2 Funktionen mehrerer Variabler

Im Folgenden betrachten wir Funktionen in mehreren Variablen. Im Hinblick auf den Bedarf in einigen Anwendungen lassen wir hier auch komplexwertige Funktionen zu. Wir betrachteten also Funktionen auf Teilmengen $D \subset \mathbb{K}^n$ mit Bildbereich in $\mathbb{K}$ oder vektor- und matrixwertige Funktionen mit Bildbereich in $\mathbb{K}^n$ bzw. in $\mathbb{K}^{n\times n}$. Dabei interessieren uns zunächst grundlegende Eigenschaften wie Stetigkeit, gleichmäßige Stetigkeit und Liptschitz-Stetigkeit, welche analog zum eindimensionalen Fall definiert sind. Viele dieser Begriffe lassen sich auch wieder ganz allgemein für Funktionen auf Teilmengen von normierten oder metrischen Räumen definieren. Auf diese formale Verallgemeinerung wird aber hier bewusst zugunsten der Anschaulichkeit verzichtet.

2.1 Stetigkeit

Wir betrachten Funktionen $f : D \subset \mathbb{K}^n \to \mathbb{K}$ mit nichtleerem Definitionsbereich $D \subset \mathbb{K}^n$ und Bildbereich $B_f \subset \mathbb{K}$. Für Teilmengen $M \subset D$ und $N \subset f(D)$ sind das „Bild“ von M bzw. das „Urbild“ von N definiert durch

$$f(M) := \{y \in \mathbb{K} : \exists x \in M, y = f(x)\}, \qquad f^{-1}(N) := \{x \in D : \exists y \in N, f(x) = y\}.$$

In diesem Sinne ist dann $B_f = f(D)$ und $D = f^{-1}(B_f)$. Die Notation $f^{-1}(\cdot)$ ist mengentheoretisch zu verstehen und darf nicht mit der für bijektive Abbildungen definierten punktweisen Umkehrabbildung $f^{-1} : B_f \to D$ verwechselt werden. Wegen der Äquivalenz aller Normen auf $\mathbb{K}^n$ sind alle im Folgenden abgeleiteten Aussagen unabhängig von der gewählten Norm. Diese wird daher einheitlich mit $\|\cdot\|$ bezeichnet und bedeutet in der Regel die euklidische Norm.

Definition 2.1 (Stetigkeit): *Eine Funktion $f : D \to \mathbb{K}$ heißt „stetig“ in einem Punkt $a \in D$, wenn für jede Folge $(x^{(k)})_{k\in\mathbb{N}}$ in D gilt:*

$$x^{(k)} \to a \ \ (k \to \infty) \quad \Rightarrow \quad f(x^{(k)}) \to f(a) \ \ (k \to \infty).$$

Sie heißt „stetig in D“, wenn sie in jedem Punkt in D stetig ist.

Die Definitionsbereiche stetiger Funktionen brauchen nicht *offen* zu sein. Für eine stetige Funktion $f : D \to \mathbb{K}$ ist offenbar auch jede Restriktion $f|_M : M \to \mathbb{K}$ auf eine Teilmenge $M \subset D$ stetig. Ferner sind mit f auch der Realteil $\mathrm{Re}\, f$ der Imaginärteil $\mathrm{Im}\, f$ sowie der Absolutbetrag $|f|$ stetig.

Lemma 2.1: *Eine Funktion $f : D \subset \mathbb{K}^n \to \mathbb{K}$ ist genau dann in einem Punkt $a \in D$ stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für $x \in D$ gilt:*

$$\|x - a\| < \delta \quad \Rightarrow \quad |f(x) - f(a)| < \varepsilon.$$

Beweis: Die Argumentation verläuft analog wie im eindimensionalen Fall. Q.E.D.

Lemma 2.2: *Für zwei stetige Funktionen $f, g : D \to \mathbb{K}$ sind auch die Summe $f + g$, das Produkt fg sowie im Falle $g(x) \neq 0,\ x \in D$, auch der Quotient f/g stetig.*

Beweis: Die Argumentation verläuft analog wie im eindimensionalen Fall. Q.E.D.

Beispiel 2.1: Wir geben ein paar Beispiele elementarer stetiger Funktionen:

i) Die Koordinatenfunktionen $f(x) := x_k,\ k = 1, \dots, n$, sind wegen $|x_k - a_k| \le \|x - a\|_2$ stetig auf ganz $\mathbb{K}^n$.

ii) Jede Norm $\|\cdot\|$ auf $\mathbb{K}^n$ definiert wegen $\big|\, \|x\| - \|a\| \,\big| \le \|x - a\|$ eine stetige Funktion.

iii) Ein „Monom" auf dem $\mathbb{K}^n$ vom Grad r ist eine Funktion der Gestalt

$$m(x) := x_1^{r_1} \dots x_n^{r_n},$$

mit $r := r_1 + \dots + r_n,\ r_i \in \mathbb{N}_0$. Eine „Polynomfunktion" (oder kurz „Polynom") ist eine Linearkombination von Monomen vom Grad $\le r$:

$$p(x) := \sum_{\substack{k_1, \dots, k_n \in \mathbb{N}_0 \\ k_1 + \dots + k_n = 0}}^{r} a_{k_1 \dots k_n} x_1^{k_1} \dots x_n^{k_n}$$

mit Koeffizienten $a_{k_1 \dots k_n} \in \mathbb{K}$. Das Polynom heißt „vom Grad r", wenn mindestens einer der Koeffizienten $a_{k_1 \dots k_n} \neq 0$ für $k_1 + \dots + k_n = r$. Ein Beispiel eines Polynoms zweiten Grades auf dem $\mathbb{K}^2$ ist

$$P(x) = \sum_{\substack{k_1, k_2 \in \mathbb{N}_0 \\ k_1 + k_2 = 0}}^{2} a_{k_1 k_2} x_1^{k_1} x_2^{k_2} = a_{00} + a_{10} x_1 + a_{01} x_2 + a_{20} x_1^2 + a_{11} x_1 x_2 + a_{02} x_2^2.$$

Nach dem eben Bewiesenen sind Polynome stetige Funktionen auf ganz $\mathbb{K}^n$.

iv) Sind $f : D \subset \mathbb{K}^n \to \mathbb{K}$ und $g : f(D) \subset \mathbb{K} \to \mathbb{K}$ stetige Funktionen, so ist auch die zusammengesetzte Funktion $g \circ f : D \to \mathbb{K}$ stetig. Z. B. ist die Funktion $f(x) := \sqrt{\|x\|}$ auf ganz $\mathbb{K}^n$ stetig.

v) Die quadratische Funktion auf dem $\mathbb{R}^2$

$$q(x_1, x_2) := a x_1^2 + 2b x_1 x_2 + c x_2^2 + 2d x_1 + 2e x_2 + f$$

hat als Nullstellenmenge $\{(x_1, x_2) \in \mathbb{R}^2 |\, q(x_1, x_2) = 0\}$ gerade die Kegelschnitte:

$$\begin{vmatrix} a & b & d \\ b & c & e \\ d & e & f \end{vmatrix} \neq 0 \ \text{ regulärer Kegelschnitt,} \qquad \begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 \begin{cases} > 0 & \text{Ellipse} \\ = 0 & \text{Parabel} \\ < 0 & \text{Hyperbel} \end{cases}$$

Wir beweisen im Folgenden zunächst einige grundlegende Eigenschaften stetiger Funktionen auf kompakten Mengen, welche bereits vom eindimensionalen Fall her bekannt sind.

