz genauso, wie für eine Gleichung:

Satz 3.6 (Existenz-Satz von Peano) *Gegeben sei ein System wie oben, mit n <u>stetigen</u> Funktionen $F_1(x, y_1, ..., y_n), ..., F_n(x, y_1, ..., y_n)$ auf der rechten Seite. Dann <u>existiert</u> zu jedem Anfangswertproblem <u>lokal</u> eine Lösung.*

Unter AWP versteht man in diesem Fall folgendes: Die Funktionen $F_1, ..., F_n$ auf der rechten Seite seien definiert und stetig auf der offenen Menge $G \subset \mathbb{R}^{n+1}$. Ein Anfangswert ist ein $(n+1)$-tupel $(x_0, c_1, ..., c_n) \in G$. Eine Lösung des zugehörigen AWP ist ein n-tupel von differenzierbaren Funktionen $y_1, ..., y_n$, welche das System von Differentialgleichungen erfüllen, und der Bedingung

$$y_1(x_0) = c_1, ..., y_n(x_0) = c_n,$$

genügen.

Und dass das AWP lokal lösbar ist, heißt: Es gibt ein Intervall $]x_0 - \epsilon, x_0 + \epsilon[$ mit $\epsilon > 0$, und Funktionen $y_1, ..., y_n$, die auf diesem Intervall differenzierbar sind und das AWP lösen.

Und die Lipschitz-Bedingung verallgemeinert sich folgendermaßen:

Definition 3.6 *Das obige System von Differentialgleichungen genügt auf seinem Definitionsbereich einer Lipschitz-Bedingung mit der Lipschitz-Konstanten L, wenn für je zwei Vektoren*

$$(x, c_1, ..., c_n),\ (x, c'_1, ..., c'_n) \in G$$

gilt

$$\left\| \begin{pmatrix} F_1(x, c_1, ..., c_n) \\ \vdots \\ F_n(x, c_1, ..., c_n) \end{pmatrix} - \begin{pmatrix} F_1(x, c'_1, ..., c'_n) \\ \vdots \\ F_n(x, c'_1, ..., c'_n) \end{pmatrix} \right\| \quad \leq \quad L \cdot \left\| \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} - \begin{pmatrix} c'_1 \\ \vdots \\ c'_n \end{pmatrix} \right\|.$$

Man muss also nur die eindimensionalen Absolut-Striche durch die Vektor-Norm ersetzen.

Und genau wie in einer Dimension gilt der

Satz 3.7 *Die Funktionen $F_1, ..., F_n$ auf der rechten Seite des obigen Systems seien auf ihrem Definitionsbereich G alle stetig nach $y_1, ..., y_n$ partiell differenzierbar. Dann genügt das System auf G lokal einer Lipschitz-Bedingung.*

Das heißt: Zu jedem Punkt $(x, y_1, ..., y_n) \in G$ gibt es einen $(n+1)$-dimensionalen Würfel

$$Q = \{(\xi, \eta_1, ..., \eta_n) \in \mathbb{R}^{n+1} : |\xi - x| < \epsilon, |\eta_1 - y_1| < \epsilon, ..., |\eta_n - y_n| < \epsilon\} \subset G$$

mit $\epsilon > 0$, auf dem eine Lipschitz-Bedingung gilt.

Auch der Beweis dieses Satzes geht wie in einer Dimension (Satz 2.3), mit Hilfe des MWS für differenzierbare Funktionen. Allerdings braucht man hier die n-dimensionale Version des MWS, Satz 2.4.

Und genau wie in einer Dimension gilt auch der Existenz- und Eindeutigkeitsssatz

Satz 3.8 (Picard-Lindelöf) *Das obige System von Differentialgleichungen erfülle auf seinem Definitionsbereich G <u>lokal</u> eine <u>Lipschitz-Bedingung</u>. Dann <u>besitzt jedes</u> AWP eine <u>lokale</u> Lösung, und je zwei Lösungen desselben AWP <u>stimmen</u> auf ihrem gemeinsamen Definitionsbereich <u>überein</u>.*

Auch der Beweis dieses Satzes geht genau wie in einer Dimension (Satz 3.4). Überall muss man skalare Funktionen $F(x, y)$ durch vektorwertige Funktionen $(F_\nu(x, y_1, ..., _n))_{\nu=1,...,n}$ ersetzen. Nur an zwei Stellen muss man etwas aufpassen, bei der Definition des Integrals über vektorwertige Funktionen und bei der Integralabschätzung für vektorwertige Funktionen.

Definition 3.7 (Integral vektorwertiger Funktionen) *Die Funktionen $F_1, ..., F_n : [a, b] \to \mathbb{R}$ seien Riemann-Integrierbar. Dann definiert man*

$$\int_a^b \begin{pmatrix} F_1(x) \\ \vdots \\ F_n(x) \end{pmatrix} dx := \begin{pmatrix} \int_a^b F_1(x)dx \\ \vdots \\ \int_a^b F_n(x)dx \end{pmatrix}.$$

Das ist also einfach komponentenweise Integration und nichts wesentlich Neues. Etwas trickreich ist dagegen der Beweis für folgenden

Satz 3.9 (Integralabschätzung) *Die Abbildung $F : [a, b] \to \mathbb{R}^n$ sei integrierbar im Sinn der eben gegebenen Definition. Dann gilt*

$$\left\| \int_a^b F(x)dx \right\| \leq \left| \int_a^b \| F(x) \| \, dx \right|.$$

Das ist genau die gleiche Integralabschätzung, wie für skalare Funktionen. Nur ist der Absolutbetrag durch die Norm ersetzt.

Beweis des Satzes. Es sei $\mathbf{v} \in \mathbb{R}^n$ ein fester Vektor. Dann haben wir mit der Abschätzung für skalare Funktionen und der Cauchy-Schwarz-Ungleichung

$$\left| \int_a^b (F(x).\mathbf{v})dx \right| \leq \left| \int_a^b |(F(x).\mathbf{v})|dx \right| \leq \left| \int_a^b \| F(x) \| \cdot \| \mathbf{v} \| \, dx \right| = \left| \int_a^b \| F(x) \| \, dx \right| \cdot \| \mathbf{v} \| .$$

Spannend wird es erst, wenn wir hier für $\mathbf{v}$ den Integralvektor

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \int_a^b F_1(t)dt \\ \vdots \\ \int_a^b F_n(t)dt \end{pmatrix}$$

einsetzen. Dann ist also das Skalarprodukt

$$(F(x), \mathbf{v}) = \sum_{\nu=1}^n F_\nu(x) \cdot \int_a^b F_\nu(t)dt$$

und das Integral hierüber

$$\int_a^b (F(x).\mathbf{v})dx = \sum_{\nu=1}^n \left(\int_a^b F_\nu(x)dx \right)^2 = \left\| \int_a^b F(x)dx \right\|^2.$$

Und die eben angegebene Ungleichung wird

$$\left\| \int_a^b F(x)dx \right\|^2 \leq \left| \int_a^b \| F(x) \| \, dx \right| \cdot \left\| \int_a^b F(x)dx \right\|.$$

Die behauptete Ungleichung ergibt sich, wenn wir hier durch den zweiten Faktor auf der rechten Seite kürzen. Wenn wir nicht kürzen können, ist aber

$$\left\| \int_a^b F(x)dx \right\| = 0,$$

und die Behauptung ist trivial. □

Das ist eigentlich alles, was man ganz allgemein zu Systemen von Differentialgleichungen sagen kann. Es gibt nicht einen noch so kleinen Ansatz zu allgemeinen Lösungsmethoden. Man kann nur dann etwas machen, wenn das Differentialgleichungs-System eine spezielle Form hat. Einen dieser Spezialfälle möchte ich hier besprechen. Er ist analog zu einem linearen Gleichungs-System, dessen Koeffizientenmatrix obere Dreiecks-Form hat. Jede Funktion F_k soll also nur von $x, y_k, ..., y_n$ abhängen.

Es handelt sich also um ein Gleichungssystem der Form

$$\begin{array}{rcl} y_1' & = & F_1(x, y_1, ..., y_n), \\ y_2' & = & F_2(x, y_2, ..., y_n), \\ & \vdots & \\ y_{n-1}' & = & F_{n-1}(x, y_{n-1}, y_n), \\ y_n' & = & F_n(x, y_n). \end{array}$$

Hier kann man etwas machen: Die letzte Gleichung ist eine ganz gewöhnlich Differentialgleichung erster Ordnung für eine Funktion y_n. Die kann man (theoretisch) lösen. Und ihre Lösungen y_n kann man in die vorletzte Gleichung einsetzen. Dann wird diese eine ganz gewöhnliche Differentialgleichung für die unbekannte Funktion y_{n-1}. Und so kann man weitermachen, und sich wie beim Gauß-Algorithmus nach oben hoch-hangeln.

Nur dass es den wesentlichen Teil des Gauß-Algorithmus, die elementaren Zeilenumformungen, mit denen man ein lineares Gleichungssystem auf obere Dreiecksform bringt, bei Systemen von Differentialgleichungen nicht gibt!

Beispiel 3.17 *So, jetzt wollen wir das alles noch durch ein Beispiel vertiefen. In den Büchern findet sich keines, weil es keine sinnvolle Theorie zum Lösen nicht-linearer Systeme von Differentialgleichungen gibt. Also habe ich mir eines aus den Fingern gesogen, und versucht, alles so hinzutricksen, dass man das System explizit lösen kann:*

$$\begin{array}{rcl} y_1' & = & y_1^2 + y_1 \cdot y_2 \cdot e^{-(x^2)}, \\ y_2' & = & 2x \cdot y_2. \end{array}$$

Die zweite Gleichung, eine Gleichung für die gesuchte Funktion y_2 alleine, löst man durch Erraten

$$y_2 = c_2 \cdot e^{(x^2)}, \quad c_2 \in \mathbb{R}.$$

Dann wird die erste Gleichung

$$y_1' = y_1^2 + y_1 \cdot c_2 e^{(x^2)} \cdot e^{-(x^2)} = y_1^2 + c_2 y_1,$$

eine Bernouillesche Differentialgleichung. Wir dividieren durch y_1^2

$$\frac{y_1'}{y_1^2} = c_2 \frac{1}{y_1} + 1$$

und substituieren $z = 1/y_1$

$$-z' = c_2 z + 1, \qquad z' = -c_2 z - 1.$$

Eine spezielle Lösung $z_0 = -1/c_2$ errät man, und die allgemeine Lösung der homogenen Gleichung $z = c_1 e^{-c_2 x}$, $c_1 \in \mathbb{R}$, sieht man auch. Damit ist die allgemeine Lösung der inhomogenen Gleichung

$$z = c_1 e^{-c_2 x} - \frac{1}{c_2}, \qquad y_1 = \frac{1}{c_1 e^{-c_2 x} - 1/c_2}.$$

Natürlich muss man auch noch $c_2 = 0$ berücksichtigen, was auf $z = c_1 - x$, $y_1 = 1/(c_1 - x)$ führt.

Ich habe dieses Beispiel eigentlich nur vorgeführt, um zu zeigen, dass die Lösungsfunktion y_1 von zwei reellen Konstanten c_1 und c_2 abhängt. Weil man ja jedes AWP (lokal) lösen können muss, kann das auch gar nicht anders sein.

Wir wenden uns jetzt ab von der ganz allgemeinen Theorie, wo man kaum mehr sagen kann, als allgemeine Existenz- und Eindeutigkeitsaussätze, und wenden uns Systemen von linearen Differentialgleichungen zu, wo man zunächst auch nicht viel mehr sagen kann.

Wenn alle Differentialgleichungen eines Systems linear sind, dann sieht es so aus:

$$\begin{array}{rcl} y_1' &=& a_{1,1}(x)y_1 + a_{1,2}(x)y_2 + ... + a_{1,n}(x)y_n + b_1(x) \\ y_2' &=& a_{2,1}(x)y_1 + a_{2,2}(x)y_2 + ... + a_{2,n}(x)y_n + b_2(x) \\ &\vdots& \\ y_n' &=& a_{n,1}(x)y_1 + a_{n,2}(x)y_2 + ... + a_{n,n}(x)y_n + b_n(x) \end{array}$$

Es heißt *homogen*, wenn $b_1 = ... = b_n \equiv 0$, sonst *inhomogen*.

Die Koeffizientenfunktionen $a_{\mu,\nu}(x)$ kann man zu einer $n \times n$-Matrix

$$A(x) = \begin{pmatrix} a_{1,1}(x) & a_{1,2}(x) & ... & a_{1,n}(x) \\ a_{2,1}(x) & a_{2,2}(x) & ... & a_{2,n}(x) \\ \vdots & \vdots & & \vdots \\ a_{n,1}(x) & a_{n,2}(x) & ... & a_{n,n}(x) \end{pmatrix}$$

von Funktionen zusammenfassen, die Inhomogenität zu einem Funktionenvektor

$$\mathbf{b}(x) = \begin{pmatrix} b_1(x) \\ b_2(x) \\ \vdots \\ b_n(x) \end{pmatrix},$$

und auch die gesuchten Lösungsfunktionen zu einem Lösungsvektor

$$\mathbf{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}.$$

Dann kann man so ein System in der kompakten Form

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x) + \mathbf{b}(x)$$

schreiben. Dadurch ist es natürlich nur einfacher hinzuschreiben, nicht einfacher zu lösen.

Zu linearen Systemen von Differentialgleichungen gibt es etwas Theorie, das meiste davon wörtlich genauso, wie bei linearen Differentialgleichungen erster Ordnung, konkrete Lösungsmethoden (wie etwa die Trennung der Variablen bei linearen Differentialgleichungen erster Ordnung) gibt es nicht. Das muss man leider in aller Deutlichkeit so sagen.

Nun, welche Theorie gab es bei linearen Differentialgleichungen erster Ordnung? Antwort:

1) den Struktursatz über homogene und inhomogene Gleichungen,

2) den Existenz- und Eindeutigkeitssatz.

3) die Variation der Konstanten,

Und diese Theorie gibt es bei Systemen ziemlich wörtlich genauso.

Satz 3.10 (Struktursatz) *a) Die Lösungen, d.h. also die Vektoren* $\mathbf{y}(x)$ *von Lösungsfunktionen eines homogenen Systems*

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x)$$

linearer Diffentialgleichungen bilden einen $\mathbb{R}$*-Vektorraum.*
b) Die allgemeine Lösung eines inhomogenen Systems

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x) + \mathbf{b}(x)$$

erhält man, indem man zu einer speziellen Lösung $\mathbf{y}_0(x)$ *dieses Systems alle Lösungen des zugehörigen homogenen Systems*

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x)$$

addiert.

Dazu ist nicht viel zu sagen. Der Beweis geht genauso, wie der von Satz 3.1. Dieser Struktursatz sagt natürlich nichts darüber, ob es überhaupt Lösungen gibt, bzw., wie viele es davon gibt. (Deswegen heißt er auch Struktursatz.) Diese Frage klärt der

Satz 3.11 (Existenz- und Eindeutigkeitssatz für homogene lineare Systeme) *Jedes AWP*

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x), \qquad y_1(x_0) = c_0, ..., y_n(x_0) = c_n,$$

wo die Funktionen $a_{\mu,\nu}(x)$ *in der Matrix* $A(x)$ *auf einem Intervall* $]\alpha, \beta[$ *stetig sind, wo* $x_0 \in]\alpha, \beta[$ *beliebig ist, und wo* $\mathbf{c} = (c_1, ..., c_n)^t \in \mathbb{R}^n$ *beliebig ist, besitzt eine Lösung* $\mathbf{y}(x)$ *auf dem ganzen Intervall* $]\alpha, \beta[$*, d.h., es gibt eine differenzierbare Abbildung* $\mathbf{y} :]\alpha, \beta[\to \mathbb{R}^n$ *mit* $\mathbf{y}(x)' \equiv A(x) \cdot \mathbf{y}(x)$ *und* $\mathbf{y}(x_0) = \mathbf{c}_0$*. Diese Lösung ist durch das AWP eindeutig bestimmt.*

Na das ist doch was! Allerdings funktioniert der Beweis, den ich in 3.2.1 gegeben habe hier überhaupt nicht. Dort habe ich ja die Lösungen einfach hingeschrieben. Hier kennt man keine Lösungen (bis auf die Null-Lösung) explizit. Zunächst kann man sich auf den Existenz-Teil konzentrieren. Die globale Eindeutigkeit beweist man wie in Satz 3.5.

Wir kürzen ab: $I :=]\alpha, \beta[$. Der Definitionsbereich der rechten Seite ist die offene Menge $U = I \times \mathbb{R}^n$. Weil das für die letzten n Koordinaten keine Einschränkung bedeutet, ist die Picard-Abbildung

$$P : C^0(I, \mathbb{R}^n) \to C^0(I, \mathbb{R}^n), \quad \mathbf{y}(\mathbf{x}) \mapsto \mathbf{c} + \int_{x_0}^{x} A(t) \cdot \mathbf{y}(t) dt$$

wohldefiniert. Und wie immer beweisen wir die Existenz der Lösung $\mathbf{y}(x)$ durch Iteration der Picard-Abbildung Jede stetige Vektor-Funktion $\mathbf{y}(x)$, die Fixpunkt dieser Abbildung ist, ist differenzierbar und Lösung unseres AWP.

Natürlich müssen wir die Abschätzung

$$\| A(x) \cdot \mathbf{y}(x) \| \leq \| A(x) \| \cdot \| \mathbf{y}(x) \|$$

verwenden. Die Matrix-Norm $\| A(x) \|$ ist stetig auf I, aber dummerweise kann sie am Rand dieses Intervalls gegen ∞ gehen. Das verhindern wir, indem wir die Situation auf kompakte Teilintervalle

$$I_1 = [\alpha_1, \beta_1] \subset I \quad \text{mit} \quad \alpha < \alpha_1 < x_0 < \beta_1 < \beta$$

einschränken. Es genügt, einen Fixpunkt $\mathbf{y}(x)$ der Picard-Abbildung auf jedem solchen Teilintervall zu finden. Das Intervall I können wir nämlich durch derartige Teilintervalle ausschöpfen. Und wenn wir auf zwei derartigen kompakten Teilintervallen I_1, I_2 Fixpunkte $\mathbf{y}(x)$ haben, dann stimmen sie wegen der globalen Eindeutigkeit auf $I_1 \cap I_2$ überein. Daraus folgt die Existenz eines Fixpunkts der Picard-Abbildung P auf dem ganzen Intervall I.

Auf jedem kompakten Teilintervall $K \subset I$ ist die stetige Funktion $\| A(x) \|$ beschränkt durch die Konstante $L_K := \| A(x) \|_K$. Damit haben wir auf K die Abschätzung

$$\| (P(\mathbf{y}))(x) - (P(\mathbf{z}))(x) \| = \left\| \int_{x_0}^{x} A(t) \cdot (\mathbf{y}(t) - \mathbf{z}(t)) dt \right\| \leq$$

$$\leq \left| \int_{x_0}^{x} \| A(t) \cdot (\mathbf{y}(t) - \mathbf{z}(t)) \| \, dt \right| \leq |x - x_0| L_K \cdot \| \mathbf{y} - \mathbf{z} \|_K .$$

Weil L_K, oder $|x - x_0|$, oder beide Konstanten ziemlich groß sein können, reicht das keineswegs dafür aus, dass P kontrahierend ist. Wir müssen etwas genauer hinschauen und betrachten die Iterierten P^k der Picard-Abbildung.

Satz 3.12 (Hilfs-Abschätzung) *Für alle $k \in \mathbb{N}$ ist*

$$\| (P^k(\mathbf{y}))(x) - (P^k(\mathbf{z}))(x) \| \leq \frac{L^k \cdot |x - x_0|^k}{k!} \cdot \| \mathbf{y} - \mathbf{z} \|_K .$$

Beweis (Vollständige Induktion nach k). Für $k = 1$ ist die Aussage die soeben durchgeführte Abschätzung. Führen wir den Schluss von k auf $k + 1$ durch:

$$\begin{aligned}
\| (P^{k+1}(\mathbf{y}))(x) - (P^{k+1}(\mathbf{z}))(x) \| &= \left\| \int_{x_0}^{x} A(t) \cdot [(P^k(\mathbf{y}))(t) - (P^k(\mathbf{z}))(t)] dt \right\| \\
&\leq \left| \int_{x_0}^{x} \| A(t) \| \cdot \| (P^k(\mathbf{y}))(t) - (P^k(\mathbf{z}))(t) \| \, dt \right| \\
&\leq \left| \int_{x_0}^{x} L \cdot \frac{L^k \cdot |x - x_0|^k}{k!} \| \mathbf{y} - \mathbf{z} \|_K \, dt \right|
\end{aligned}$$

$$= \frac{L^{k+1}}{k!} \cdot \left| \int_{x_0}^{x} |t - x_0|^k dt \right| \cdot \| \mathbf{y} - \mathbf{z} \|_K$$

$$= \frac{L^{k+1} \cdot |x - x_0|^{k+1}}{(k+1)!} \cdot \| P^k(\mathbf{y}) - P^k(\mathbf{z}) \|_K . \qquad \square$$

Auf unserem kompakten Intervall K ist sicher $|x - x_0| \leq r$ für ein festes $r \in \mathbb{R}$. Wegen der Konvergenz der Exponentialreihe

$$e^{L \cdot r} = \sum_{k=0}^{\infty} \frac{L^k \cdot r^k}{k!}$$

ist sicher

$$\frac{L^k \cdot r^k}{k!} < 1$$

wenn k genügend groß ist. Für ein solches k ist dann die k-te Iterierte P^k der Picard-Abbildung kontrahierend. Nach dem Banachschen Fixpunktsatz gibt es einen Fixpunkt $\mathbf{y} \in C^0(K, \mathbb{R}^n)$ für P^k. Nun ist

$$P^k(P\mathbf{y}) = P^{k+1}(\mathbf{y}) = P(P^k(\mathbf{y})) = P(\mathbf{y}).$$

Also ist auch die Vektorfunktion $P(\mathbf{y})$ ein Fixpunkt von P^k. Wegen der Eindeutigkeit des Fixpunkts von P^k muss $P(\mathbf{y}) = \mathbf{y}$ sein. $\square$

Die Aussage des Existenz- und Eindeutigkeitssatzes kann man umformulieren, als Vektorraum-Isomorphismus ausdrücken. Die Lösungen $\mathbf{y}(x)$ einer homogenen lineare Differentialgleichung $\mathbf{y}'(x) = A(x) \cdot \mathbf{y}(x)$ bilden ja einen $\mathbb{R}$-Vektorraum, ihren Lösungsraum. Nun sei

- $A(x) = (a_{\mu,\nu}(x))$ eine $n \times n$-Matrix von auf dem Intervall $]\alpha, \beta[\subset \mathbb{R}$ stetigen Funktionen,
- V der Lösungsraum des homogenen Systems $\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x)$,
- $x_0 \in]\alpha, \beta[$ beliebig.

Dann gilt:

Satz 3.13 *Die lineare Abbildung*

$$V \to \mathbb{R}^n, \qquad \mathbf{y}(x) \mapsto \mathbf{y}(x_0)$$

ist ein Vektorraum-Isomorphismus.

In der Tat! Der Existenz-Teil des Existenz- und Eindeutigkeitssatzes sagt, dass diese Abbildung surjektiv ist, und der Eindeutigkeitsteil, dass sie injektiv ist.

Satz 3.14 (Korollar) *Der Lösungsraum eines homogenen $n \times n$-Systems von linearen Differentialgleichungen hat Dimension n.*

Beispiel 3.18 *Betrachten wir als Beispiel das System*

$$\begin{aligned} y_1' &= \quad y_2, \\ y_2' &= -y_1. \end{aligned}$$

Es hat die beiden Lösungen

$$\mathbf{y}_1(x) = \begin{pmatrix} sin(x) \\ cos(x) \end{pmatrix} \qquad und \qquad \mathbf{y}_2(x) = \begin{pmatrix} cos(x) \\ -sin(x) \end{pmatrix}.$$

Damit sind auch alle Linearkombinationen

$$c_1\mathbf{y}_1(x) + c_2\mathbf{y}_2(x) = \begin{pmatrix} c_1 sin(x) + c_2 cos(x) \\ c_1 cos(x) - c_2 sin(x) \end{pmatrix}, \qquad c_1, c_2 \in \mathbb{R},$$

Lösungen des Systems. Und das ist der ganze Lösungsraum. Äquivalent dazu ist: Die Lösungen $\mathbf{y}_1(x)$ *und* $\mathbf{y}_2(x)$ *sind linear unabhängig. Machen wir den Test: Sei*

$$c_1\mathbf{y}_1(x) + c_2\mathbf{y}_2(x) \equiv \mathbf{0}.$$

Da setzen wir $x = 0$ *ein und finden*

$$c_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Die beiden Vektorfunktionen $\mathbf{y}_1, \mathbf{y}_2 : \mathbb{R} \to \mathbb{R}^2$ *bilden also eine Basis des Lösungsraumes* V *für unser System.*

Das Auffinden aller Lösungen eines homogenen Systems ist äquivalent mit dem Auffinden einer *Basis für dessen Lösungsraum.* Weil eine Basis so wichtig ist, bekommt sie hier einen eigenen Namen:

Definition 3.8 *Ein* Lösungsfundamentalsystem *für das homogene lineare* $n \times n$ *Differentialgleichungssystem* $\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x)$ *auf dem Intervall* $]\alpha, \beta[$ *ist ein* n*-tupel von Vektorfunktionen* $\mathbf{y}_1, ..., \mathbf{y}_n :]\alpha, \beta[\to \mathbb{R}^n$*, welches eine Basis für den Lösungsraum* V *des Systems ist.*

Weil der Lösungsraum V die Dimension n hat, ist ein n-tupel von Lösungen $\mathbf{y}_1, ..., \mathbf{y}_n$ schon dann ein Lösungsfundamentalsystem, wenn dieses n-tupel linear unabhängig ist. Und dazu gibt es das folgende, sehr effektive

Satz 3.15 (Kriterium für lineare Unabhängigkeit) *Ein* n*-tupel* $\mathbf{y}_1, ..., \mathbf{y}_n$ *von Lösungen des homogenen* $n \times n$*-Systems* $\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x)$ *ist genau dann linear unabhängig (d.h. ein Lösungsfundamentalsystem), wenn für einen einzigen Punkt* x_0 *des Definitionsintervalls die Vektoren*

$$\mathbf{y}_1(x_0), ..., \mathbf{y}_n(x_0)$$

linear unabhängig sind.

Beweis. a) Seien die Lösungen $\mathbf{y}_1, ..., \mathbf{y}_n$ linear unabhängig. Wir müssen zeigen: Für jeden Punkt x_0 sind die Vektoren $\mathbf{y}_1(x_0), ..., \mathbf{y}_n(x_0) \in \mathbb{R}^n$ linear unabhängig. Dazu setzen wir den Test

$$c_1\mathbf{y}_1(x_0) + ... + c_n\mathbf{y}_n(x_0) = \mathbf{0}, \quad c_1, ..., c_n \in \mathbb{R}$$

an. Mit diesen Koeffizienten $c_1, ..., c_n$ basteln wir die Vektorfunktion

$$\mathbf{y}(x) = c_1\mathbf{y}_1(x) + ... + c_n\mathbf{y}_n(x).$$

Sie ist eine Linearkombinationen der Lösungen $\mathbf{y}_1(x), ..., \mathbf{y}_n(x)$, also selbst eine Lösung des Systems. Im Punkt x_0 hat sie den Wert $\mathbf{y}(x_0) = \mathbf{0}$. Wegen des Eindeutigkeits-Satzes folgt also daraus

$$c_1\mathbf{y}_1(x) + ... + c_n\mathbf{y}_n(x) = \mathbf{y}(x) \equiv \mathbf{0}.$$

Weil aber die Vektorfunktionen $\mathbf{y}_1, ..., \mathbf{y}_n$ linear unabhängig waren, muss $c_1 = ... = c_n = 0$ gelten.

b) Sei jetzt x_0 ein Punkt, in dem die Vektoren $\mathbf{y}_1(x_0), ..., \mathbf{y}_n(x_0) \in \mathbb{R}^n$ linear unabhängig sind. Wir müssen zeigen: auch die Vektorfunktionen $\mathbf{y}_1(x), ..., \mathbf{y}_n(x)$ sind es. Wir benutzen wieder unseren Test: Sei

$$c_1\mathbf{y}_1(x) + ... + c_n\mathbf{y}_n(x) \equiv \mathbf{0}, \quad c_1, ..., c_n \in \mathbb{R}.$$

Hier setzen wir $x = x_0$ ein und finden

$$c_1\mathbf{y}_1(x_0) + ... + c_n\mathbf{y}_n(x_0) = \mathbf{0}.$$

Weil diese Vektoren aber linear unabhängig waren, muss auch jetzt $c_1 = ... = c_n = 0$ gelten. □

Ein Lösungsfundamentalsystem kann man auffassen als eine $n \times n$-Matrix

$$Y(x) = (\mathbf{y}_1(x), ..., \mathbf{y}_n(x)) = \begin{pmatrix} y_{1,1}(x) & ... & y_{n,1}(x) \\ \vdots & & \vdots \\ y_{n,1}(x) & ... & y_{n,n}(x) \end{pmatrix}$$

von differenzierbaren Funktionen $y_{\mu,\nu}(x)$. Und das eben bewiesene Kriterium sagt

$$det(Y(x_0)) \neq 0 \quad \text{für alle} \quad x_0 \text{ im Definitionsintervall.}$$

In unserem obigen Beispiel war

$$Y(x) = \begin{pmatrix} sin(x) & cos(x) \\ cos(x) & -sin(x) \end{pmatrix}$$

und $det(Y(x)) \equiv -1$.