Satz 2.1 (Beschränktheit): *Eine stetige Funktion* $f : D \subset \mathbb{K}^n \to \mathbb{K}$ *ist auf jeder kompakten Menge* $K \subset D$ *beschränkt, d. h.: Es gibt eine Konstante* M_K *mit*

$$|f(x)| \le M_K, \quad x \in K. \tag{2.1.1}$$

Beweis: Die Argumentation ist analog zum eindimensionalen Fall. Angenommen, die stetige Funktion $f(x)$ ist nicht beschränkt auf K. Dann gibt es zu jedem $k \in \mathbb{N}$ ein $x^{(k)} \in K$ mit $|f(x^{(k)})| > k$. Die Folge $(x^{(k)})_{k\in\mathbb{N}}$ aus der kompakten Menge K besitzt eine konvergente Teilfolge $(x^{(k_j)})_{j\in\mathbb{N}}$ mit Limes $x \in K$. Da f stetig ist, folgt

$$|f(x^{(k_j)})| \to |f(x)| < \infty \quad (j \to \infty),$$

im Widerspruch zur Annahme $|f(x^{(k)})| \to \infty \ (k \to \infty)$. Q.E.D.

Satz 2.2 (Extremum): *Eine stetige Funktion* $f : D \subset \mathbb{K}^n \to \mathbb{R}$ *nimmt auf jeder (nichtleeren) kompakten Menge* $K \subset D$ *ihr Maximum und Minimum an, d. h.: Es gibt Punkte* $x^{\max}$ *und* $x^{\min}$, *so dass*

$$f(x^{\max}) = \sup_{x\in K} f(x), \qquad f(x^{\min}) = \inf_{x\in K} f(x). \tag{2.1.2}$$

Beweis: Die Argumentation ist analog zum eindimensionalen Fall. Die auf K stetige Funktion f ist nach Satz 2.1 beschränkt, d. h.: Sie besitzt eine obere Grenze

$$\kappa := \sup_{x\in K} f(x) < \infty.$$

Dazu gibt es eine Folge $(x^{(k)})_{k\in\mathbb{N}}$ von Punkten aus K mit $f(x^{(k)}) \to \kappa \ (k \to \infty)$. Diese Folge besitzt wieder eine konvergente Teilfolge $(x^{(k_j)})_{j\in\mathbb{N}}$ mit Limes $x^{\max} \in K$. Da f stetig ist, folgt $f(x^{(k_j)}) \to f(x^{\max}) \ (j \to \infty)$, d. h.: Es gilt $f(x^{\max}) = \kappa$. Das Argument für die untere Grenze ist analog. Q.E.D.

Anwendung 2.1.1: Seien $K_1, K_2 \subset \mathbb{K}^n$ (nichtleere) kompakte Mengen. Dann ist die Menge $K_1 \times K_2$ kompakt im Produktraum $\mathbb{K}^n \times \mathbb{K}^n$. Die Funktion

$$f(x, y) := \|x - y\|$$

ist stetig auf der kompakten Menge $K_1 \times K_2 \subset \mathbb{K}^n \times \mathbb{K}^n$. Dies folgt aus der Abschätzung

$$\big|\|x - y\| - \|x' - y'\|\big| \le \|x - y - x' + y'\| \le \|x - x'\| + \|y - y'\|.$$

Dann gibt es Punkte $a \in K_1$ und $b \in K_2$ mit

$$\|a - b\| = \inf_{x\in K_1, y\in K_2} \|x - y\|,$$

d. h.: Damit ist der „Abstand“ $d(K_1, K_2) := \|a - b\|$ der Mengen K_1 und K_2 definiert. Im Fall $K_1 \cap K_2 = \emptyset$ folgt $\inf_{x\in K_1, y\in K_2} \|x - y\| > 0$ (Übungsaufgabe). Insbesondere für eine einpunktige Menge $K_1 = \{a\}$ ist dann $b \in K_2$ eine sog. „Projektion“ des Punktes a auf die Menge K_2. Diese Projektion ist i. Allg. nicht eindeutig bestimmt.

Satz 2.3 (Gleichmäßige Stetigkeit): *Eine stetige Funktion $f : D \subset \mathbb{K}^n \to \mathbb{K}$ ist auf einer kompakten Menge $K \subset D$ „gleichmäßig stetig", d. h.: Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass für $x, y \in K$ gilt:*

$$\|x - y\| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon. \tag{2.1.3}$$

Beweis: Die Argumentation ist analog zum eindimensionalen Fall. Angenommen, f sei nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$ derart, dass zu jedem $k \in \mathbb{N}$ Punkte $x^{(k)}, y^{(k)} \in D$ existieren mit

$$\|x^{(k)} - y^{(k)}\| < \frac{1}{k}, \qquad |f(x^{(k)}) - f(y^{(k)})| \geq \varepsilon\,. \tag{2.1.4}$$

Wegen der Kompaktheit von K besitzt die Folge $(x^{(k)})_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(x^{(k_j)})_{j\in\mathbb{N}}$ mit Limes $x \in K$. Wegen $|x^{(k)} - y^{(k)}| < 1/k$ ist auch $x = \lim_{j\to\infty} y^{(k_j)}$. Da f stetig ist, folgt daraus

$$|f(x^{(k_j)}) - f(y^{(k_j)})| \to |f(x) - f(x)| = 0 \quad (j \to \infty),$$

im Widerspruch zu (2.1.4). Q.E.D.

Definition 2.2: *Eine Folge von Funktionen $f_k : D \subset \mathbb{K}^n \to \mathbb{K}, k \in \mathbb{N}$ konvergiert „punktweise" gegen eine Funktion $f : D \to \mathbb{K}$, wenn für alle $x \in D$ gilt:*

$$f_k(x) \to f(x) \quad (k \to \infty).$$

Sie konvergiert „gleichmäßig", wenn gilt:

$$\sup_{x\in D} |f_k(x) - f(x)| \to 0 \quad (k \to \infty).$$

Satz 2.4 (Gleichmäßige Konvergenz): *Konvergiert eine Folge stetiger Funktionen $f_k : D \subset \mathbb{K}^n \to \mathbb{K}, k \in \mathbb{N}$, gleichmäßig gegen eine Funktion $f : D \to \mathbb{K}$, so ist auch diese stetig.*

Beweis: Die Argumentation ist analog zum eindimensionalen Fall. Seien ein Punkt $x \in D$ sowie ein $\varepsilon > 0$ beliebig vorgegeben. Wegen der gleichmäßigen Konvergenz gibt es ein $n = n(\varepsilon) \in \mathbb{N}$, so dass

$$\sup_{y\in D} |f_n(y) - f(y)| < \tfrac{1}{3}\varepsilon.$$

Da f_n stetig ist, gibt es ein $\delta > 0$, so dass für alle $y \in D$ mit $\|x - y\| < \delta$ gilt:

$$|f_n(x) - f_n(y)| < \tfrac{1}{3}\varepsilon.$$

Für alle solche $y \in D$ folgt damit

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < \varepsilon,$$

d. h.: Die Funktion f ist in x stetig. Q.E.D.

Bemerkung 2.1: Wir haben gesehen, dass die Beweise der Sätze von der Beschränktkeit, vom Extremum, der gleichmäßigen Stetigkeit sowie der gleichmäßigen Konvergenz für Funktionen in mehreren Variablen praktisch identisch sind mit den entsprechenden Beweisen für Funktionen in einer Variable. Dies legt nahe, dass es sich hierbei um Aussagen handelt, die in noch viel allgemeinerem Kontext gültig sind. Tatsächlich gelten analoge Sätze allgemein für Funktionen auf kompakten Mengen in einem normierten oder metrischen Raum $(V, \|\cdot\|)$ bzw. $(X, d(\cdot,\cdot))$. Dabei sind die Begriffe „offen“, „abgeschlossen“ und „kompakt“ ganz analog wie im normierten Raum $\mathbb{K}^n$ definiert. Eine ähnliche Allgemeingültigkeit hat auch der weiter unten formulierte Zwischenwertsatz.