Mit einem Lösungsfundamentalsystem kann man *Variation der Konstanten* machen. Ist $Y(x) = (\mathbf{y}_1(x), ..., \mathbf{y}_n(x))$ ein Lösungsfundamentalsystem des Systems

$$\mathbf{y}(x)' = A(x) \cdot \mathbf{y}(x),$$

so gilt also

$$Y(x)' = A(x) \cdot Y(x)$$

und die allgemeine Lösung des homogenen Systems hat die Form

$$c_1\mathbf{y}_1(x)+...+c_n\mathbf{y}_n(x)=Y(x)\cdot\mathbf{c}$$

mit einem Vektor

$$\mathbf{c}=\begin{pmatrix}c_1\\ \vdots\\ c_n\end{pmatrix}\in\mathbb{R}^n.$$

Diesen konstanten Vektor variieren wir. Wir setzen also an

$$\mathbf{y}_0(x)=Y(x)\cdot\mathbf{c}(x)=c_1(x)\mathbf{y}_1(x)+...+c_n(x)\mathbf{y}_n(x).$$

Dieser Funktionenvektor hat den Ableitungsvektor

$$\mathbf{y}_0(x)'=Y(x)'\cdot c(x)+Y(x)\cdot\mathbf{c}(x)'=c_1(x)\mathbf{y}_1(x)'+...+c_n(x)\mathbf{y}_n(x)'+c_1(x)'\mathbf{y}_1(x)+...+c_n(x)'\mathbf{y}_n(x).$$

Wegen $Y(x)'=A(x)\cdot Y(x)$ können wir den auch

$$\mathbf{y}_0(x)'=A(x)\cdot Y(x)\cdot\mathbf{c}(x)+Y(x)\cdot\mathbf{c}(x)'=A(x)\cdot\mathbf{y}_0(x)+Y(x)\cdot\mathbf{c}(x)'$$

schreiben. Und $\mathbf{y}_0(x)$ löst das inhomogene System

$$\mathbf{y}(x)'=A(x)\cdot\mathbf{y}(x)+\mathbf{b}(x)$$

genau dann wenn

$$Y(x)\cdot\mathbf{c}(x)'=\mathbf{b}(x)$$

gilt. Und weil die Lösungsfundamentalmatrix $Y(x)$ in jedem Punkt $Y(x)$ invertierbar ist ($det(Y(x))\neq 0$), können wir sie invertieren, und diese Bedingung

$$\mathbf{c}(x)'=Y^{-1}(x)\cdot\mathbf{b}(x)$$

schreiben. Die rechte Seite dieser Gleichung ist ein Vektor aus Funktionen. Und wenn wir die alle integrieren, ja dann haben wir $\mathbf{c}(x)$ und damit die spezielle Lösung $\mathbf{y}_0(x)$.

Aufgabe 3.21 *a) Lösen Sie das Differentialgleichungssystem*

$$\begin{aligned}y_1'&=2xy_2,\\ y_2'&=-2xy_1,\end{aligned}$$

indem sie y_1y_1' und y_2y_2' vergleichen.
b) Lösen Sie das AWP $y_1(0)=1, y_2(0)=1$.

Aufgabe 3.22 *Finden Sie alle Lösungen des Differentialgleichungssystems*

$$\begin{aligned}y_1'&=cos(x)\cdot y_1+sin(x)\cdot y_2\\ y_2'&=cos(x)\cdot y_2\end{aligned}$$

Aufgabe 3.23 *Bestimmen Sie ein Lösungsfundamentalsystem für das Differentialgleichungssystem*

$$\mathbf{y}(x)' = \begin{pmatrix} 2/x & x^2 & 1 \\ 0 & -1 & 1/x \\ 0 & 0 & 1/x \end{pmatrix} \cdot \mathbf{y}(x).$$

Aufgabe 3.24 *Zeigen Sie, dass die Vektorfunktionen*

$$\mathbf{y}_1(x) = \begin{pmatrix} cosh(t) \\ sinh(t) \end{pmatrix}, \quad \mathbf{y}_2(x) = \begin{pmatrix} sinh(t) \\ cosh(t) \end{pmatrix}$$

auf jedem Intervall linear unabhängig sind, und geben Sie ein lineares 2×2*-Differentialgleichungssystem an, welches* $\mathbf{y}_1(x), \mathbf{y}_2(x)$ *als Lösungsfundamentalsystem besitzt.*

3.5 Systeme von linearen Differentialgleichungen mit konstanten Koeffizienten

In diesem Abschnitt betrachten wir Systeme von Differentialgleichungen der Form

$$\mathbf{y}'(x) = A \cdot \mathbf{y}(x),$$

wo $A \in M(n \times n, \mathbb{R})$ eine konstante $n \times n$-Matrix ist. Die kann man alle systematisch und vollständig lösen. Der entscheidende Punkt ist

Satz 3.16 *Die Spalten der Matrix* $e^{A\cdot x}$ *bilden ein Lösungsfundamentalsystem für das obige System.*

Beweis. Nach Aufgabe 2.27 ist

$$\frac{d}{dx} e^{A\cdot x} = A \cdot e^{A\cdot x}.$$

Ist $\mathbf{v} \in \mathbb{R}^n$ ein konstanter Vektor und

$$\mathbf{y}(x) = e^{A\cdot x} \cdot \mathbf{v},$$

so folgt daraus

$$\mathbf{y}'(x) = \frac{d}{dx} e^{A\cdot x} \cdot \mathbf{v} = A \cdot e^{A\cdot x} \cdot \mathbf{v} = A \cdot \mathbf{y}(x).$$

Die Vektorfunktion $\mathbf{y}(x)$ ist also eine Lösungsfunktion unseres Systems. Insbesondere sind alle Spaltenvektoren der Matrix $e^{A\cdot x}$ Lösungen. Nach Aufgabe 1.24 a) ist die Matrix $e^{A\cdot x}$ invertierbar für alle x. Ihre Spaltenvektoren sind deswegen linear unabhängig und bilden ein Lösungsfundamentalsystem. $\square$

Es kommt also darauf an, die Exponentialmatrix wirklich auszurechnen.

Beispiel 3.19 *Der einfachste Fall ist der, wo*

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$$

eine Diagonalmatrix ist. Dann ist

$$e^{A\cdot x} = \begin{pmatrix} e^{a_1\cdot x} & & \\ & \ddots & \\ & & e^{a_n\cdot x} \end{pmatrix}$$

und die Vektorfunktionen

$$\mathbf{y}_1(x) = e^{a_1\cdot x}\cdot \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, ..., \mathbf{y}_n(x) = e^{a_n\cdot x}\cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

bilden ein Lösungsfundamentalsystem.

Diese Art, Differentialgleichungssysteme zu lösen, verträgt sich glänzend mit der Ähnlichkeit von Matrizen.

Satz 3.17 *Die $n\times n$-Matrix A sei ähnlich zur Matrix $B^{-1}\cdot A\cdot B$. Ist $\mathbf{z}_1(x), ..., \mathbf{z}_n(x)$ ein Lösungsfundamentalsystem von*

$$\mathbf{z}'(x) = (B^{-1}\cdot A\cdot B)\cdot \mathbf{z},$$

so ist $\mathbf{y}_1(x) = B\cdot \mathbf{z}_1(x), ..., \mathbf{y}_n(x) = B\cdot \mathbf{z}_n(x)$ ein Lösungsfundamentalsystem von

$$\mathbf{y}'(x) = A\cdot \mathbf{y}(x).$$

Beweis. Für $\nu = 1, ..., n$ ist

$$\begin{aligned} \mathbf{z}'_\nu &= B^{-1}\cdot A\cdot B\cdot \mathbf{z}_\nu \\ B\cdot \mathbf{z}'_\nu(x) &= A\cdot (B\cdot \mathbf{z}'_\nu(x) \\ \mathbf{y}'_\nu(x) &= A\cdot \mathbf{y}_\nu(x). \end{aligned}$$

Die Vektorfunktionen $\mathbf{y}_1, ..., \mathbf{y}_n$ sind also Lösungen. Wegen

$$(\mathbf{y}_1(x), ..., \mathbf{y}_n(x)) = B\cdot (\mathbf{z}_1(x), ..., \mathbf{z}_n(x))$$

sind sie auch linear unabhängig. □

Ist insbesondere A diagonalisierbar und

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = B^{-1}\cdot A\cdot B$$

eine Diagonalmatrix, so sind die λ_ν die Eigenwerte von A und die Spaltenvektoren $\mathbf{b}_\nu$ von B zugehörige Eigenvektoren. Ein Lösungsfundamentalsystem von $\mathbf{z}' = D \cdot \mathbf{z}$ ist dann

$$e^{\lambda_1 \cdot x} \cdot \mathbf{e}_1, ..., e^{\lambda_n \cdot x} \cdot \mathbf{e}_n.$$

Und ein Lösungsfundamentalsystem von $\mathbf{y}' = A \cdot \mathbf{y}$ ist

$$e^{\lambda_1 \cdot x} \cdot \mathbf{b}_1, ..., e^{\lambda_n \cdot x} \cdot \mathbf{b}_n.$$

Das klärt die Situation wenn A diagonalisierbar ist.

Beispiel 3.20 *Die Matrix*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$

ist symmetrisch und damit diagonalisierbar. Indem man eine quadratische Gleichung löst berechnet man ihre Eigenwerte und dazugehörige Eigenvektoren

$$\lambda_1 = \frac{1}{2}(5+\sqrt{5}), \quad \mathbf{v}_1 = \begin{pmatrix} 2 \\ 1+\sqrt{5} \end{pmatrix},$$

$$\lambda_2 = \frac{1}{2}(5-\sqrt{5}), \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1-\sqrt{5} \end{pmatrix}.$$

Weil die Eigenvektoren linear unabhängig sind, erhält man das Lösungsfundamentalsystem

$$\mathbf{y}_1(x) = e^{\lambda_1 x}\mathbf{v}_1 = e^{\frac{1}{2}(5+\sqrt{5})\cdot x} \cdot \begin{pmatrix} 2 \\ 1+\sqrt{5} \end{pmatrix}, \quad \mathbf{y}_2(x) = e^{\lambda_2 x}\mathbf{v}_2 = e^{\frac{1}{2}(5-\sqrt{5})\cdot x} \cdot \begin{pmatrix} 2 \\ 2-\sqrt{5} \end{pmatrix}.$$

Aber es gibt auch Matrizen, die nicht diagonalisierbar sind. Der Musterfall ist ein Jordanblock

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

Glücklicherweise ist $e^{J_\lambda \cdot x}$ leicht auszurechenen: Wir schreiben

$$J_\lambda \cdot x = \lambda \mathbb{1} \cdot x + N \cdot x$$

mit

$$N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

Weil die Einheitsmatrix $\mathbb{1}$ mit N kommutiert, kommutiert auch $\lambda\mathbb{1} \cdot x$ mit $N \cdot x$ und nach Aufgabe 1.24 a) haben wir

$$e^{J_\lambda \cdot x} = e^{\lambda\mathbb{1}\cdot x + N\cdot x} = e^{\lambda\mathbb{1}\cdot x} \cdot e^{N\cdot x} = e^{\lambda\cdot x} \cdot e^{N\cdot x}.$$

Wir müssen also nur noch $e^{N\cdot x}$ ausrechnen. Wegen $N^n = 0$ bricht diese Exponentialreihe nach n Summanden ab und liefert

$$e^{N\cdot x} = \begin{pmatrix} 1 & x & \frac{x^2}{2} & \dots & \dots & \frac{x^{n-1}}{(n-1)!} \\ & 1 & x & \frac{x^2}{2} & & \vdots \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & 1 & x & \frac{x^2}{2} \\ & & & & 1 & x \\ & & & & & 1 \end{pmatrix}.$$

Damit erhalten wir als Lösungsfundamentalsystem die Spalten der Matrix $e^{\lambda x} \cdot e^{N\cdot x}$

$$\begin{pmatrix} e^{\lambda x} \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} xe^{\lambda x} \\ e^{\lambda x} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \begin{pmatrix} \frac{x^{n-1}}{(n-1)!}e^{\lambda x} \\ \frac{x^{n-2}}{(n-2)!}e^{\lambda x} \\ \vdots \\ xe^{\lambda x} \\ e^{\lambda x} \end{pmatrix}.$$

Mit diesem Verfahren kann man alle Systeme behandeln, deren Koeffizientenmatrix A durch eine *reelle* Koordinatentransformation auf Jordansche Normalform gebracht werden kann. Das geht genau dann, wenn das charakteristische Polynom von A lauter reelle Nullstellen hat.

Beispiel 3.21 *Der einfachste Fall, wo das nicht geht, ist eine Matrix*

$$A = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}, \quad \omega \neq 0,$$

mit dem charakteristischen Polynom $\lambda^2 + \omega^2$ ohne reelle Nullstellen. Aber die Matrix $e^{A\cdot x}$ kann man trotzdem ausrechnen:

$$\begin{aligned} e^{A\cdot x} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\omega x \\ \omega x & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} -(\omega x)^2 & 0 \\ 0 & -(\omega x)^2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 0 & (\omega x)^3 \\ -(\omega x)^3 & 0 \end{pmatrix} + \dots \\ &= \begin{pmatrix} cos(\omega x) & -sin(\omega x) \\ sin(\omega x) & cos(\omega x) \end{pmatrix}. \end{aligned}$$

Das sieht ganz anders aus, als das Verfahren mit Eigenwerten und Exponentialfunktion. Aber nach Übergang zu komplexen Eigenwerten und Eigenvektoren stellt es sich als im Wesentlichen identisch heraus. Unsere Matrix A hat die komplexen Eigenwerte und Eigenvektoren

$$\lambda_1 = i\cdot\omega, \quad \mathbf{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}, \qquad \lambda_2 = -i\cdot\omega, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ +i \end{pmatrix}.$$

Mit der Eulerschen Formel erhält man die komplexen Lösungen

$$\mathbf{z}_1(x) = e^{i\omega\cdot x}\cdot\begin{pmatrix} 1 \\ -i \end{pmatrix} = \begin{pmatrix} cos(\omega\cdot x) + i\cdot sin(\omega\cdot x \\ -i\cdot cos(\omega\cdot x) + sin(\omega\cdot x \end{pmatrix},$$

$$\mathbf{z}_2(x) = e^{-i\omega\cdot x} \cdot \begin{pmatrix} 1 \\ i \end{pmatrix} = \overline{\mathbf{z}_1(x)}.$$

Was komplexe Lösungen unseres Differentialgleichungssytems sein sollen, das definieren wir gar nicht erst, sondern gehen sofort über zu Real- und Imaginär-Teil

$$\mathbf{y}_1(x) = Re(\mathbf{z}_1(x)) = \begin{pmatrix} cos(\omega x) \\ sin(\omega x) \end{pmatrix}, \quad \mathbf{y}_2(x) = -Im(\mathbf{z}_1(x)) = \begin{pmatrix} -sin(\omega x) \\ cos(\omega x) \end{pmatrix}.$$

Das ist genau das schon vorher angegebene Lösungs-Fundamentalsystem.