Definition 2.3: *i) Eine Teilmenge $M \subset G \subset \mathbb{K}^n$ heißt „relativ-offen (bzgl. G)“, wenn zu jedem Punkt $a \in M$ eine Kugelumgebung $K_r(a)$ mit $K_r(a) \cap G \subset M$ existiert.*

ii) Eine Teilmenge $M \subset G \subset \mathbb{K}^n$ heißt „relativ-abgeschlossen (bzgl. G)“, wenn die Menge $M^c \cap G \subset G$ relativ offen bzgl. G ist.

iii) Eine Menge $G \subset \mathbb{K}^n$ heißt „zusammenhängend“, wenn es keine relativ-offene Zerlegung $G = U \cup V$ gibt mit $U, V \neq \emptyset$ und $U \cap V = \emptyset$.

iv) Eine offene und zusammenhängende Menge $G \subset \mathbb{K}^n$ heißt „Gebiet“.

v) Eine (beliebige) Teilmenge $M \subset \mathbb{K}^n$ heißt „konvex”, wenn mit je zwei Punkten $x, x' \in M$ auch jede Linearkombination der Art $\lambda x + (1-\lambda)x'$ für $\lambda \in [0,1]$ (d. h. geometrisch die gesamte „Verbindungslinie“ zwischen x und x') in M enthalten ist. Eine abgeschlossene Teilmenge $M \subset \mathbb{K}^n$ heißt „strikt konvex“, wenn jede der Linearkombinationen $\lambda x + (1-\lambda)x'$ für $\lambda \in (0,1)$ im offenen Kern M^o liegt.

Lemma 2.3 (Konvexe Mengen): *i) Die Einheitskugel $K_1(0)$ im $\mathbb{K}^n$ bzgl. einer beliebigen Vektornorm $\|\cdot\|$ is konvex.*

ii) Die abgeschlossene Einheitskugel $\overline{K_1(0)}$ im $\mathbb{K}^n$ bzgl. jeder der l_p-Normen $\|\cdot\|_p$ für $1 < p < \infty$ is strikt konvex. Für die Extremfälle $p = 1$ und $p = \infty$ ist $\overline{K_1(0)}$ nicht strikt konvex.

Beweis: i) Seien $x, x' \in K_1(0)$. Für $\lambda \in [0,1]$ gilt dann mit Hilfe der Dreiecksungleichung

$$\|\lambda x + (1-\lambda)x'\| \le \lambda\|x\| + (1-\lambda)\|x'\| = \lambda + (1-\lambda) = 1,$$

d. h.: Die Punkte $\lambda x + (1-\lambda)x'$ liegen in $K_1(0)$. Also ist $K_1(0)$ konvex.

ii) Für den Beweis wird auf die einschlägige Literatur verwiesen. Q.E.D.

Bemerkung 2.2: Der Begriff der „relativen Offenheit“ einer Menge M bzgl. einer Obermenge $G \subset \mathbb{K}^n$ wird benötigt, da wir „offen“ nur für Mengen im ganzen normierten Raum $\mathbb{K}^n$ definiert haben. Wird die Teilmenge $M \subset \mathbb{K}^n$ als eigenständiger metrischer Raum mit der Metrik $d(x,y) := \|x-y\|_2$ aufgefasst, so sind die „relativ-offenen“ Teilmengen $M \subset G \subset \mathbb{K}^n$ gerade die „offenen“ Teilmengen dieses metrischen Raumes. Als Beispiel betrachten wir die Einheitssphäre $S_1(0) := \{x \in \mathbb{K}^n : \|x\|_2 = 1\}$ im $\mathbb{K}^n$, welche als Teilmenge von $\mathbb{K}^n$ abgeschlossen ist. Für eine offene Menge $M \subset \mathbb{K}^n$ ist dann der nichtleere Schnitt $M \cap S_1(0)$ als Teilmenge von $\mathbb{K}^n$ weder offen noch abgeschlossen, aber als Teilmenge von $S_1(0)$ ist er relativ-offen. Das Extrembeispiel ist hier die Menge $M = S_1(0)$, welche, obwohl im $\mathbb{K}^n$ abgeschlossen, dennoch bzgl. ihrer selbst sowohl *relativ offen* als auch *relativ-abgeschlossen* ist. Der Schnitt $M \cap G$ einer *offene* Menge $M \subset \mathbb{K}^n$ mit $G \subset \mathbb{K}^n$ ist stets *relativ-offen* bzgl. G. Umgekehrt ist auch der Schnitt $M \cap G$ einer *abgeschlossenen* Menge $M \subset \mathbb{K}^n$ mit $G \subset \mathbb{K}^n$ stets *relativ-abgeschlossen* bzgl. G.

Bemerkung 2.3: Die obige Definition der Eigenschaft „zusammenhängend“ von Mengen wird üblicherweise auch mit „topologisch zusammenhängend“ bezeichnet. Damit unterscheidet man diesen Zusammenhangsbegriff von der intuitiv zunächst naheliegenderen Eigenschaft „wegzusammenhängend“. Dabei nennt man eine Teilmenge $M \subset \mathbb{K}^n$ „wegzusammenhängend“, wenn es zu je zwei Punkten $x, x' \in M$ einen verbindenden „Weg“ (bzw. „parametrisierte Kurve“) gibt, der ganz in M verläuft. Man kann im $\mathbb{R}^2$ Beispiele von Mengen angeben, die zwar topologisch zusammenhängend aber *nicht* wezusammenhängend sind (Übungsaufgabe).

Bemerkung 2.4: Bei der Verwendung der Begriffe „abgeschlossen“ und „relativ-abgeschlossen“ für allgemeine normierte Räume $(V, \|\cdot\|)$ oder metrische Räume $(X, d(\cdot,\cdot))$ ist etwas Vorsicht geboten. Die Definitionen dieser Begriffe können zwar direkt vom $\mathbb{K}^n$ übernommen werden, aber bei der Charakterisierung dieser Eigenschaften über Folgenkonvergenz muss beachtet werden, dass der zugrunde liegende normierte Raum $(V, \|\cdot\|)$ eventuell *nicht vollständig* ist, d. h. dass nicht jede Cauchy-Folge in ihm einen Limes hat. Für eine *abgeschlossene* oder „relativ-abgeschlossene“ Teilmenge $A \subset V$ muss daher nur der Limes einer solchen Cauchy-Folge aus A auch in A enthalten sein, für welche überhaupt ein Limes in V existiert. Im endlich dimensionalen (und damit automatisch *vollständigen*) Raum $\mathbb{K}^n$ besteht diese Unterscheidung natürlich nicht; ebenso nicht im unendlich dimensionalen (vollständigen) Banach-Raum $(C[a,b], \|\cdot\|_\infty)$ mit der Maximumnorm $\|\cdot\|_\infty$. Dagegen ist der normierte Raum $(C[a,b], \|\cdot\|_2)$ mit der L^2-Norm $\|\cdot\|_2$ bekanntlich *nicht* vollständig, so dass in seinen *abgeschlossenen* Teilmengen nicht jede Cauchy-Folge einen Limes zu haben braucht.

Bemerkung 2.5: Mengen $M \subset \mathbb{K}^n$ mit disjunkten Komponenten, welche einen positiven Abstand haben, können offenbar nicht zusammenhängend sein:

$$M = M_1 \cup M_2, \quad M_1, M_2 \neq \emptyset, \quad d(M_1, M_2) = \inf_{x \in M_1, y \in M_2} \|x - y\| > 0.$$

Bei sich „berührenden“ Mengen können beide Situationen auftreten. Die Menge bestehend aus den beiden (offenen) Einheitskugeln

$$M := K_1(0) \cup K_1(2) \subset \mathbb{K}^2$$

ist, obwohl $d(K_1(0), K_1(2)) = 0$, *nicht* zusammenhängend, da ihr der „verbindende“ Punkt $(1, 0)$ fehlt. Dagegen ist die Vereinigung der zugehörigen Abschlüsse

$$M := \overline{K_1(0)} \cup \overline{K_1(2)} \subset \mathbb{K}^2$$

zusammenhängend.