Das soeben behandelte Beispiel war besonders einfach gleich in zweierlei Hinsicht:

1) Die Eigenwerte $\pm i\omega$ hatten keinen Realteil. Jedes reelle Polynom hat immer Paare konjugiert komplexer Nullstellen. Komplexe Eigenwerte einer reellen Matrix treten immer paarweise auf als

$$\lambda \pm i \cdot \omega, \quad \lambda, \omega \in \mathbb{R}.$$

Wenn der Realteil $\lambda \neq 0$ ist, dann wird die komplexe Exponentialfunktion

$$e^{(\lambda \pm i\omega)\cdot x} = e^{\lambda \cdot x} \cdot e^{\pm i\omega\cdot x} = e^{\lambda\cdot x}(cos(\omega x) \pm i \cdot sin(\omega x)).$$

Zu den Winkelfunktionen bekommt man noch eine reelle Exponentialfunktion als Faktor dazu. Das ist nur eine unwesentliche Komplikation. Sie macht das Leben einfach etwas reichhaltiger.

2) Die komplexen Eigenwerte traten nur einfach auf. Bei mehrfachen Linearfaktoren im charakteristischen Polynom braucht die Matrix auch komplex nicht mehr diagonalisierbar zu sein. Das Führt auf die Jordansche Normalform im Komplexen.

Dises zweite angesprochen Problem möchte ich noch diskutieren, allerdings nicht in voller Allgemeinheit. Der wesentliche Punkt ist folgender: Echt komplexe Eigenwerte führen auf echt komplexe Jordan-Blöcke. Aber wie die Eigenwerte treten auch die Jordan-Blöcke als komplex-konjugierte Paare auf. Solche Paare kan man zusammenfassen und auf eine 'reelle Jordan-Normalform' kommen. Die Theorie gehört in die Lineare Algebra und ich will deswegen hier nur den aller-einfachsten Fall behandeln:

Ich nehmen an, die reelle Matrix A habe den komplexen Eigenwert $\lambda + i\omega, \omega \neq 0$, mit einer Vielfachheit n und dazu gehöre nur ein einziger komplexer $n \times n$ Jordan-Block $J_{\lambda+i\omega}$. Um die Matrix A in komplexe Jordan-Normalform zu transformieren verwendet man n linear unabhängige Vektoren (eine 'Kette')

$$\mathbf{v}, (A - (\lambda + i\omega)\mathbb{1})\mathbf{v}, ..., (A - (\lambda + i\omega)\mathbb{1})^{n-1}\mathbf{v},$$

wobei

$$(A - (\lambda + i\omega)\mathbb{1})^n \mathbf{v} = \mathbf{0}$$

ist. Es gibt noch den komplex-konjugierten Eigenwert $\lambda - i\omega$. Und es gibt dazu die komplex-konjugierte Kette

$$\mathbf{v}, (A - (\lambda - i\omega)\mathbb{1})\mathbf{v}, ..., (A - (\lambda + i\omega)\mathbb{1})^{n-1}\mathbf{v}.$$

Sie führt auf einen $n \times n$-Jordan-Block $J_{\lambda - i\omega}$. Ich nehme weiter, vereinfachend, an, dass

$$J := \begin{pmatrix} J_{\lambda+i\omega} & 0 \\ 0 & J_{\lambda-i\omega} \end{pmatrix}$$

schon die ganze Jordan-Normalform der Matrix A ist. Dann hat die Matrix A also das Format $2n \times 2n$.

In die komplexe Jordan-Normalform transformiert man mit einer komplexen $2n \times 2n$-Matrix, deren Spaltenvektoren die $2n$ Vektoren der beiden komplex-konjugierten Ketten sind. Diese Transformations-Matrix hat also die Form

$$(B, \overline{B}), \quad B \in M(2n \times n, \mathbb{C}).$$

Und es ist

$$(B, \overline{B})^{-1} \cdot A \cdot (B, \overline{B}) = J.$$

Das Resultat ist zu komplex. Wir gehen über zu Real- und Imaginär-Teil der Jordan-Normalform in Matrizen-Schreibweise. Dazu verwenden wir die komplexe $2n \times 2n$-Transformationsmatrix

$$X := \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} \mathbb{1}_n & i \cdot \mathbb{1}_n \\ \mathbb{1}_n & -i \cdot \mathbb{1}_n \end{pmatrix} \quad \text{mit} \quad X^{-1} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} \mathbb{1}_n & \mathbb{1}_n \\ -i \cdot \mathbb{1}_n & i \cdot \mathbb{1}_n \end{pmatrix}.$$

Damit wird

$$X^{-1} \cdot J \cdot X = \begin{pmatrix} Re(J_{\lambda+i\omega}) & -Im(J_{\lambda+i\omega}) \\ Im(J_{\lambda+i\omega}) & Re(J_{\lambda+i\omega}) \end{pmatrix} = \begin{pmatrix} J_\lambda & -\omega\mathbb{1} \\ \omega\mathbb{1} & J_\lambda \end{pmatrix}.$$

Diese reelle Matrix heißt die *reelle Jordan-Normalform* von A. Sie ist die zu A ähnliche Matrix

$$X^{-1} \cdot (B, \overline{B})^{-1} \cdot A \cdot (B, \overline{B}) \cdot X = \Big((B, \overline{B}) \cdot X\Big)^{-1} \cdot A \cdot \Big((B, \overline{B}) \cdot X\Big).$$

Wesentlich ist hier, dass die verwendete Transformationsmatrix

$$(B, \overline{B}) \cdot X = \frac{1}{\sqrt{2}} \cdot (B + \overline{B}, i \cdot (B - \overline{B}))$$

reell ist. Man kann von A übergehen zu dessen reeller Jordan-Normalform durch eine reelle Transformation im $\mathbb{R}^{2n}$. Nach dieser Transformation genügt es also, das Differential-Gleichungssystem mit Koeffizientenmatrix

$$\begin{pmatrix} J_\lambda & -\omega\mathbb{1} \\ \omega\mathbb{1} & J_\lambda \end{pmatrix} = \begin{pmatrix} J_\lambda & 0 \\ 0 & J_\lambda \end{pmatrix} + \begin{pmatrix} 0 & -\omega\mathbb{1} \\ \omega E & 0 \end{pmatrix}$$

zu behandeln.

Und wieder einmal fügt sich alles am Ende sehr harmonisch: Man rechnet sehr leicht nach, dass die beiden Matrizen auf der rechten Seite der obigen Gleichung kommutieren. Ein Lösungsfundamentalsystem für dieses System sind also die Spalten der Matrix

$$\begin{aligned} exp \begin{pmatrix} J_\lambda \cdot x & -\omega\mathbb{1} \cdot x \\ \omega\mathbb{1} \cdot x & J_\lambda \cdot x \end{pmatrix} &= exp \begin{pmatrix} (\lambda\mathbb{1} + N) \cdot x & 0 \\ 0 & (\lambda\mathbb{1} + N) \cdot x \end{pmatrix} \cdot exp \begin{pmatrix} 0 & -\omega\mathbb{1} \cdot x \\ \omega\mathbb{1} \cdot x & 0 \end{pmatrix} \\ &= e^{\lambda \cdot x} \cdot \begin{pmatrix} cos(\omega \cdot x) \cdot e^{N \cdot x} & -sin(\omega \cdot x) \cdot e^{N \cdot x} \\ sin(\omega \cdot x) \cdot e^{N \cdot x} & cos(\omega \cdot x) \cdot e^{N \cdot x} \end{pmatrix} \end{aligned}$$

Dabei ist N die nilpotente Matrix mit Einsern auf der Nebendiagonale und $e^{N \cdot x}$ ist die weiter oben angegebene Matrix mit Polynom-Einträgen bis zum Grad $n-1$. Die Komponenten der Lösungsvektoren $\mathbf{y}_1(x), ..., \mathbf{y}_{2n}(x)$ sind also Produkte

- der Exponentialfunktion $e^{\lambda \cdot x}$,

- von Winkelfunktionen $cos(\omega \cdot x)$, $sin(\omega \cdot x)$,
- Polynomen in x vom Grad $\leq n-1$.

Aufgabe 3.25 *Es seien $a, b \in \mathbb{R}$.*
a) Finden Sie ein Lösungsfundamentalsystem für das System

$$\begin{aligned} y_1' &= a \cdot y_1 + b \cdot y_2, \\ y_2' &= b \cdot y_1 + a \cdot y_2. \end{aligned}$$

b) Finden Sie eine spezielle Lösung des Systems

$$\begin{aligned} y_1' &= a \cdot y_1 + b \cdot y_2 + (1-a) \cdot e^x, \\ y_2' &= b \cdot y_1 + a \cdot y_2 - b \cdot e^x. \end{aligned}$$

Aufgabe 3.26 *Finden Sie alle Lösungen des Systems*

$$\begin{aligned} y_1' &= y_2 + x, \\ y_2' &= y_1 + x. \end{aligned}$$

Aufgabe 3.27 *Finden Sie die allgemeine Lösung des Systems*

$$\mathbf{y}' = \begin{pmatrix} 1 & 12 \\ 3 & 1 \end{pmatrix} \cdot \mathbf{y} + \begin{pmatrix} e^x \\ 0 \end{pmatrix}.$$

Aufgabe 3.28 *Lösen Sie die AWProbleme*
a) $\mathbf{y}' = \begin{pmatrix} 1 & -1 \\ -3 & 0 \end{pmatrix} \cdot \mathbf{y}, \quad \mathbf{y}(0) = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$ *b)* $\mathbf{y}' = \begin{pmatrix} -3 & 5 \\ -2 & 4 \end{pmatrix} \cdot \mathbf{y}, \quad \mathbf{y}(0) = \begin{pmatrix} 13 \\ 1 \end{pmatrix}.$

Aufgabe 3.29 *Finden Sie ein Lösungsfundamentalsystem für*

$$\begin{aligned} y_1' &= - y_2 + y_3 \\ y_2' &= y_1 - y_3 \\ y_3' &= -y_1 + y_2 \end{aligned}$$

3.6 Differentialgleichungen höherer Ordnung

Eine Differentialgleichung höherer Ordnung ist eine Differentialgleichung einer Ordnung ≥ 2. Sie enthält nicht nur die erste Ableitung der gesuchten Funktion, sondern auch Ableitungen dieser Funktion von einer Ordnung ≥ 2. Eine Differentialgleichung n-ter Ordnung hat also die Form

$$F(x, y, y', ..., y^{(n)}) = 0.$$

Dabei ist die Funktion F definiert auf einer offenen Menge $U \subset \mathbb{R}^{n+2}$. Wenn F stetig partiell differenzierbar ist und die partielle Ableitung von F nach der letzten Variablen nicht verschwindet, kann man lokal nach $y^{(n)}$ auflösen (Satz 2.17) und erhält eine Gleichung

$$y^{(n)}(x) = g(x, y, y', ..., y^{(n-1)}).$$

Sowas heißt eine *explizite Differentialgleichung höherer Ordnung.*

Satz 3.18 *Eine explizite Differentialgleichung höherer Ordnung*

$$y^{(n)}(x) = g(x, y, y', ..., y^{(n-1)})$$

ist äquivalent zu dem System

$$\begin{array}{rcl} z_1' & = & z_2 \\ z_2' & = & \quad z_3 \\ \vdots & \vdots & \qquad \ddots \\ z_{n-1}' & = & \qquad\qquad z_n \\ z_n' & = & g(x, z_1, z_2, ..., z_{n-1}\ ) \end{array}$$

Beweis. Man setzt

$$z_1 = y, \quad z_2 = y', \quad ..., \quad z_n = y^{(n-1)}$$

und kann damit zwischen der Differentialgleichung n-ter Ordnung und dem System hin- und hertransformieren. □

Damit kann man zunächst die allgemeine Theorie der Systeme auf Differentialgleichungen höherer Ordnung umformulieren:

Satz 3.19 *Die Funktion $g(x, y_1, ..., y_n)$ sei stetig, und stetig partiell differenzierbar nach $y_1, ..., y_n$ auf $U \subset \mathbb{R}^{n+1}$. Dann gilt*
a) (Lokale Existenz) Es sei $(x, c_0, c_1, ..., c_{n-1}) \in U$. Dann gibt es ein $r > 0$ und eine n-mal differenzierbare Funktion $y :]x_0 - r, x_0 + r[\to \mathbb{R}$ mit

$$y^{(n)}(x) = g(x, y(x), y'(x), ..., y^{(n-1)}(x)) \quad \textit{für alle} \quad x \in]x_0 - r, x_0 + r[$$

und

$$y(x_0) = c_0, \quad y'(x_0) = c_1, \quad , ..., \quad y^{(n-1)}(x_0) = c_{n-1}.$$

b) (Globale Eindeutigkeit) Zwei Lösungen $y(x)$ und $z(x)$ der Differentialgleichung mit

$$y(x_0) = z(x_0),\ y'(x_0) = z'(x_0),\ ...,\ y^{(n-1)}(x_0) = z^{(n-1)}(x_0)$$

stimmen auf ihrem gemeinsamen Definitonsintervall überein.

Mehr kann man in dieser Allgemeinheit nicht sagen. Wenden wir uns deswegen den linearen Differentialgleichungen höherer Ordnung zu.

Definition 3.9 *Eine* lineare Differentialgleichung der Ordnung n *für eine Funktion* y *hat die Form*

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + ... + a_1(x)y' + a_0(x)y = b(x).$$

Dabei sind die Koeffizienten-Funktionen $a_\nu(x)$ *und die rechte Seite* $b(x)$ *stetig auf einem Definitionsintervall* $]\alpha, \beta[\subset \mathbb{R}$. *Die Differentialgleichung heißt* homogen, *wenn* $b(x) \equiv 0$, *sonst* inhomogen.

Wo der Leitkoeffizient $a_n(x) \neq 0$ ist, kann man die Gleichung durch ihn dividieren, und danach dann also $a_n(x) \equiv 1$ annehmen. Die Differentialgleichung ist explizit geworden. Hat der Leitkoeffizient Nullstellen, so teilen diese das Definitionsintervall in Teilintervalle, innerhalb deren $a_n(x) \neq 0$ ist, und wo man deshalb dividieren kann. Die im folgenden beschriebene Theorie funktioniert nur auf diesen Teilintervallen.