Lemma 2.4: *Für stetige Funktionen* $f : \overline{D} \subset \mathbb{K}^n \to \mathbb{K}$ *gilt:*

i) Das Urbild $f^{-1}(O)$ *einer relativ-offenen Menge* $O \subset f(D)$ *ist relativ-offen in* D *.*

ii) Das Urbild $f^{-1}(A)$ *einer abgeschlossenen Menge* $A \subset f(\overline{D})$ *ist abgeschlossenen.*

iii) Das Bild $f(K)$ *einer kompakten Menge* $K \subset D$ *ist kompakt.*

iv) Das Bild $f(G)$ *einer zusammenhängenden Menge* $G \subset D$ *ist zusammenhängend.*

Beweis: i) Eine relativ-offene Menge $O \subset f(D)$ ist (relative) Umgebung eines jeden ihrer Punkte $f(a)$, d. h.: Es gibt eine (relative) Kugelumgebung $K_\varepsilon(f(a)) \cap f(D) \subset O$. Zu diesem $\varepsilon > 0$ gibt es aufgrund der Stetigkeit von f ein $\delta > 0$, so dass für die (relative) Kugelumgebung $K_\delta(a) \cap D$ gilt

$$f(K_\delta(a) \cap D) \subset K_\varepsilon(f(a)) \cap f(D) \subset O.$$

Also ist $K_\delta(a) \cap D \subset f^{-1}(O)$ und $f^{-1}(O)$ demnach relativ-offen.

ii) Wir verwenden die Charakterisierung von „abgeschlossen" über die Folgenkonvergenz. Sei $(x^{(k)})_{k\in\mathbb{N}}$ irgendeine konvergente Folge in $f^{-1}(A)$ mit Limes $x \in \overline{D}$. Dann ist die Bildfolge $(f(x^{(k)}))_{k\in\mathbb{N}}$ konvergent mit $f(x^{(k)}) \to f(x)$ $(k \to \infty)$. Wegen der Abgeschlossenheit von A in $f(\overline{D})$ ist $f(x) \in A$, d. h.: $x \in f^{-1}(A)$. Also ist $f^{-1}(A)$ abgeschlossen.

iii) Wir zeigen, dass $f(K) \subset \mathbb{K}$ beschränkt und abgeschlossen und damit nach dem Satz von Bolzano-Weierstraß auch kompakt ist. Satz 2.1 ergibt die Beschränktheit von $f(K)$. Zum Beweis der Abgeschlossenheit betrachten wir eine beliebige konvergente Folge $(y^{(k)})_{k\in\mathbb{N}}$ aus $f(K)$ mit Limes $y \in \mathbb{K}$. Die Urbildfolge $(x^{(k)})_{k\in\mathbb{N}}$ hat dann wegen der Kompaktheit von K eine konvergente Teilfolge $(x^{(k_j)})_{j\in\mathbb{N}}$ mit Limes $x \in K$. Wegen der Stetigkeit von f ist $f(x) = \lim_{j\to\infty} f(x^{(k_j)}) = y$ und somit $y \in f(K)$. Also ist $f(K)$ abgeschlossen.

Bemerkung: Eine etwas elegantere Beweisführung verwendet den Heine-Borelschen Überdeckungssatz wie folgt: Sei $\{O_\lambda\}_{\lambda\in\Lambda}$ irgendeine offene Überdeckung von $f(K)$. Dann bilden die relativ-offenen Mengen

$$O'_\lambda := \{x \in D : \; f(x) \in O_\lambda\} \subset D$$

eine Überdeckung von K. Nach dem Satz von Heine-Borel (für relativ-offene Überdeckungen) wird die kompakte Menge K bereits durch endlich viele der O'_λ überdeckt. Dasselbe gilt dann auch für $f(K)$, d. h.: $f(K)$ ist kompakt.

iv) Wäre $f(G)$ nicht zusammenhängend, so gäbe es nicht leere, relativ-offene Mengen $U, V \subset \mathbb{K}$ mit $f(G) = U \cup V$ und $U \cap V = \emptyset$. Die Urbildmengen $U' := \{x \in G : f(x) \in U\}$ und $V' := \{x \in G : \; f(x) \in V\}$ sind dann ebenfalls disjunkt, nicht leer, nach (i) relativ-offen, und es gilt $G = U' \cup V'$, im Widerspruch zur Annahme, dass G zusammenhängend ist. Folglich ist $f(G)$ zusammenhängend. Q.E.D.

Satz 2.5 (Zwischenwertsatz): *Sei* $f : D \subset \mathbb{K}^n \to \mathbb{R}$ *stetig und* D *zusammenhängend. Dann nimmt* f *für je zwei Punkte* $a, b \in D$ *jeden Wert zwischen* $f(a)$ *und* $f(b)$ *an. Insbesondere hat* f *im Falle* $f(a)f(b) < 0$ *also eine Nullstelle in* D *.*

Beweis: Das Resultat von Lemma 2.4 Teil (iv) impliziert, dass mit D auch der Bildbereich $f(D) \subset \mathbb{R}$ zusammenhängend ist. Wir wollen zeigen, dass $f(D)$ dann notwendig ein (zusammenhängendes) Intervall ist. Hieraus ergibt sich dann unmittelbar die Richtigkeit der Behauptungen. Angenommen, $f(D)$ ist kein Intervall. Dann gibt es Punkte $f(x), f(y) \in f(D)$ und zwischen diesen einen Punkt $z \notin f(D)$. Die Mengen $U := f(D) \cap (-\infty, z)$ und $V := f(D) \cap (z, \infty)$ sind disjunkt, nicht leer und relativ zu $f(D)$ offen, und es gilt $U \cup V = f(D)$. Somit ist $f(D)$ nicht zusammenhängend, im Widerspruch zur Annahme . Q.E.D.

2.2 Vektor- und matrixwertige Funktionen

Im Folgenden betrachten wir Funktionen (bzw. Abbildungen)

$$f : D \subset \mathbb{K}^n \to \mathbb{K}^m, \quad f : D \subset \mathbb{K}^n \to \mathbb{K}^{r\times s}, \quad f : D \subset \mathbb{K}^{m\times n} \to \mathbb{K}^{r\times s}.$$

Beispiel 2.2: Einfachste Beispiele vektor- und matrixwertiger Funktionen sind:

i) Translation um Vektor $b \in \mathbb{R}^2$ und anschließende Drehung um Winkel $\varphi \in (0, 2\pi]$ (im Uhrzeigersinn) in der Ebene $\mathbb{R}^2$:

$$f(x) := \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

ii) Dyadisches Produkt $f : \mathbb{K}^n \to \mathbb{K}^{n\times n}$ eines Vektors $v \in \mathbb{K}^n$ mit sich selbst:

$$f(v) := (v_i \overline{v}_j)_{i,j=1}^n.$$

iii) Quadrierung $f : \mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}$ von Matrizen $A = (a_{ij})_{i,j=1}^n$:

$$f(A) := A^2 = \Big(\sum_{k=1}^n a_{ik} a_{kj} \Big)_{i,j=1}^n.$$

Mit Hilfe von Normen auf $\mathbb{K}^n$ und $\mathbb{K}^{n\times n}$ lassen sich die Begriffe „beschränkt" und „stetig" für solche Abbildungen ganz analog wie oben für Funktionen $f : D \subset \mathbb{K}^n \to \mathbb{K}$ definieren. Insbesondere gelten sinngemäß die Sätze von der Beschränktheit und gleichmäßigen Stetigkeit, d. h.: Solche stetige Abbildungen sind auf kompakten Mengen gleichmäßig stetig und beschränkt. Wir stellen fest, dass eine Abbildung $f : D \subset \mathbb{K}^n \to \mathbb{K}^m$ oder $f : D \subset \mathbb{K}^n \to \mathbb{K}^{r\times s}$ genau dann stetig ist, wenn alle ihre Komponentenfunktionen $f_k : D \to \mathbb{K}$ bzw. $f_{jk} : D \to \mathbb{K}$ stetig sind. Im Folgenden sei $\|\cdot\|$ irgendeine Norm auf $\mathbb{K}^n$.