Zunächst einmal ist das System, das zu einer linearen Differentialgleichung höherer Ordnung gehört, wieder linear.

Satz 3.20 *Eine lineare Differentialgleichung*

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + ... + a_1(x)y' + a_0(x)y = b(x)$$

n*-ter Ordnung ist äquivalent zu einem System von* n *linearen differentialgleichungen der Form*

$$\begin{array}{lcccccccccc} z_1' & = & & & z_2 & & & & & & \\ z_2' & = & & & & & z_3 & & & & \\ & \vdots & & & & & & \ddots & & & \\ z_{n-1}' & = & & & & & & & z_n & & \\ z_n' & = & -a_0 z_1 & - & a_1 z_2 & - & ... & - & a_{n-1} z_n & - & b \end{array}$$

Dieses System hat eine ganz besondere Form: Die ersten $n-1$ Zeilen bedeuten nichts anderes als

$$z_2 = z_1', \ z_3 = z_1'', \ ..., \ z_n = z^{(n-1)}.$$

Und die letzte Zeile alleine ist dann die Differentialgleichung n-ter Ordnung für z_1. Also gehört auch jedes System von linearen Gleichungen, das diese spezielle Form hat, zu einer linearen Differentialgleichung n-ter Ordnung. Das System von linearen Gleichungen für $z_1, ..., z_n$ und die Gleichung n-ter Ordnung für y sind äquivalent.

Beispiel 3.22 *Betrachten wir als Beispiel etwa die Differentialgleichung zweiter Ordnung*

$$x^2 f'' + x f' + f = 0.$$

Nach der Division durch x^2 *wird sie*

$$f'' + \frac{1}{x} f' + \frac{1}{x^2} f = 0.$$

Das zugehörige 2×2*-System ist*

$$\begin{array}{lcccc} y_1' & = & & & y_2, \\ & & & & \\ y_2' & = & -\frac{1}{x^2} y_1 & - & \frac{1}{x} y_2. \end{array}$$

Weil man Systeme linearer Differentialgleichungen i.A. nicht explizit lösen kann (außer, sie haben konstante Koeffizienten), kann man auch Differentialgleichungen höherer Ordnung i.A. nicht explizit lösen (außer, sie haben konstante Koeffizienten). Das ist die schlechte Nachricht. Aber es gibt auch gute Nachrichten:

Satz 3.21 (Struktursatz) *a) Die Lösungen $y :]\alpha, \beta[\to \mathbb{R}$ einer homogenen linearen Differentialgleichung n-ter Ordnung bilden einen $\mathbb{R}$-Vektorraum der Dimension n.*

b) Die allgemeine Lösung einer inhomogenen Differentialgleichung n-ter Ordnung erhält man, indem man zu einer speziellen Lösung y_0 der inhomogenen Gleichung die allgemeine Lösung der homogenen Gleichung addiert.

c) (Globale Existenz) Ist $I \subset \mathbb{R}$ das maximale Definitonsintervall, auf dem alle Koeffizienten $a_0, ..., a_{n-1}, b$ stetig sind, so ist jede lokale Lösung der Gleichung ganz auf dieses Intervall fortsetzbar.

Außerdem gilt der AWP-Satz in folgender Form:

In jedem Punkt x_0 des Definitionsintervalls einer linearen Differentialgleichung n-ter Ordnung kann man Anfangswerte $c_0, c_1, ..., c_{n-1} \in \mathbb{R}$ vorgeben. Es gibt dann immer eine eindeutig bestimmte Lösung y dieser Differentialgleichung mit diesen Anfangswerten:

$$y(x_0) = c_0,\ y'(x_0) = c_1,\ ...,\ y^{(n-1)}(x_0) = c_{n-1}.$$

Für homogene lineare Differentialgleichungen bilden die Lösungen einen Lösungsraum V. Für jedes x_0 in Definitons-Intervall ist die $\mathbb{R}$-lineare Abbildung

$$V \to \mathbb{R}^n, \quad y \mapsto \begin{pmatrix} y(x_0) \\ y'(x_0) \\ \vdots \\ y^{(n-1)}(x_0) \end{pmatrix}$$

ein Isomorphismus.

Diese Sätze sind ganz einfache Übersetzungen der entsprechenden Aussagen für Systeme linearer Differentialgleichungen erster Ordnung in die Sprache der Differentialgleichungen n-ter Ordnung. Ich möchte aber nochmal ausdrücklich darauf hinweisen, dass sie nur für Gleichungen der Form

$$f^{(n)}(x) + a_{n-1}(x) f^{(n-1)}(x) + ...$$

gelten, wo der Leitkoeffizient $= 1$ ist. Bei Gleichungen der Form

$$a_n(x) f^{(n)}(x) + a_{n-1}(x) f^{(n-1)}(x) + ...$$

gelten Sie nicht. Man muss erst durch $a_n(x)$ dividieren. Deswegen muss man die Nullstellen der Funktion $a_n(x)$ ausschließen, und eventuell zu Teilintervallen des Definitionsintervalls $]\alpha, \beta[$ übergehen.

Wir müssen jetzt noch den Begriff des Lösungsfundamentalsystems von den Systemen erster Ordnung in die Sprache der Gleichungen n-ter Ordnung übersetzen. Denn das Lösen eines homogenen Differentialgleichungs-Systems bedeutet, ein solches Lösungsfundamentalsystem anzugeben. Das sind Lösungsvektoren $\mathbf{z}_1(x), ..., \mathbf{z}_n(x)$, die in jedem Punkt des Definitions-Intervalls linear unabhängig sind.

Zu n Lösungsvektoren

$$\mathbf{z}_1 = \begin{pmatrix} y_1 \\ \vdots \\ y_1^{(n-1)} \end{pmatrix}, \quad ..., \quad \mathbf{z}_n = \begin{pmatrix} y_n \\ \vdots \\ y_n^{(n-1)} \end{pmatrix}$$

des Systems gehören n Lösungen $y_1, ..., y_n$ der Gleichung n-ter Ordnung. Dass die Lösungsvektoren in jedem Punkt x_0 linear unabhängig sind, bedeutet: Die $n \times n$ Fundamentalmatrix

$$W(x) = \begin{pmatrix} y_1(x) & y_2(x) & ... & y_n(x) \\ y_1'(x) & y_2'(x) & ... & y_n'(x) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & ... & y_n^{(n-1)}(x) \end{pmatrix}$$

hat in jedem Punkt x_0 den Rang n. Diese Matrix heißt übrigens *Wronski-Matrix*

$$W(y_1, ..., y_n)(x).$$

Mann nennt n Lösungsfunktionen $y_1, ..., y_n$ einer homogenen linearen Differentialgleichung n-ter Ordnung ein *Lösungsfundamentalsystem*, wenn die entsprechenden Lösungsvektoren des zugehörigen homogenen $n \times n$-Systems erster Ordnung ein Lösungsfundamentalsystem bilden. Das Lösen einer Differentialgleichung n-ter Ordnung ist äquivalent damit, so ein Lösungsfundamentalsystem anzugeben. Deswegen muss man wissen, wann Lösungen $y_1, ..., y_n$ ein solches bilden

Satz 3.22 (Kriterium) *Für n Lösungen $y_1, ..., y_n$ einer homogenen linearen Differentialgleichung n-ter Ordnung*

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + ... + a_0(x)y = 0$$

sind äquivalent:

1) Sie bilden ein Lösungsfundamentalsystem.

2) Die Wronski-Determinante $det(W(x))$ hat keine Nullstellen.

3) Es gibt einen Punkt x_0 des Definitions-Intervalls, in dem $det(W(x_0)) \neq 0$ ist,

4) Die Funktionen $y_1, ..., y_n$ sind linear unabhängig.

Beweis. Die Spalten der Wronski-Matrix $W(x)$ sind Lösungsvektoren des $n \times n$-Systems erster Ordnung. Deswegen folgt aus unserem Kriterium für Systeme erster Ordnung

$$1) \Leftrightarrow 2) \Leftrightarrow 3)$$

Wir zeigen 1) $\Rightarrow$ 4): Seien die Funktionen linear abhängig, es gebe also $c_1, ..., c_n \in \mathbb{R}$ mit

$$c_1 y_1(x) + ... + c_n y_n(x) \equiv 0.$$

Durch Differenzieren folgt daraus

$$\begin{aligned} c_1 y_1'(x) + ... + c_n y_n'(x) &= 0 \\ c_1 y_1''(x) + ... + c_n y_n''(x) &= 0 \\ &\vdots \\ c_1 y_1^{(n-1)}(x) + ... + c_n y_n^{(n-1)}(x) &= 0 \end{aligned}$$

Für die Wronski-Matrix hat das die Konsequenz

$$W(x) \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \mathbf{0}.$$

Weil diese Matrix aber in jedem Punkt x Maximalrang hat, muss

$$c_1 = ... = c_n = 0$$

gelten.

Seien jetzt umgekehrt die Funktionen $y_1, ..., y_n$ linear unabhängig. Wenn die Wronski-Determinante in einem Punkt x_0 verschwinden würde, gäbe es also Koeffizienten $c_1, ..., c_n \in \mathbb{R}$, nicht alle $= 0$ mit

$$W(x_0) \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \mathbf{0}.$$

Der Lösungsvektor

$$\mathbf{z}(x) = c_1\mathbf{z}_1(x) + ... + c_n\mathbf{z}_n(x)$$

des $n \times n$-Systems hätte also die Nullstelle x_0. Wegen des Eindeutigkeitssatzes für lineare Systeme würde er identisch verschwinden:

$$c_1\mathbf{z}_1(x) + ... + c_n\mathbf{z}_n(x) \equiv 0.$$

Die erste Zeile dieser Vektorgleichung lautet

$$c_1 y_1(x) + ... + c_n y_n(x) \equiv 0.$$

Das würde bedeuten, die Funktionen $y_1, ..., y_n$ sind linear abhängig. Widerspruch! $\square$

Falls man eine nicht-triviale Lösung y_1 einer Differentialgleichung n-ter Ordnung kennt, so gibt es ein Verfahren, weitere Lösungen zu ermitteln, indem man nur noch eine Gleichung $n-1$-ter Ordnung löst. Dieses Verfahren heißt *Reduktion der Ordnung* und ist von beschränktem Wert. Denn eine erste Lösung muss man erstmal haben. Ich möchte es nur für Gleichungen zweiter Ordnung besprechen:

Sei also die Differentialgleichung

$$y'' + a_1(x)y' + a_0(x)y = 0$$

gegeben, von der wir eine Lösung $y_1 \not\equiv 0$ kennen. Um eine weitere Lösung zu finden macht man den Ansatz

$$y(x) = g(x)y_1(x).$$

Die ersten beiden Ableitungen von einem solchen y sind

$$\begin{aligned} y' &= g'y_1 + gy_1', \\ y'' &= g''y_1 + 2g'y_1' + gy_1''. \end{aligned}$$

Dass y eine Lösung der Differentialgleichung ist, bedeutet deswegen

$$\begin{array}{rcl} g''y_1 + 2g'y_1' + gy_1'' & + & \\ a_1[g'y_1 + gy_1'] & + & \\ a_0 g y_1 & = & 0, \\ g''y_1 + g'(2y_1' + a_1 y_1) + g(y_1'' + a_1 y_1' + a_0 y_1) & = & \\ g''y_1 + g'(2y_1' + a_1 y_1) & = & 0. \end{array}$$

Wo $y_1 \neq 0$ ist, dort ist $u := g'$ eine Lösung der homogenen linearen Differentialgleichung

$$u' + \left(2\frac{y_1'}{y_1} + a_1\right) u = 0.$$

Die kann man lösen und bekommt mit $y_2 := (\int u) \cdot y_1$ eine weiter Lösung der Differentialgleichung zweiter Ordnung. Natürlich muss man eine Lösung $u \neq 0$ nehmen. Dann ist $\int u \neq \mathit{const}$ und y_2 ist linear unabhängig von y_1.

Für inhomogene Differentialgleichungen n-ter Ordnung gibt es, genauso wie für Systeme, das Verfahren der Variation der Konstanten, um eine spezielle Lösung der inhomogenen Gleichung zu ermitteln, wenn man ein Lösungsfundamentalsystem der homogenen Gleichung kennt. Ich möchte aber die Formeln dafür nicht aus der Sprache der Systeme erster Ordnung in die Sprache der Gleichung n-ter Ordnung übersetzen. Die Formeln sind mir einfach zu kompliziert. Es ist einfacher, von der Gleichung n-ter Ordnung zum System überzugehen, dort die Variation der Konstanten zu machen, und zurück zu transformieren.

Es gibt eine einzige Sorte von linearen Differentialgleichungen mit nicht-konstanten Koeffizienten, bei denen es sich lohnt, sich einen Lösungsansatz zu merken: die *Eulerschen Differentialgleichungen*

$$x^n y^{(n)} + a_{n-1} x^{n-1} y^{(n-1)} + ... + a_1 x y' + a_0 y = 0, \quad a_{n-1}, ..., a_1, a_0 \in \mathbb{R}.$$

Hier führt der Ansatz

$$y(x) = x^r$$

zu einer Lösung. Denn damit wird

$$\begin{array}{lcllcl} y' & = & r x^{r-1} & xy' & = & r x^r \\ y'' & = & r(r-1)x^{r-2} & x^2 y'' & = & r(r-1)x^r \\ & \vdots & & & \vdots & \\ y^{(n)} & = & r(r-1)...(r-n+1)x^{r-n} & x^n y^{(n)} & = & r(r-1)...(r-n+1)x^r \end{array}$$

Aus der ganzen Differentialgleichung kann man dann x^r ausklammern

$$[r(r-1)...(r-n+1) + ... + a_2 r(r-1) + a_1 r + a_0] x^r = 0$$

und kommt zu einer Polynomgleichung vom Grad n für r. Hat diese r verschiedene reelle Lösungen $r_1, ..., r_n$, so ergeben sich daraus n verschiedene Lösungen $x^{r_1}, ..., x^{r_n}$ der Eulerschen Differentialgleichung. Aber dabei können eine ganze Reihe von Komplikationen auftreten.

1) Die Exponenten $r_1, ..., r_n$ brauchen keine ganzen Zahlen zu sein. Dann ist x^{r_ν} nur für $x > 0$ definiert.

2) Es können einige der Exponenten r_ν zusammenfallen. Dann bekommen wir auf diese Weise weniger als n Lösungen.

3) Es können komplexe Exponenten r_ν herauskommen.

Im Fall 1) kann man sich aus der Affäre ziehen, indem man die Intervalle $x > 0$ und $x < 0$ gesondert behandelt. Durch eine Transformation $x = -t$ wird der Fall $x < 0$ auf den Fall $x > 0$ zurückgeführt.