Lemma 2.5: *Für stetige Funktionen $g : D \subset \mathbb{K}^n \to B \subset \mathbb{K}^m$ und $f : B \to \mathbb{K}^r$ ist auch die Komposition $f \circ g : D \subset \mathbb{K}^n \to \mathbb{K}^r$ stetig.*

Beweis: Die Argumentation ist wieder analog zum eindimensionalen Fall. Sei $x \in D$ und $x^{(k)} \in D$ mit $x^{(k)} \to x$ $(k \to \infty)$. Wegen der Stetigkeit von g konvergiert $y^{(k)} = g(x^{(k)}) \to g(x) = y$ $(k \to \infty)$ und dann wegen der Stetigkeit von f:

$$(f \circ g)(x^{(k)}) = f(g(x^{(k)})) = f(y^{(k)}) \to f(y) \quad (x^{(k)} \to x).$$

Mit $f(y) = f(g(x)) = (f \circ g)(x)$ ergibt sich die Behauptung. Q.E.D.

Lemma 2.6: *Die auf einer beschränkten, abgeschlossenen (d. h. kompakten) Teilmenge $D \subset \mathbb{K}^n$ definierte Funktion $f : D \to B \subset \mathbb{K}^n$ sei injektiv und stetig. Dann ist auch ihre Umkehrfunktion $f^{-1} : B \to D$ stetig.*

Beweis: Die Argumentation ist wieder analog zum eindimensionalen Fall. Sei $(y^{(k)})_{k\in\mathbb{N}}$ eine Folge in B mit $y^{(k)} \to y \in B$ $(k \to \infty)$. Wir haben zu zeigen, dass dann $x^{(k)} := f^{-1}(y^{(k)}) \to f^{-1}(y) =: x$ $(k \to \infty)$. Die Urbildfolge $(x^{(k)})_{k\in\mathbb{N}}$ ist beschränkt, da in der beschränkten Menge D enthalten. Sei $(x^{(k_j)})_{j\in\mathbb{N}}$ eine konvergente Teilfolge mit $x^{(k_j)} \to \xi \in D$. Wegen der Stetigkeit von f konvergiert dann $f(x^{(k_j)}) \to f(\xi)$. Es gilt aber auch $f(x^{(k_j)}) = y^{(k_j)} \to y$, d. h.: $f(x) = f(\xi)$. Wegen der Injektivität von f folgt $\xi = x$. Also sind alle Häufungswerte der (beschränkten) Folge $(x^{(k)})_{k\in\mathbb{N}}$ gleich x, so dass notwendig $x^{(k)} \to x$ $(k \to \infty)$. Q.E.D.

Bemerkung 2.6: In physikalischer Sprache werden skalarwertige Abbildungen $f : \mathbb{R}^n \to \mathbb{R}$ „Skalarfelder“ (oder „Tensorfelder 0-ter Stufe“), vektorwertige Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^n$ „Vektorfelder“ (oder „Tensorfelder 1-ter Stufe“) und matrixwertige Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ „Tensorfelder“ (oder „Tensorfelder 2-ter Stufe“)“ genannt. Beispiele für Skalarfelder sind „Temperatur“, „Dichte“ und „Druck“; „Geschwindigkeit, „Schwerkraft“ und „elektrische Feldstärke“ sind Vektorfelder, während „Verzerrungen“ und „Spannungen“ durch Tensorfelder 2-ter Stufe beschrieben werden. Abbildungen $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ treten z. B. als sog. „Materialtensoren“ in der Elastizitätstheorie auf. Die Bezeichnung „Tensor“ kommt dabei eigentlich den durch diese Skalar-, Vektor- und Matrizenfeldern bzgl. des willkürlich gewählten, kartesischen Koordinatensystems dargestellten Abbildungen zu. Da letztere aber nicht von der Wahl des Koordinatensystems abhängen dürfen, gehören zur Tensoreigenschaft von Skalar-, Vektor- und Matrizenfeldern noch gewisse Invarianzeigenschaften gegenüber Koordinatentransformationen (→ Lineare Algebra).

2.2.1 Lineare und nichtlineare Gleichungssysteme

Wir betrachten allgemeine (quadratische) Gleichungssystme der Form

$$\begin{aligned} f_1(x_1, \dots, x_n) &= b_1, \\ &\vdots \\ f_n(x_1, \dots, x_n) &= b_n, \end{aligned} \tag{2.2.5}$$

bzw. in kompakter Schreibweise

$$f(x) = b, \tag{2.2.6}$$

mit der Abbildung $f = (f_1, \dots, f_n) : D \subset \mathbb{K}^n \to \mathbb{K}^n$ und der vorgegebenen rechten Seite $b = (b_1, \dots, b_n) \in \mathbb{K}^n$. Für Gleichungssysteme dieser Art lassen sich in der Regel keine Lösungen in expliziter Form $x = f^{-1}(b)$ angeben. Man kann vielmehr nur hoffen, eine Folge von Vektoren $x^{(k)}$ zu konstruieren, welche gegen eine Lösung konvergiert. Als Ausgangspunkt für eine solche Iteration macht man den Ansatz

$$g(x) := x - \sigma(f(x) - b) = x,$$

mit einem geeigneten $\sigma \in \mathbb{K} \setminus \{0\}$ d. h.: man sucht einen sog. „Fixpunkt“ der neuen Abbildung $g : D \to \mathbb{K}^n$. Dies motiviert eine sog. „Fixpunktiteration“ (oder „sukzessive Approximation“)

$$x^{(k)} = g(x^{(k-1)}), \quad k \in \mathbb{N}, \tag{2.2.7}$$

welche mit einem geeigneten Startvektor $x^{(0)} \in D$ beginnt. Mit f ist auch g stetig. Im Falle der Konvergenz $x^{(k)} \to x$ $(k \to \infty)$ konvergiert also $g(x^{(k-1)}) \to g(x)$ $(k \to \infty)$, d. h. der Limes x ist Fixpunkt,

$$x = g(x) = x - \sigma(f(x) - b),$$

und löst damit die Gleichung $f(x) = b$. Die Frage ist also, unter welchen Bedingungen die Konvergenz der Fixpunktiteration garantiert werden kann.

Definition 2.4: *Eine Abbildung $g : D \subset \mathbb{K}^n \to \mathbb{K}^n$ heißt „Lipschitz-stetig“, wenn mit einer Konstante $L > 0$, der sog. „Lipschitz-Konstante“, gilt:*

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad x, y \in D. \tag{2.2.8}$$

Im Falle $L < 1$ heißt g „Kontraktion“ (bzgl. der gewählten Norm $\|\cdot\|$).

Für Kontraktionen gilt der folgende fundamentale „Banachsche Fixpunktsatz“, der für eine großen Klasse nichtlinearer Gleichungen neben einer Existenzaussage für Lösungen vor allem auch ein Verfahren zu deren näherungsweisen Berechnung liefert. Dieser Satz kann ganz allgemein in Banachschen Räumen (oder sogar in vollständigen metrischen Räumen) formuliert werden. Wir beschränken uns hier aber auf seine einfache Variante im $\mathbb{K}^n$ und verlagern Verallgemeinerungen in Übungsaufgaben.

Satz 2.6 (Banachscher Fixpunktsatz): *Sei $g : D \subset \mathbb{K}^n \to \mathbb{K}^n$ eine Abbildung, für welche die folgenden Bedingungen erfüllt sind:*

i) g bildet eine abgeschlossene Teilmenge $M \subset D$ in sich ab.

ii) Auf M ist g eine Kontraktion mit Lipschitz-Konstante $q \in (0,1)$.