Für die Komplikationen 2) und 3) gibt es Formeln. Die sind schwer zu merken. Am einfachsten ist es in diesen Fällen, nicht den Ansatz $y = x^r$ zu machen, sondern die Transformation $x = e^t$. Dies führt dann auf eine Differentialgleichung mit konstanten Koeffizienten, wie wir sie uns anschließend ansehen.

Sind die ermittelten Koeffizienten $r_1, ..., r_n$ allerdings alle reell und voneinander verschieden, so sind die Lösungen $x^{r_1}, ..., x^{r_n}$ linear unabhängig und bilden ein Lösungsfundamentalsystem.

Beweis. Es seien also die Exponenten $r_1, ..., r_n$ alle voneinander verschieden. Wenn die Funktionen $x^{r_1}, ..., x^{r_n}$ linear abhängig wären, dann gäbe es eine lineare Relation

$$c_1 x^{r_1} + ... + c_n x^{r_n} \equiv 0, \quad c_1, ..., c_n \in \mathbb{R}.$$

Wir dividieren diese Relation durch x^{r_n}

$$c_1 x^{r_1 - r_n} + ... + c_{n-1} x^{r_{n-1} - r_n} + c_n \equiv 0$$

und differenzieren:

$$c_1(r_1 - r_n) x^{r_1 - r_n - 1} + ... + c_{n-1} x^{r_{n-1} - r_n - 1} \equiv 0.$$

Hier haben wir die $n-1$ Exponenten

$$r_1 - r_n - 1, ..., r_{n-1} - r_n - 1.$$

Weil die r_ν alle voneinander verschieden sind, sind auch diese $n-1$ Exponenten alle voneinander verschieden. Wir machen deshalb vollständige Induktion nach n und können annehmen, dass die Funktionen

$$x^{r_1 - r_n - 1}, ..., x^{r_{n-1} - r_n}$$

linear unabhängig sind. Dann folgt

$$c_1(r_1 - r_n) = ... = c_{n-1}(r_{n-1} - r_n) = 0$$

und weil $r_1 \neq r_n, ..., r_{n-1} \neq r_n$, auch

$$c_1 = ... = c_{n-1} = 0.$$

Also ist unsere ursprüngliche Relation $c_n x^{r_n} \equiv 0$. Hier setzen wir $x = 1$ ein und finden $c_n = 0$. □

Beispiel 3.23 *Jetzt noch ein Beispiel:*

$$x^2 y'' + xy' - y = 0.$$

Mit

$$y = x^r, \quad y' = r x^{r-1}, \quad y'' = r(r-1) x^{r-2}$$

wird daraus

$$r(r-1)x^r + rx^r - x^r = 0.$$

Nach Division durch x^r erhalten wir die Polynomgleichung

$$r(r-1) + r - 1 = r^2 - 1 = 0$$

für r. Sie hat die Lösungen $r_1 = 1$ und $r_2 = -1$. Die führen auf die Funktionen

$$y_1(x) = x, \quad y_2(x) = \frac{1}{x},$$

welche die gegebene Eulersche Differentialgleichung lösen.

Ich habe es natürlich (durch das Minus-Zeichen vor y) so hingetrickst, dass aus der Polynomgleichung zwei reelle Lösungen rauskamen. Für die Eulersche Gleichung

$$x^2 f'' + xf' + f = 0$$

wäre die Exponentengleichung

$$r^2 + 1 = 0$$

gewesen, mit den beiden komplexen Lösungen $\pm i$. Wenn man die Eulersche Formel

$$e^{i\varphi} = cos(\varphi) + isin(\varphi)$$

anwendet, kann man das bearbeiten:

$$x^i = e^{i \cdot ln(x)} = cos(ln(x)) + isin(ln(x)), \quad x^{-i} = e^{-i \cdot ln(x)} = cos(ln(x)) - isin(ln(x)).$$

Daraus kann man die beiden reellen Lösungen

$$\frac{1}{2}(x^i + x^{-i}) = cos(ln(x)), \quad \frac{1}{2i}(x^i - x^{-i}) = sin(ln(x))$$

herleiten.

Schließlich betrachten wir Differentialgleichungen n-ter Ordnung

$$f^{(n)} + a_{n-1} f^{(n-1)} + ... + a_1 f' + a_0 f = b(x)$$

mit konstanten Koeffizienten $a_0, ..., a_{n-1} \in \mathbb{R}$. Zunächst behandeln wir, wie immer, den homogenen Fall $b(x) \equiv 0$.

Genauso wie bei Systemen von linearen Differentialgleichungen ist hier das Zauberwort: Der e-hoch-Ansatz

$$f(x) := e^{\lambda \cdot x}.$$

Dann ist also

$$f' = \lambda e^{\lambda x}, \; f'' = \lambda^2 e^{\lambda x}, \; ..., \; f^{(n)} = \lambda^n e^{\lambda x},$$

Und die homogene Differentialgleichung für $f = e^{\lambda x}$ lautet

$$\lambda^n e^{\lambda x} + a_{n-1} \lambda^{n-1} e^{\lambda x} + ... + a_1 \lambda e^{\lambda x} + a_0 e^{\lambda x} = 0.$$

Wenn man hier $e^{\lambda x}$ ausklammert, und anschließend durch diese Funktion dividiert, stößt man auf die Polynom-Gleichung

$$\lambda^n + a_{n-1} \lambda^{n-1} + ... + a_1 \lambda + a_0 = 0.$$

Grad n in λ. Sie hat genau die gleichen Koeffizienten, wie die ursprüngliche Differentialgleichung.

Definition 3.10 (Charakteristisches Polynom) *Das Polynom*

$$\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0$$

heißt das charakteristische Polynom *der Differentialgleichung*

$$f^{(n)} + a_{n-1}f^{(n-1)} + ... + a_1 f' + a_0 f = 0.$$

Dieses charakteristische Polynom ist im Wesentlichen das charakteristische Polynom der Koeffizientenmatrix des zugehörigen $n \times n$-Systems von linearen Differentialgleichungen erster Ordnung mit konstanten Koeffizienten. Und wenn λ eine Nullstelle dieses Polynoms ist, dann ist die Funktion $e^{\lambda x}$ eine Lösung der Differentialgleichung.

Beispiel 3.24 *Die Differentialgleichung*

$$f'' + 2f' + f = 0$$

hat das charakteristische Polynom

$$\lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$$

mit der Nullstelle $\lambda = -1$. Also ist e^{-x} eine Lösung der Differentialgleichung.

Wenn das charakterisitsche Polynom n verschiedene reelle Nullstellen $\lambda_1, ..., \lambda_n$ hat, dann hat man mit diesem billigen Ansatz die Differentialgleichung gelöst:

Satz 3.23 *Wenn das charakteristische Polynom n verschiedene reelle Nullstellen $\lambda_1, ..., \lambda_n$ besitzt, dann hat die zugehörige lineare Differentialgleichung das Lösungsfundamentalsystem*

$$e^{\lambda_1 x}, ..., e^{\lambda_n x}.$$

Beweis (Induktion nach n). Nach obigem ist klar, dass die n Funktionen $e^{\lambda_1 x}, ..., e^{\lambda_n x}$ Lösungen der Differentialgleichung sind. Wir müssen nur noch zeigen: Sie sind linear unabhängig. Dazu benützen wir den Test

$$\gamma_1 e^{\lambda_1 x} + ... + \gamma_n e^{\lambda_n x} \equiv 0 \quad \text{mit} \quad \gamma_1, ..., \gamma_n \in \mathbb{R}.$$

Wir dividieren durch $e^{\lambda_n x}$

$$\gamma_1 e^{(\lambda_1 - \lambda_n)x} + ... + \gamma_{n-1} e^{(\lambda_{n-1} - \lambda_n)x} + \gamma_n \equiv 0$$

und differenzieren

$$(\lambda_1 - \lambda_n)\gamma_1 e^{(\lambda_1 - \lambda_n)x} + ... + (\lambda_{n-1} - \lambda_n)\gamma_{n-1} e^{(\lambda_{n-1} - \lambda_n)x} \equiv 0.$$

Das ist jetzt eine lineare Relation zwischen $n-1$ Exponentialfunktionen. Auch hier sind die Koeffizienten in den Exponenten $\lambda_1 - \lambda_n, ..., \lambda_{n-1} - \lambda_n$ alle voneinander verschieden. Mit der Induktionsannahme folgt

$$(\lambda_1 - \lambda_n)\gamma_1 = ... = (\lambda_{n-1} - \lambda_n)\gamma_{n-1} = 0.$$

Und weil $\lambda_n \neq \lambda_\nu$ mit $\nu < n$ ist, folgt daraus

$$\gamma_1 = ... = \gamma_{n-1} = 0$$

und schließlich auch $\gamma_n = 0$. $\square$

Das ist ja sehr schön. Wieder einmal ist ein lineares Problem auf das Lösen einer Polynom-Gleichung zurückgeführt worden. Aber wieder einmal gibt es die üblichen Probleme damit: Die Wurzeln der Polynomgleichung brauchen nicht alle voneinander verschieden zu sein, es kann komplexe Nullstellen geben, und beides kann kombiniert auftreten.

Um diese Probleme zu entkoppeln, benutzt man vorteilhaft einen neuen Begriff.

Definition 3.11 (Differentialoperator) *Gegeben sei ein Polynom*

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0 \quad \textit{mit reellen Koeffizienten} \quad a_{n-1}, ..., a_0.$$

Der zugehörige Differentialoperator

$$\begin{aligned} P(\frac{d}{dx}) &= \left(\frac{d}{dx}\right)^n + a_{n-1}\left(\frac{d}{dx}\right)^{n-1} + ... + a_1\frac{d}{dx} + a_0 \\ &= \left(\frac{d^n}{dx^n}\right) + a_{n-1}\left(\frac{d^{n-1}}{dx^{n-1}}\right) + ... + a_1\frac{d}{dx} + a_0 \end{aligned}$$

ist die (lineare) Abbildung $C^n(\mathbb{R}) \to C^n(\mathbb{R})$ die jeder genügend oft differenzierbaren Funktion f die Funktion

$$f \mapsto P\left(\frac{d}{dx}\right)[f] := \left(\frac{d^n f}{dx^n}\right) + a_{n-1}\left(\frac{d^{n-1}f}{dx^{n-1}}\right) + ... + a_1\frac{df}{dx} + a_0 \cdot f.$$

zuordnet.

Ist P das charakteristische Polynom einer Differentialgleichung, dann ist $P(d/dx)[f]$ die linke Seite der Differentialgleichung.

Satz 3.24 *Für je zwei Polynome P und Q mit reellen Koeffizienten sind die beiden Differentialoperatoren*

$$(P \cdot Q)\left(\frac{d}{dx}\right) \qquad \textit{und} \quad P\left(\frac{d}{dx}\right) \circ Q\left(\frac{d}{dx}\right)$$

gleich. D.h., für jede genügend oft differenzierbare Funktion f gilt

$$(P \cdot Q)\left(\frac{d}{dx}\right)[f] = P\left(\frac{d}{dx}\right)\left[Q\left(\frac{d}{dx}\right)[f]\right].$$

Beweis. Sowohl das Produkt-Polynom $P(\lambda)\cdot Q(\lambda)$ als auch der Produkt-Operator $P(d/dx)\circ Q(d/dx)$ sind linear bezüglich P und Q (geradezu bilinear). Es genügt deswegen die Behauptung für Monome $P(\lambda) = \lambda^m$ und $Q(\lambda) = \lambda^n$ zu beweisen. Wir rechnen nach

$$(P \cdot Q)\left(\frac{d}{dx}\right)[f] = \left(\frac{d^{m+n}}{dx^{m+n}}\right)[f] = f^{(m+n)} = \frac{d^m}{dx^m}f^{(n)} = P\left(\frac{d}{dx}\right)Q\left(\frac{d}{dx}\right)[f].$$

$\square$

Diese Rechenregel ist sehr schön, weil sie sehr nützlich ist. Aber es ist ganz wesentlich, dass die vorkommenden Koeffizienten konstant sind. Hierzu das einfachste Gegenbeispiel.

Beispiel 3.25 (Unschärferelation) *Wir betrachten die Differentialoperatoren*

$$P = \frac{d}{dx} \quad \textit{und} \quad Q = x.$$

(Das soll heißen $Q(f) = x \cdot f$). Dann ist für $f \not\equiv 0$

$$(P \circ Q)[f] = \frac{d}{dx}(x \cdot f) = f + x \cdot f' \neq x \cdot f' = (Q \circ P)[f].$$

Wozu ist diese allgemeine Theorie gut? Nach dem Fundamentalsatz der Algebra ist das charakteristische Polynom der Differentialgleichung ein Produkt komplexer Linearfaktoren

$$P(\lambda) = (\lambda - c_1)^{r_1} \cdot \ldots \cdot (\lambda - c_k)^{r_k}.$$

Wenn zwei komplex-konjugierte Nullstellen (mit notwendig gleicher Vielfachheit!) $c = a + ib$ und $\bar{c} = a - ib$ vorkommen, fassen wir die zugehörigen Linearfaktoren zu dem reellen quadratischen Polynom

$$(\lambda - a - ib) \cdot (\lambda - a + ib) = (\lambda - a)^2 + b^2 = \lambda^2 - 2a\lambda + a^2 + b^2$$

zusammen. Dann ist P ein Produkt von Potenzen von Faktoren vom Grad 1 und Grad 2. Und jede dieser Potenzen behandeln wir separat. Ist nämlich $P = P_1 \cdot P_2$, wo P_2 so eine Potenz ist, dann ist jede Lösung der Differentialgleichung $P_2(d/df)[f] = 0$ auch eine Lösung der Differentialgleichung $P(d/dx)[f] = 0$.