Dann besitzt g in M genau einen Fixpunkt x^*, und für jeden Startpunkt $x^{(0)} \in M$ konvergiert die Folge der durch (2.2.7) definierten Iterierten $x^{(k)} \in M$ gegen diesen Fixpunkt $x^* \in M$ mit der Fehlerabschätzung

$$\|x^{(k)} - x^*\| \le \frac{q^k}{1-q}\|x^{(1)} - x^{(0)}\|. \tag{2.2.9}$$

Beweis: i) Wir zeigen zunächst die Eindeutigkeit des Fixpunktes (falls er existiert). Seien $x, x' \in M$ zwei Fixpunkte. Dann gilt wegen der Kontraktionseigenschaft

$$\|x - x'\| = \|g(x) - g(x')\| \le q\|x - x'\|,$$

was wegen $q < 1$ notwendig $x = x'$ impliziert.

ii) Da g die Menge M in sich abbildet, sind für jeden Startpunkt $x^{(0)}$ die Iterierten $x^{(k)}$, $k \in \mathbb{N}$, wohl definiert. Wir wollen zeigen, dass die Folge $(x^{(k)})_{k\in\mathbb{N}}$ eine Cauchy-Folge ist. Dann hat sie einen Limes $x^* \in \mathbb{K}^n$, der wegen der Abgeschlossenheit von M auch in M liegt. Wegen der Stetigkeit von g folgt

$$x^* = \lim_{k\to\infty} x^{(k)} = \lim_{k\to\infty} g(x^{(k-1)}) = g(\lim_{k\to\infty} x^{(k-1)}) = g(x^*),$$

d. h.: $x^* \in M$ ist Fixpunkt von g. Für beliebige $k, m \in \mathbb{N}$ gilt:

$$\begin{aligned}
\|x^{(k+m)} - x^{(k)}\| &= \|x^{(k+m)} - x^{(k+m-1)} + \ldots + x^{(k+1)} - x^{(k)}\| \\
&\le \|x^{(k+m)} - x^{(k+m-1)}\| + \cdots + \|x^{(k+1)} - x^{(k)}\| \\
&= \|g^{m-1}(x^{(k+1)}) - g^{m-1}(x^{(k)})\| + \cdots + \|x^{(k+1)} - x^{(k)}\| \\
&\le (q^{m-1} + q^{m-2} + \cdots + 1)\|x^{(k+1)} - x^{(k)}\| \\
&= (q^{m-1} + q^{m-2} + \cdots + 1)\|g^k(x^{(1)}) - g^k(x^{(0)})\| \\
&\le (q^{m-1} + q^{m-2} + \cdots + 1)q^k\|x^{(1)} - x^{(0)}\| \\
&= \frac{1-q^m}{1-q} q^k\,\|x^{(1)} - x^{(0)}\| \le \frac{q^k}{1-q}\,\|x^{(1)} - x^{(0)}\|.
\end{aligned}$$

Wegen $q < 1$ wird die rechte Seite kleiner als jedes vorgegebene $\varepsilon > 0$, wenn nur k groß genug ist. Also ist $(x^{(k)})_{k\in\mathbb{N}}$ eine Cauchy-Folge.

iii) Zum Nachweis der Fehlerabschätzung (2.2.9) betrachten wir in der obigen Abschätzung den Grenzprozess $m \to \infty$ und erhalten wegen $x^{(k+m)} \to x^*$ $(m \to \infty)$ das gewünschte Ergebnis. Q.E.D.

Bemerkung 2.7: Im Beweis des Banachschen Fixpunktsatzes werden die endliche Dimension sowie im Prinzip auch die Vektorraumstruktur des zugrunde liegenden normierten Raumes nicht verwendet. Er lässt sich also ohne Probleme auf die Situation eines allgemeinen normierten Raumes $(V, \|\cdot\|)$ oder sogar metrischen Raumes $(X, d(\cdot,\cdot)$ übertragen. Allerdings muss der Grundraum *vollständig* sein, damit für die Folge der sukzessiven Approximationen überhaupt die Existenz eines Limes gesichert ist. (Übungsaufgabe)

Bemerkung 2.8: Die Anwendung des obigen Fixpunktprinzips für die Lösung konkreter nichtlinearer Gleichungssysteme ist Gegenstand von Texten zur „Numerischen Mathematik“. Eine Schwierigkeit ist dabei überhaupt die Bestimmung einer abgeschlossenen Teilmenge $M \subset D$, welche von g in sich abgebildet wird. Dies erübrigt sich aber in dem Fall, dass g sogar auf ganz $\mathbb{K}^n$ definiert ist. Liegt dann auch noch die Kontraktionseigenschaft vor, so ist wegen der Vollständigkeit von $\mathbb{K}^n$ der Banachsche Fixpunktsatz direkt anwendbar. Im Folgenden werden wir zwei Klassen von Problemen kennenlernen, bei denen diese Idealsituation vorliegt.

Anwendung 2.2.1 (Lineare Gleichungssysteme): Wir betrachten zunächst als einfachsten Spezialfall ein *lineares* Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n. \end{aligned} \tag{2.2.10}$$

mit gegebener rechter Seite $b = (b_i)_{i=1}^n \in \mathbb{K}^n$ und Koeffizientenmatrix $A = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n\times n}$. In kompakter Form lautet dies

$$Ax = b. \tag{2.2.11}$$

Die Matrix A sei als regulär angenommen, so dass eine eindeutig Lösung existiert. Der obigen Philosophie folgend machen wir den Ansatz einer Fixpunktgleichung

$$g(x) := x - \sigma(Ax - b) = x$$

mit einem $\sigma \in \mathbb{K} \setminus \{0\}$. Die zugehörige Fixpunktiteration

$$x^{(k)} = x^{(k-1)} - \sigma(Ax^{(k-1)} - b), \quad k \in \mathbb{N},$$

konvergiert dann nach dem Banachschen Fixpunktsatz für jeden Startvektor $x^{(0)} \in \mathbb{K}^n$ gegen einen Fixpunkt von g bzw. gegen die Lösung des Gleichungssystems, wenn g eine Kontraktion ist. Wir verwenden jetzt die euklidische Norm $\|\cdot\|_2$. Wegen

$$\begin{aligned} \|g(x) - g(y)\|_2 &= \|x - \sigma(Ax - b) - y + \sigma(Ay - b)\|_2 \\ &= \|(I - \sigma A)(x - y)\|_2 \le \|I - \sigma A\|_2 \, \|x - y\|_2 \end{aligned}$$

ist dies der Fall, wenn $q := \|I - \sigma A\|_2 < 1$ ist.

Korollar 2.1: *Seien $A \in \mathbb{K}^{n\times n}$ hermitesch und positiv definit (und damit regulär) und $b \in \mathbb{K}^n$ gegeben. Dann konvergiert die Fixpunktiteration (sog. „Richardson*[1]*-Iteration“)*

$$x^{(k)} = x^{(k-1)} - \|A\|_\infty^{-1}\big(Ax^{(k-1)} - b\big), \quad k \in \mathbb{N}, \tag{2.2.12}$$

für jeden Startwert $x^{(0)}$ gegen die Lösung des Gleichungssystems

$$Ax = b. \tag{2.2.13}$$

[1] Lewis Fry Richardson (1881–1953): Englischer Mathematiker und Physiker; wirkte an verschiedenen Institutionen in England und Schottland; typischer „angewandter Mathematiker“; leistete Pionierbeiträge zur Modellierung und Numerik in der Wettervorhersage.

Beweis: Die Eigenwerte der Matrix A sind positiv und beschränkt: $0 < \lambda \le \|A\|_\infty$. Folglich gilt für die Eigenwerte der Matrix $I - \|A\|_\infty^{-1} A$:

$$0 \le 1 - \|A\|_\infty^{-1} \lambda < 1.$$

Dies impliziert dann für die Spektralnorm

$$\big\| I - \|A\|_\infty^{-1} A \big\|_2 < 1.$$

Satz 2.6 liefert daher die behauptete Konvergenzaussage. Q.E.D.

Korollar 2.1 zeigt einen einfachen Weg zur sukzessiven Approximation von Lösungen von Gleichungssystemen mit hermitescher, positiv definiter Koeffizientenmatrix. Dieses ist der proto-typische, aber auch langsamste Vertreter einer großen Klasse von sog. „iterativen“ Lösungsverfahren (im Gegensatz zu sog. „direkten“ Verfahren wie z. B. der Gauß-Elimination). Für nicht-hermitesche oder indefinite Matrizen ist die einfache Richardson-Iteration i. Allg. nicht geeignet; in diesen Fällen sind raffiniertere Iterationen erforderlich. Die Konstruktion und Analyse solcher Verfahren ist Gegenstand von Texten zur „Numerischen Mathematik“.