Betrachten wir die beiden Fälle:

1) $P(\lambda) = (\lambda - c)^k$, *Potenz eines reellen Linearfaktors:* Die zugehörige Differentialgleichung

$$\left(\frac{d}{dx} - c\right)^k [f] = 0$$

hat sicher die Lösung $e^{c \cdot x}$. Um weitere Lösungen zu finden machen wir den Ansatz

$$f(x) = g(x) \cdot e^{c \cdot x}.$$

Was passiert? Wir berechnen

$$\begin{aligned} \left(\frac{d}{dx} - c\right)^k [g \cdot e^{cx}] &= \left(\frac{d}{dx} - c\right)^{k-1} \left[\left(\frac{d}{dx} - c\right) [g \cdot e^{cx}]\right] \\ &= \left(\frac{d}{dx} - c\right)^{k-1} [g' \cdot e^{cx} + cg \cdot e^{cx} - cg \cdot e^{cx}] \\ &= \left(\frac{d}{dx} - c\right)^{k-1} [g' \cdot e^{cx}]. \end{aligned}$$

Und so können wir weitermachen:

$$\begin{aligned} \left(\frac{d}{dx} - c\right)^k [g \cdot e^{cx}] &= \left(\frac{d}{dx} - c\right)^{k-1} g' \cdot e^{cx} \\ &= \left(\frac{d}{dx} - c\right)^{k-2} g'' \cdot e^{cx} \\ &\vdots \\ &= g^{(k)} \cdot e^{cx}. \end{aligned}$$

Für g können wir also alle Funktionen mit k-ter Ableitung

$$g^{(k)} \equiv 0$$

nehmen um Lösungen zu erhalten. Die k linear unabhängigen Polynome

$$1,\ x,\ ...,\ x^{k-1}$$

führen so auf k linear unabhängige Lösungen

$$e^{cx},\ x \cdot e^{cx},\ x^2 \cdot e^{cx},\ ...,\ x^{k-1} e^{cx}.$$

Beispiel 3.26 *Die oben betrachtete Differentialgleichung*

$$f'' + 2f' + f = 0$$

hat das charakteristische Polynom

$$P(\lambda) = (\lambda + 1)^2$$

mit der doppelten Nullstelle $\lambda = -1$. Deswegen bilden die beiden Funktionen

$$e^{-x}, \quad x \cdot e^{-x}$$

ein Lösungsfundamentalsystem dieser Diferentialgleichung.

2) $P(\lambda) = (\lambda^2 - 2a\lambda + a^2 + b^2)^k$, *Potenz eines quadratischen Polynoms ohne reelle Nullstelle*: Jetzt bringt die Verwendung der komplexen Exponentialfunktion eine solche Vereinfachung, dass ich nicht mehr darauf verzichten möchte. Wir verwenden also für $c = a + bi \in \mathbb{C}$ die komplex-wertige Funktion

$$e^{c \cdot x} = e^{(a+ib) \cdot x}.$$

Zunächst gilt die Funktionalgleichung der Exponentialfunktion

$$e^{(a+ib) \cdot x} = e^{a \cdot x} \cdot e^{ib \cdot x}$$

genau wie im Reellen. Der Beweis ist identisch, nur lässt man in der Potenzreihe auch komplexe Argumente zu. Damit, und mit der Eulerschen Formel haben wir also

$$e^{(a+ib) \cdot x} = e^{a \cdot x} \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x)).$$

Diese komplex-wertige Funktion differenzieren wir nach x, indem wir Real- und Imaginär-Teil differenzieren:

$$\begin{aligned}
\frac{d}{dx} e^{c \cdot x} &= \left(\frac{d}{dx} e^{a \cdot x} \right) \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x) + e^{a \cdot x} \cdot \frac{d}{dx} (cos(b \cdot x) + i \cdot sin(b \cdot x)) \\
&= a \cdot e^{a \cdot x} \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x) + e^{a \cdot x} \cdot b \cdot (-sin(b \cdot x) + i \cdot cos(b \cdot x)) \\
&= (a + ib) \cdot e^{a \cdot x} \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x)) \\
&= c \cdot e^{c \cdot x}
\end{aligned}$$

Für den komplexen Differentialoperator

$$P\left(\frac{d}{dx}\right) = \frac{d^2}{dx^2} - 2a\frac{d}{dx} + a^2 + b^2 = \left(\frac{d}{dx} - c\right) \cdot \left(\frac{d}{dx} - \bar{c}\right)$$

gilt dann also

$$P\left(\frac{d}{dx}\right) e^{c \cdot x} = P\left(\frac{d}{dx}\right) e^{\bar{c} \cdot x} \equiv 0.$$

Die beiden komplexen Funktionen

$$e^{c \cdot x} = e^{a \cdot x} \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x)),$$

$$e^{\bar{c} \cdot x} = e^{a \cdot x} \cdot (cos(b \cdot x) + i \cdot sin(b \cdot x))$$

sind damit Lösungen. Wir rekombinieren sie zu den beiden linear unabhängigen reellen Lösungen

$$f_1(x) := Re\,(e^{c \cdot x}) = e^{a \cdot x} \cdot cos(b \cdot x), \quad f_2(x) := Im\,(e^{c \cdot x}) = e^{a \cdot x} \cdot sin(b \cdot x).$$

Und weiter geht es genau wie in Fall 1): Mit f_1 und f_2 sind auch die Funktionen

$$x \cdot f_1(x), ..., x^{k-1} \cdot f_1(x), x \cdot f_2(x), ..., x^{k-1} \cdot f_2(x)$$

Lösungen der Differentialgleichung. Damit haben wir den folgenden Satz schon zum Teil bewiesen.

Satz 3.25 *Das reelle Polynom $P(\lambda)$ vom Grad n zerfalle in Linearfaktoren*

$$P(\lambda) = (\lambda - c_1)^r_1 \cdot ... \cdot (\lambda - c_k)^{r_k} \cdot ((\lambda - (a_1 + ib_1)(\lambda - (a_1 - ib_1))^{s_1} \cdot ... \cdot ((\lambda - (a_l + ib_l)(\lambda - (a_l - ib_l))^{s_l},$$

$$r_1 + ... + r_k + 2(s_1 + ... + s_l) = n.$$

Dann bilden die folgenden n Funktionen

$$e^{c_1 \cdot x}, ..., x^{r_1 - 1} \cdot e^{c_1 \cdot x},$$

$$\vdots$$

$$e^{c_k \cdot x}, ..., x^{r_k - 1} \cdot e^{c_k \cdot x},$$

$$e^{a_1 \cdot x} \cdot cos(b_1 \cdot x), ..., x^{s_1 - 1} \cdot e^{a_1 \cdot x} \cdot cos(b_1 \cdot x),\ e^{a_1 \cdot x} \cdot sin(b_1 \cdot x), ..., x^{s_1 - 1} \cdot e^{a_1 \cdot x} \cdot sin(b_1 \cdot x),$$

$$\vdots$$

$$e^{a_l \cdot x} \cdot cos(b_l \cdot x), ..., x^{s_l - 1} \cdot e^{a_l \cdot x} \cdot cos(b_l \cdot x),\ e^{a_l \cdot x} \cdot sin(b_l \cdot x), ..., x^{s_l - 1} \cdot e^{a_l \cdot x} \cdot sin(b_l \cdot x)$$

ein Lösungsfundamentalsystem der Differentialgleichung

$$P\left(\frac{d}{dx}\right)[f] = 0.$$

Beweis. Dass die angegebenen Funktionen Lösungen sind, das wissen wir schon. Es bleibt zu zeigen, dass sie linear unabhängig sind. Wir müssen also eine beliebige Linearkombination all dieser Funktionen untersuchen, welche die Nullfunktion ergibt, und zeigen, alle Koeffizienten in dieser Linearkombination sind $= 0$. In der Linearkombination kommen jede Exponentialfunktion und Winkelfunktion, bzw. jedes Produkt von Exponentialfunktion und Winkelfunktion mit Potenzen von x vor, die wir zu einem Polynom zusammenfassen können. Dann sieht die Linearkombination so aus:

$$p_1(x)e^{c_1\cdot x} + ... + p_k(x)\cdot e^{c_k\cdot x} +$$

$$+u_1(x)\cdot e^{a_1\cdot x}\cdot cos(b_1\cdot x)+v_1(x)\cdot e^{a_1\cdot x}\cdot sin(b_1\cdot x)+...+u_l(x)\cdot e^{a_l\cdot x}\cdot cos(b_l\cdot x)+v_l(x)\cdot e^{a_l\cdot x}\cdot sin(b_l\cdot x)\equiv 0.$$

Jetzt ist es wieder vorteilhaft von den Winkelfunktionen zu komplexen Exponentialfunktionen überzugehen vermöge

$$\begin{aligned} u(x)\cdot e^{a\cdot x}\cdot cos(b\cdot x)+v(x)\cdot e^{a\cdot x}\cdot sin(b\cdot x) &= u(x)\cdot\frac{1}{2}\cdot(e^{c\cdot x}+e^{\bar{c}\cdot x})v(x)\cdot\frac{1}{2i}\cdot(e^{c\cdot x}-e^{\bar{c}\cdot x}) \\ &= \frac{1}{2}(u(x)+\frac{1}{i}v(x))\cdot e^{c\cdot x}+\frac{1}{2}(u(x)-\frac{1}{i}v(x))\cdot e^{\bar{c}\cdot x} \end{aligned}$$

Dann bleibt folgende Aussage zu zeigen:

Satz 3.26 (Lemma) *Es seien $c_1, ..., c_k$ paarweise verschiedene (reelle oder komplexe) Zahlen und $p_1, ..., p_k$ Polynome (mit reellen oder komplexen Koeffizienten). Ist die Funktion*

$$p_1(x)\cdot e^{c_1\cdot x} + ... + p_k(x)\cdot e^{c_k\cdot x} \equiv 0$$

identisch die Nullfunktion, dann sind alle Polynome $p_1, ..., p_k$ das Nullpolynom.

Beweis (Induktion nach k). Die Polynome $p_1, ..., p_k$ mögen den Grad $r_1, ..., r_k$ haben. Wir wenden auf die obige Gleichung den Differentialoperator

$$P = \left(\frac{d}{dx} - c_k\right)^{r_k+1}$$

an. Dann verschwindet der letzte Summand spurlos aus der Gleichung. Für jedes $l < k$ ist aber

$$\left(\frac{d}{dx} - c_k\right) p_l(x)e^{c_l\cdot x} = (p_l'(x) + (c_l - c_k)\cdot p_l(x))\cdot e^{c_l\cdot x}.$$

Wegen $c_l \neq c_k$ hat hier das Polynom $p_l' + (c_l - c_k)p_l$ den gleichen Grad r_l wie das Polynom p_l. Durch Iteration finden wir

$$P(p_l(x)\cdot e^{c_l\cdot x}) = q_l(x)\cdot e^{c_l\cdot x}$$

mit einem Polynom q_l vom gleichen Grad r_l wie p_l. Auf die Gleichung

$$q_1(x)\cdot e^{c_1\cdot x} + ... + q_{k-1}(x)\cdot e^{c_{k-1}}\cdot x \equiv 0$$

wenden wir die Induktionsannahme an und finden, dass jedes Polynom $q_1, ..., q_{k-1}$ das Nullpolynom ist. Dann müssen auch die Polynome $p_1 = ... = p_{k-1} = 0$ das Nullpolynom gewesen sein. Unsere Ausgangsgleichung schnurrt zusammen auf

$$p_k(x)\cdot e^{c_k\cdot x} \equiv 0,$$

woraus $p_k(x) \equiv 0$ folgt. $\square$

Schließlich müssen wir uns noch um die inhomogenen Gleichungen

$$f^{(n)} + a_{n-1}f^{(n-1)} + ... + a_1 f' + a_0 f = b(x)$$

kümmern. Im Prinzip gibt es die Möglichkeit, sie in ein System linearer Differentialgleichgungen erster Ordnung umzuschreiben und, wenn man ein Lösungsfundamentalsystem kennt, mit der Variation der Konstanten eine spezielle Lösung zu ermitteln. Es gibt aber direktere Methoden zur Ermittlung einer speziellen Lösung, wenn die rechte Seite $b(x)$ eine besonders einfache Form hat. Und zwar, wenn $b(x)$ eine Linearkombination von folgenden Funktionen ist:

- Exponentialfunktionen e^{cx},
- Winkelfunktionen $sin(bx)$ oder $cos(bx)$,
- Potenzen x^k multipliziert mit einer Exponential- oder Winkelfunktion.

Am einfachsten funktioniert es mit Exponentialfunktionen

$$b(x) = e^{cx}.$$

Wendet man nämlich den Differentialoperator $P(d/dx)$ auf die Funktion e^{cx} an, so kommt $P(c)e^{kx}$ heraus: Wegen

$$\left(\frac{d}{dx}\right)^k [e^{cx}] = c^k \cdot e^{cx}$$

ist

$$\begin{aligned} P\left(\frac{d}{dx}\right)[e^{cx}] &= \left(\frac{d}{dx}\right)^n [e^{cx}] + a_{n-1}\left(\frac{d}{dx}\right)^{n-1} [e^{cx}] + ... + a_1 \frac{d}{dx} e^{cx} + a_0 \cdot e^{cx} \\ &= P(c) \cdot e^{cx}. \end{aligned}$$

Und somit ist

$$f_0 := \frac{1}{P(c)} e^{cx}$$

eine spezielle Lösung der Gleichung

$$P\left(\frac{d}{dx}\right)[f] = e^{cx}.$$

Natürlich muss hier $P(c) \neq 0$ sein. Aber wenn $P(c) = 0$ doch der Fall sein sollte, hilft ein Ansatz

$$f_1(x) = x^k \cdot e^{cx},$$

wo k die Vielfachheit der Nullstelle c von P ist:

$$\begin{aligned} P\left(\frac{d}{dx}\right) &= P_1\left(\frac{d}{dx}\right)\left(\frac{d}{dx} - c\right)^k [x^k e^{cx}] \\ &= P_1\left(\frac{d}{dx}\right)\left[(x^k)^{(k)} e^{cx}\right] \\ &= k! \cdot P_1(c) e^{cx} \end{aligned}$$

und man erhält eine spezielle Lösung

$$\frac{1}{k! \cdot P_1(c)} \cdot e^{cx}.$$

Beispiel 3.27

$$f'' + 2f' + f = e^{2x}.$$

Eine spezielle Lösung ist

$$\frac{1}{P(2)}e^{2x} = \frac{1}{(2+1)^2} \cdot e^{2x} = \frac{1}{9}e^{2x}.$$

Wenn Winkelfunktionen auf der rechten Seite stehen geht es ganz ähnlich: Man wendet $P(d/dx)$ auf diese Winkelfunktionen an, und sucht eine geeignete Linearkombination aus. Die allgemeinen Formeln sind unschön hinzuschreiben, und unmöglich zu merken. Ich möchte das Prinzip nur an einem Beispiel illustrieren:

Beispiel 3.28 *Gesucht sei eine spezielle Lösung der Gleichung*

$$f'' + 2f' + f = sin(2x).$$

Es ist

$$\begin{aligned}
P\left(\frac{d}{dx}\right)[sin(2x)] &= -4sin(2x) + 4cos(2x) + sin(2x) \\
&= -3sin(2x) + 4cos(2x) \\
P\left(\frac{d}{dx}\right)[cos(2x)] &= -4cos(2x) - 4sin(2x) + cos(2x) \\
&= -4sin(2x) - 3cos(2x), \\
P\left(\frac{d}{dx}\right)[3sin(2x) + 4cos(2x)] &= -25sin(2x),
\end{aligned}$$

und eine spezielle Lösung ist

$$f_0(x) = -\frac{1}{25}(3sin(2x) + 4cos(2x)).$$

Auch wenn die rechte Seite $b(x)$ ein Polynom ist, funktioniert das Verfahren ganz ähnlich.