Anwendung 2.2.2 (Nichtlineare Gleichungssysteme): Die Lösung *nichtlinearer* Gleichungssysteme ist i. Allg. viel schwieriger als die *linearer* Systeme. Um ein entsprechendes Analogon zu Korollar 2.1 zu erhalten, führen wir den Begriff der „monotonen“ Abbildung ein. Dies ist eine direkte Verallgemeinerung der Eigenschaft „positiv definit“ für Matrizen auf nichtlineare Abbildungen.

Definition 2.5: *Eine Abbildung $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ heißt „stark monoton“, wenn es eine Konstante $m > 0$ gibt, so dass für $x, y \in D$ gilt:*

$$(f(x) - f(y), x - y)_2 \ge m\|x - y\|_2^2 . \tag{2.2.14}$$

Für Funktionen $f : \mathbb{R} \to \mathbb{R}$ impliziert die eben definierte „starke Monotonie“ die Gültigkeit der Beziehung

$$(f(x) - f(y))(x - y) > 0,$$

woraus im Fall $x > y$ notwendig $f(x) > f(y)$ folgt, d. h. die „strenge“ Monotonie von f im ursprünglichen Sinne.

Korollar 2.2: *Seien $f : \mathbb{R}^n \to \mathbb{R}^n$ eine Lipschitz-stetige, stark monotone Abbildung mit Lipschitz-Konstante L und Monotoniekonstante $m > 0$ sowie ein $b \in \mathbb{R}^n$ gegeben. Dann hat die Gleichung*

$$f(x) = b \tag{2.2.15}$$

eine eindeutige Lösung x^*. Für jeden Startpunkt $x^{(0)}$ konvergiert die sukzessive Iteration

$$x^{(k)} = x^{(k-1)} - \theta\big(f(x^{(k-1)}) - b\big) \tag{2.2.16}$$

für jedes $\theta \in (0, 2m/L^2)$ gegen x^*.

Beweis: i) Wir zeigen zunächst die Eindeutigkeit der Lösung (sofern sie existiert). Seien x, x' zwei Lösungen. Für diese gilt dann:

$$0 = (f(x) - b + b - f(x'), x - x')_2 = (f(x) - f(x'), x - x')_2 \geq m\|x - x'\|_2^2 ,$$

woraus $x = x'$ folgt.

ii) Die Existenz einer Lösung wird mit Hilfe des Banachschen Fixpunktsatzes gezeigt. Wir betrachten die zur gestellten Gleichung äquivalente Fixpunktgleichung

$$g(x) := x - \theta(f(x) - b) = x.$$

Wir wollen zeigen, dass die Abbildung $g : \mathbb{R}^n \to \mathbb{R}^n$ eine Kontraktion ist. Dann folgt über den Banachschen Fixpunktsatz die Existenz eines eindeutig bestimmten Fixpunktes x^* , der nach Konstruktion auch Lösung der Aufgabe (2.2.15) ist. Für beliebige $x, y \in \mathbb{R}^n$ betrachten wir

$$\begin{aligned} \|g(x) - g(y)\|_2^2 &= \|x - \theta f(x) - y + \theta f(y)\|_2^2 \\ &= \|x - y\|_2^2 - 2\theta(x - y, f(x) - f(y))_2 + \theta^2\|f(x) - f(y)\|_2^2 \\ &\leq (1 - 2m\theta + L^2\theta^2)\|x - y\|_2^2 \end{aligned}$$

Für $\theta \in (0, 2m/L^2)$ ist also g eine Kontraktion. Diese Lösung ist Limes der durch die sukzessive Iteration

$$x^{(k)} = g(x^{(k-1)})$$

erzeugten Folge $(x^{(k)})_{k\in\mathbb{N}}$ für beliebigen Startpunkt $x^{(0)} \in \mathbb{R}^n$. Q.E.D.

2.2.2 Matrixfunktionen

Mit Hilfe der Normkonvergenz von Matrixfolgen lassen sich sog. „Matrixfunktionen“ definieren, welche z. B. eine wichtige Rolle bei der Analyse von approximativen Lösungsverfahren von Differentialgleichungen spielen.

Matrixpolynome und rationale Funktionen

Für eine Matrix $A \in \mathbb{K}^{n\times n}$ und ein Polynom $p(x) = \sum_{k=0}^r a_k x^k$ vom Grad $r \geq 0$ ist das Matrixpolynom

$$p(A) := \sum_{k=0}^r a_k A^k.$$

definiert. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A mit Eigenvektor $z \in \mathbb{C}^n$, $z \neq 0$. Dann gilt

$$p(A)z = \sum_{k=0}^r a_k A^k z = \sum_{k=0}^r a_k \lambda^k z = p(\lambda)z. \tag{2.2.17}$$

Also ist $p(\lambda)$ Eigenwert von $p(A)$ mit demselben Eigenvektor z .

Lemma 2.7: *Sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix mit (ihrer Vielheiten entsprechend oft gezählten) Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ und p ein Polynom. Die Matrix $p(A)$ ist genau dann regulär, wenn keiner der Eigenwerte von A Nullstelle von p ist.*

Beweis: Ist ein Eigenwert λ von A mit Eigenvektor $z \in \mathbb{K}^n \setminus \{0\}$ Nullstelle von p, so folgt wegen (2.2.17) $p(A)z = 0$, d. h.: Die Matrix $p(A)$ ist singulär. Ist umgekehrt $p(A)$ singulär, so gibt es ein $z \neq 0$ mit $p(A)z = 0$. Sei $\{z^{(i)},\, i = 1, \dots, n\}$ eine zu den Eigenwerten $\lambda_i, i = 1, \dots, n$, gehörende Orthonormalbasis von Eigenvektoren von A. Damit gilt

$$\begin{aligned} 0 = \|p(A)z\|_2^2 &= \Big\| p(A) \sum_{i=1}^n (z, z^{(i)})_2\, z^{(i)} \Big\|_2^2 = \Big\| \sum_{i=1}^n (z, z^{(i)})_2\, p(A) z^{(i)} \Big\|_2^2 \\ &= \Big\| \sum_{i=1}^n (z, z^{(i)})_2\, p(\lambda_i) z^{(i)} \Big\|_2^2 = \sum_{i=1}^n |(z, z^{(i)})_2|^2\, |p(\lambda_i)|^2. \end{aligned}$$

Da wegen $z \neq 0$ nicht alle Produkte $(z, z^{(i)})_2$ Null sein können, muss mindestens für ein $i \in \{1, \dots, n\}$ der Eigenwert λ_i Nullstelle von p sein. Q.E.D.

Als Folgerung aus Lemma 2.7 sehen wir, dass für eine rationale Funktion $r(x) = p(x)/q(x)$ mit Polynomen $p(x)$ und $q(x)$ und eine hermitesche Matrix $A \in \mathbb{K}^{n\times n}$ die zugehörige Matrixfunktion

$$r(A) = p(A)q(A)^{-1} = q(A)^{-1}p(A)$$

wohl definiert ist, wenn kein Eigenwert von A Nullstelle des Nennerpolynoms q ist. Unter modifizierten Bedingungen lässt sich dies auch für nichthermitesche Matrizen definieren.