Schließlich kann man spezielle Lösungen auch linear kombinieren, wenn Linearkombinationen auf der rechten Seite stehen. So hat z.B. die Gleichung

$$f'' + 2f + f = e^{2x} + sin(2x)$$

die spezielle Lösung

$$f_0(x) = \frac{1}{9}e^{2x} - \frac{1}{25}(3sin(2x) + 4cos(2x).$$

Als krönenden Abschluss dieses Paragraphen, ja dieses Kapitels, ja sogar dieses Semesters möchte ich jetzt die aus der Schule wohlbekannte Differentialgleichung des harmonischen Oszillators diskutieren. Nach unserer Bekanntschaft mit der ganzen Theorie der Differentialgleichungen haben wir jetzt einen höheren Standpunkt, von dem aus wir die bekannten Eigenschaften dieser Differentialgleichung und ihrer Lösungen betrachten wollen.

Die Differentialgleichung des ungedämpften harmonischen Oszillators ist

$$y'' + \omega_0^2 y = 0.$$

Die wird in der Oberstufe im Physik-Unterricht diskutiert, weil der harmonische Oszillator so unheimlich wichtig für die Anwendungen ist. Der einfachste harmonische Oszillator ist das Fadenpendel. Aber Vorsicht! Die Gleichung des Fadenpendels ist in Wirklichkeit nicht

$$y'' = -k \cdot y$$

sondern

$$y'' = -k \cdot sin(y).$$

Das ist gar keine lineare, sondern eine nichtlineare Differentialgleichung zweiter Ordnung. Na, sei's drum, sagt man in der Schule, und in der Experimentalphysik, für kleine y ist $sin(y) \simeq y$, und auf diese kleinen y kommt es uns an. Na ja, selbst wenn das so ist, hat das Fadenpendel, spätestens seit Erfindung der Uhrfeder vor fünfhundert Jahren, oder so, (in Nürnberg?), stark an Bedeutung für technische Anwendungen eingebüßt.

Na gut, aber da war noch der elektrische Schwingkreis, ohne den kein Radio und kein Fernseher auskommt. Aber auch der ist ein harmonischer Oszillator nur in einem kleinen Bereich. Wenn wir an ihn ein paar Tausend Volt anlegen, schmort er zusammen, und schwingt gar nicht mehr. Irgendwie erinnert das ans Fadenpendel mit seinen moderaten Schwingungen.

In Wirklichkeit ist es eben so, dass die Gleichung des harmonischen Oszillators deswegen wichtig ist, weil man sie so schön mathematisch analysieren kann. Und wenn sie dann noch ein paar physikalische Systeme näherungsweise beschreibt, dann ist das eben noch schöner. Wirkliche, in der Technik vorkommende Schwingungen sind, wie vieles im Leben, meist nicht harmonisch. Ihre Differentialgleichungen sind wesentlich komplizierter.

So, nach diesen philosophischen, aber trotzdem wichtigen Bemerkungen zurück zur Gleichung

$$y'' + \omega_0^2 y = 0.$$

Ihr charakteristisches Polynom ist

$$P(\lambda) = \lambda^2 + \omega_0^2$$

mit den beiden komplexen Nullstellen $\pm i\omega_0$. Ein Lösungsfundamentalsystem enthält deswegen auch keine Exponentialfunktionen, sondern wird z.B. von den beiden Winkelfunktionen

$$sin(\omega_0 t),\ cos(\omega_0 t)$$

gebildet.

Ein realistischerer harmonischer Oszillator ist der *gedämpfte harmonische Oszillator*

$$y'' + 2\mu y' + \omega_0^2 y = 0, \quad 0 < \mu \in \mathbb{R}.$$

Sein charakteristisches Polynom ist

$$P(\lambda) = \lambda^2 + 2\mu \cdot \lambda + \omega_0^2.$$

Seine beiden Lösungen

sind		falls
komplex-konjugiert	$-\mu \pm i\sqrt{\omega_0^2 - \mu^2}$	$\mu < \omega_0$
zusammenfallend	$-\mu$	$\mu = \omega_0$
reell und < 0	$-\mu \pm \sqrt{\mu^2 - \omega_0^2}$	$\mu > \omega_0$

Der erste Fall heißt der Fall *schwacher Dämpfung*. Ein Lösungsfundamentalsystem sind die beiden Funktionen

$$y_1(t) = e^{-\mu t} sin(\omega t),\ y_2(t) = e^{-\mu t} cos(\omega t), \qquad \omega := \sqrt{\omega_0^2 - \mu^2}.$$

Es sind Schwingungen mit Frequenz ω, deren Amplitude wegen der Dämpfung exponentiell kleiner wird.

Der zweite Fall ist der *Grenzfall* mit dem Lösungsfundamentalsystem

$$y_1(t) = e^{-\mu t},\ y_2(t) = t \cdot e^{-\mu t}.$$

Hier schwingt nichts mehr, die Lösung fällt unaufhaltsam gegen 0. Aber wenn bei einer Lösung

$$y = c_1 y_1 + c_2 y_2 = c_1 e^{-\mu t} + c_2 t \cdot e^{-\mu t}$$

die Koeffizienten verschiedenes Vorzeichen haben, geht die Lösung bei $t = -c_1/c_2$ durch die Null-Lage, ein einziges mal, bevor sie vorübergehend nochmal anwächst, und sich dann unwiderruflich monoton fallend zur Null-Lage hinbewegt.

Der dritte Fall heißt der Fall *starker Dämpfung*. Beide Wurzeln

$$\mu_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}$$

sind negativ. Beide Lösungen

$$y_1(t) = e^{-\mu_1 t},\ y_2(t) = e^{-\mu_2 t}$$

beschreiben eine exponentielle Abnahme der Auslenkung. Auch hier kann die Lösung, für Koeffizienten c_1 und c_2 mit verschiedenem Vorzeichen, eine Nullstelle haben, aber nur eine einzige.

Damit ein harmonischer Oszillator nicht durch die in der Realität unausweichliche Dämpfung abgebremst wird, stößt man ihn manchmal periodisch an. Eine Differentialgleichung dafür ist

$$y'' + 2\mu y' + \omega_0^2 y = acos(t).$$

Es kommt darauf an, eine spezielle Lösung $y_0(t)$ zu finden. Denn weil die Lösungen der homogenen Gleichung auf 0 zu fallen, wird sich jede Lösung langfristig in die Nähe dieser speziellen Lösung hin einpendeln. Wir wenden mal

$$P\left(\frac{d}{dt}\right) = \frac{d^2}{dt^2} + 2\mu\frac{d}{dt} + \omega_0^2$$

auf die beiden Winkelfunktionen $sin(t)$ und $cos(t)$ an:

$$\begin{aligned}
P\left(\frac{d}{dx}\right)[sin(t)] &= -sin(t) + 2\mu cos(t) + \omega_0^2 sin(t)\\
&= (\omega_0^2 - 1)sin(t) + 2\mu cos(t),\\
P\left(\frac{d}{dx}\right)[cos(t)] &= -cos(t) - 2\mu sin(t) + \omega_0^2 cos(t)\\
&= (\omega_0^2 - 1)cos(t) - 2\mu sin(t).
\end{aligned}$$

Setzen wir erst einmal

$$z(t) := 2\mu sin(t) + (\omega_0^2 - 1)cos(t),$$

so wird also

$$P\left(\frac{d}{dx}\right)[z] = (4\mu^2 + (\omega_0^2 - 1)^2)cos(t).$$

Deswegen ist

$$y_0(t) := \frac{a}{4\mu^2 + (\omega_0^2 - 1)^2}(2\mu sin(t) + (\omega_0^2 - 1)cos(t))$$

eine spezielle Lösung.

Aufgabe 3.30 *Gegeben sei die Differentialgleichung*

$$x^2 y''(x) + xy'(x) + (x^2 - \frac{1}{4})y(x) = 0 \quad (x > 0).$$

Durch eine Transformation $y(x) =: x^\rho z(x)$ mit einem geeignet zu wählenden $\rho \in \mathbb{R}$ bestimme man für $z(x)$ eine Differentialgleichung mit konstanten Koeffizienten.

Aufgabe 3.31 *Gegeben seien die Differentialgleichungen*

$$\begin{array}{ll} (1) & x^4 \dfrac{d^2y}{dx^2} + 2(x^3 + x^2)\dfrac{dy}{dx} + y = e^{\frac{1}{x}},\ x > 0 \\ (2) & \dfrac{d^2z}{dt^2} - 2\dfrac{dz}{dt} + z = e^t,\ t > 0. \end{array}$$

Man zeige, dass sich die Differentialgleichung (1) durch Substitution auf die Differentialgleichung (2) zurückführen lässt.

Aufgabe 3.32 *Gegeben sei die Differentialgleichung*

$$x^3 y''' - 6xy' + 12y = 0 \quad (x > 0).$$

a) Bestimmen sie alle reellen Zahlen $r \in \mathbb{R}$ derart, dass die Funktion $x \mapsto x^r$ für $x > 0$ eine Lösung ist. Bestimmen Sie ein Fundamentalsystem für die Differentialgleichung.
b) Bestimmen Sie alle Lösungen der Differentialgleichung mit $y(1) = 1$ und $y'(1) = 0 = y''(1)$.

Aufgabe 3.33 *Gegeben sei die Differentialgleichung*

$$y'' + p(x)y' + q(x)y = f(x), \quad x \in \mathbb{R}.$$

wobei $p, q, f : \mathbb{R} \to \mathbb{R}$ reelle Funktionen seien und $f(x_0) \neq 0$ für ein $x_0 \in \mathbb{R}$ gelte. Es seien $y_1, y_2 : \mathbb{R} \to \mathbb{R}$ Lösungen dieser Gleichung. Man bestimme alle Zahlen $c_1, c_2 \in \mathbb{R}$, für welche $c_1 y_1 + c_2 y_2$ ebenfalls Lösung der Differentialgleichung ist.

Aufgabe 3.34 *Die Exponentialfunktion* e^x *und ihre* k*-te Partialsumme*

$$e_k(x) = \sum_{\nu=0}^{k} \frac{x^\nu}{\nu!}$$

mit $k \in \mathbb{N} \setminus (1)$ *seien Lösungen einer Differentialgleichung 2. Ordnung der Form:*

$$xy''(x) + p(x)y'(x) + q(x)y(x) = 0.$$

Man bestimme die Koeffizientenfunktionen $p, r : \mathbb{R} \to \mathbb{R}$*. Ist auch der* k*-te Reihenrest* $r_k(x) := e^x - e_k(x)$ *eine Lösung dieser Differentialgleichung?*

Aufgabe 3.35 *Die Funktionen*

$$\mathbb{R} \to \mathbb{R},\ x \mapsto 3e^x + e^{x^2},\ \mathbb{R} \to \mathbb{R},\ x \mapsto 7e^x + e^{x^2} \quad \textit{und}\ \mathbb{R} \to \mathbb{R},\ x \mapsto 5e^x + e^{-x^2} + e^{x^2}$$

seien Lösungen einer inhomogenen linearen Differentialgleichung 2. Ordnung mit stetigen Koeffizienten. Bestimmen Sie alle Lösungen $\varphi : \mathbb{R} \to \mathbb{R}$ *mit* $\varphi(0) = 1$ *und* $\varphi'(0) = 2$.

Aufgabe 3.36 *Zeigen Sie, dass jedes lineare homogene* 2×2*-System*

$$\begin{aligned} y_1' &= a_{1,1}(x)y_1 + a_{1,2}(x)y_2, \\ y_2' &= a_{2,1}(x)y_1 + a_{2,2}(x)y_2, \end{aligned}$$

äquivalent zu einer Differentialgleichung zweiter Ordnung ist.

Aufgabe 3.37 *Bestimmen Sie jene Lösung der Differentialgleichung*

$$y''(x) + 5y'(x) = x + e^{-5x},$$

die der Bedingung $y(0) = y'(0) = 0$ *genügt.*

Aufgabe 3.38 *Bestimmen Sie für die Differentialgeichung*

$$y''' + y'' + 8y' - 10y = 18e^{-x}$$

1. die allgemeine reelle Lösung.
2. diejenige reelle Lösung, die die Bedingungen

$$y(0) = 0; \quad y\left(\frac{\pi}{6}\right) = 0; \quad \lim_{x\to\infty} y(x) = 0$$

erfüllt.

Aufgabe 3.39 *Im Intervall* $]0,\infty[$ *betrachte man die linearen Differentialgleichungen*

$(*)\qquad y'' - 6y' + 9y = e^x$ *und*

$(**)\qquad x^2y'' - 5xy' + 9y = x.$

a) *Bestimmen Sie alle Lösungen der Differentialgleichung* $(*)$.

b) *Bestimmen Sie alle Lösungen der Differentialgleichung* $(**)$.

Hinweis: Zeigen Sie, dass für jede Lösung $\varphi:]0,\infty[\to \mathbb{R}$ *der Differentialgleichung (**) die Funktion* $\psi:]0,\infty[\to \mathbb{R}$ *mit* $\psi(x) := \varphi(e^x)$ *für* $x \in]0,\infty[$ *eine Lösung der Differentialgleichung (*) ist.*

Aufgabe 3.40 *Bestimmen Sie alle Lösungen der Differentialgleichung*

a) $y^{(4)} + 2y^{(3)} - 8y^{(2)} = exp(x)$,

b) $y'' + 4y' - 2y = x$,

c) $y'' + y' - 2y = x^2$ *mit der Anfangsbedingung* $y(0) = 0$,

d) $y''' + y'' + y' + y = 2e^{-x}$,

e) $y'' + y = e^x$ *mit der Anfangsbedingung* $y(0) = y(\frac{\pi}{2}) = 0$,

f) $y'' + 16y = -sin(4t)$.

Aufgabe 3.41 *Für welche positiven reellen Zahlen c besitzt die Differentialgleichung*

$$y'' + c^2y = 0$$

eine Lösung $\varphi: \mathbb{R} \to \mathbb{R}$ *mit* $\varphi \neq 0,\ \varphi(0) = 0$ *und* $\varphi'(1) = 0$*?*

Aufgabe 3.42 *Bestimmen Sie alle* $\alpha \in \mathbb{R}$*, für die die Differentialgleichung*

$$y'' + 2y' + \alpha y = 0$$

eine von 0 *verschiedene Lösung mit*

$$y(0) = y(2\pi) = 0$$

hat.

Aufgabe 3.43 *Man bestimme alle homogenen linearen Differentialgleichungen (D) 2. Ordnung mit konstanten Koeffizienten, welche die folgenden beiden Eigenschaften haben:*

a) Das charakterisitische Polynom von (D) besitzt eine Doppelnullstelle.

b) (D) besitzt eine Lösung $\varphi: \mathbb{R} \to \mathbb{R}$ *mit*

$$\varphi(0) = 1,\ \varphi'(0) = 1 \text{ und } \varphi(1) = 0.$$

Index