Wurzelfunktion

Auf analogem Wege lassen sich auch allgemeinere, nicht rationale Matrixfunktionen definieren. Z. B. erhält man für eine hermitesche, positiv-definite Matrix $A \in \mathbb{K}^{n\times n}$ mit Eigenwerten $\lambda_i \in \mathbb{R}_+$ und zugehöriger Orthonormalbasis von Eigenvektoren $\{z^{(i)},\, i = 1, \dots, n\}$ durch

$$\mathcal{B}x := \sum_{i=1}^n \lambda_i^{1/2} (x, z^{(i)})_2 z^{(i)}, \quad x \in \mathbb{K}^n, \tag{2.2.18}$$

eine lineare Abbildung $\mathcal{B} : \mathbb{K}^n \to \mathbb{K}^n$ mit der Eigenschaft

$$\mathcal{B}^2 x = \sum_{i=1}^n \lambda_i (x, z^{(i)})_2 z^{(i)} = Ax, \quad x \in \mathbb{K}^n,$$

Bei dieser Schreibweise wird die Matrix A ebenfalls als lineare Abbildung in $\mathbb{K}^n$ aufgefasst, wobei hier „Punkt im $\mathbb{K}^n$“ und zugehöriger „kartesischer Koordinatenvektor“

identifiziert werden. Die Abbildung $\mathcal{B}$ hat die Eigenschaften einer (positiven) Quadratwurzel von A. Aus der Darstellung (2.2.18) gewinnt man auch eine Matrixdarstellung $B = (b_{jk})_{j,k=1}^n$ der Abbildung $\mathcal{B}$, indem man mit den kartesischen Basisvektoren $e^{(j)}$, $j = 1,\dots,n$, bildet:

$$b_{jk} := (\mathcal{B}e^{(j)}, e^{(k)})_2 = \sum_{i=1}^n \lambda_i^{1/2}(e^{(j)}, z^{(i)})_2(z^{(i)}, e^{(k)})_2, \quad j,k = 1,\dots,n.$$

Dass diese Matrix B tatsächlich dieselbe Wirkung auf einen kartesischen Vektor $x = (x_1,\dots,x_n)$ hat wie die Abbildung $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$, sieht man wie folgt:

$$\begin{aligned}(Bx)_j &= \sum_{k=1}^n b_{jk}x_k = \sum_{k=1}^n \sum_{i=1}^n \lambda_i^{1/2}(e^{(j)}, z^{(i)})_2(z^{(i)}, e^{(k)})_2 x_k \\ &= \sum_{k=1}^n \sum_{i=1}^n \lambda_i^{1/2} z_j^{(i)} z_k^{(i)} x_k = \sum_{i=1}^n \lambda_i^{1/2}(x, z^{(i)})_2 z_j^{(i)}.\end{aligned}$$

Es gibt noch weitere Matrizen $\tilde{B} \in \mathbb{R}^{n\times n}$ mit der Eigenschaft $\tilde{B}^2 = A$, z. B. die Matrix $\tilde{B} := -A^{1/2}$. Die „positive Wurzel“ $B = A^{1/2}$ ist aber die einzige symmetrische und positiv definite. Für eine zweite „positive Wurzel“ $\tilde{B}$ gilt $\tilde{B}A = \tilde{B}^3 = A\tilde{B}$, d. h. sie kommutiert mit A. Nach Lemma 2.8 (s. weiter unten) besitzen folglich $\tilde{B}$ und A eine gemeinsame Orthonormalbasis $\{z^{(i)}, i = 1,\dots,n\}$ von Eigenvektoren. Die Eigenwerte von $\tilde{B}$ sind dann gerade $\lambda_i^{1/2} > 0$. Damit gilt für alle $x \in \mathbb{R}^n$:

$$\tilde{B}x = \sum_{i=1}^n \lambda_i^{1/2}(x, z^{(i)})_2 z^{(i)} = A^{1/2}x,$$

d. h. $\tilde{B}$ und $A^{1/2}$ stimmen überein.

Bemerkung 2.9: Für große Matrizen ist die Darstellung (2.2.18) kaum zur praktischen Berechnung der Quadratwurzel $B =: A^{1/2}$ geeignet, da man dazu alle Eigenwerte von A kennen müsste. Stattdessen kann man sich der Fixpunktiteration

$$X^k := \tfrac{1}{2}\big(X^{(k-1)} + (X^{(k-1)})^{-1}A\big), \quad k \in \mathbb{N},$$

mit einem geeigneten Startwert, z. B.: $X^{(0)} = A$, bedienen, welche beliebig gute Approximationen zu $A^{1/2}$ liefert (Übungsaufgabe). Deren Durchführung erfordert „nur“ Matrixinvertierungen.

Analytische Funktionen

Ein allgemeinerer, mehr analytischer Weg zur Definition von Matrixfunktionen bedient sich der Taylor-Entwicklung. Die Potenzreihe

$$s_\infty(x) = \sum_{k=0}^\infty a_k x^k, \quad x \in \mathbb{R},$$

habe den Konvergenzradius $\rho > 0$. Im Falle $x := \|A\|_2 < \rho$ folgt aus der Abschätzung

$$\Big\| \sum_{k=r}^{m} a_k A^k \Big\|_2 \le \sum_{k=r}^{m} |a_k| \, \|A\|_2^k \le \sum_{k=r}^{m} |a_k| x^k,$$

dass die Folge der Partialsummen $s_m(A) := \sum_{k=0}^{m} a_k A^k$ eine Cauchy-Folge in $\mathbb{K}^{n\times n}$ ist. Ihr Limes wird geschrieben als

$$s_\infty(A) := \lim_{m\to\infty} s_m(A) = \sum_{k=0}^{\infty} a_k A^k.$$

In diesem Sinne definieren wir nun die Matrixfunktionen e^A, $\cos(A)$ und $\sin(A)$ durch formales Einsetzen der Matrix A in die Taylor-Reihe der entsprechenden Funktion:

$$e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k, \quad \cos(A) := \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k)!} A^{2k}, \quad \sin(A) := \sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k+1)!} A^{2k+1}.$$

Die Konvergenz der Partialsummenfolgen in $\mathbb{K}^{n\times n}$ ergibt sich dabei nach obiger Argumentation aus der absoluten Konvergenz der jeweiligen Taylor-Reihen mit Konvergenzradius $\rho = \infty$.

Bemerkung 2.10: Für die skalare Exponentialfunktion e^x gilt

$$e^{x+y} = e^x e^y.$$

Beim Beweis dieser Beziehung über die Multiplikationsregel für Potenzreihen wird die Kommutativität der Multiplikation in $\mathbb{K}$ verwendet, d. h.: $xy = yx$. Für die Multiplikation von Matrizen gilt i. Allg. $AB \neq BA$, so dass in diesem Fall auch die obige Funktionalgleichung i. Allg. nicht gilt:

$$e^{A+B} \neq e^A e^B.$$

Wir illustrieren dies durch ein Beispiel:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = BA.$$

$$e^A e^B = \Big(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \Big) \Big(\sum_{k=0}^{\infty} \frac{1}{k!} B^k \Big) = \begin{bmatrix} e & e \\ 0 & 1 \end{bmatrix} \neq \begin{bmatrix} e & e-1 \\ 0 & 1 \end{bmatrix} = \sum_{k=0}^{\infty} \frac{1}{k!} (A+B)^k = e^{A+B}.$$

Es gilt aber stets $e^{A+A} = e^A e^A$ und allgemeiner $e^{A+B} = e^A e^B$. wenn $AB = BA$ (Übungsaufgabe).

Lemma 2.8: *Zwei hermitesche Matrizen $A, B \in \mathbb{K}^{n\times n}$ kommutieren, d. h. erfüllen $AB = BA$, genau dann, wenn sie eine gemeinsame Orthonormalbasis von Eigenvektoren besitzen.*

Beweis: i) Sei $\{z^{(i)},\, i=1,\dots,n\}$ ein gemeinsames Orthonormalsystem von Eigenvektoren von A und B mit zugehörigen Eigenwerten λ_i bzw. μ_i. Dann gilt für jedes $x \in \mathbb{K}^n$:

$$ABx = \sum_{i=1}^{n} (x, z^{(i)})_2 ABz^{(i)} = \sum_{i=1}^{n} \lambda_i \mu_i (x, z^{(i)})_2 z^{(i)} = \sum_{i=1}^{n} (x, z^{(i)})_2 BAz^{(i)} = BAx,$$

d. h.: Es ist $AB = BA$.

ii) Sei nun umgekehrt $AB = BA$. Dann gilt mit den Eigenwerten λ_i und einem zugehörigen Orthonormalsystem $\{z^{(i)},\, i=1,\dots,n\}$ von Eigenvektoren von A:

$$ABz^{(i)} = BAz^{(i)} = \lambda_i Bz^{(i)},$$

d. h. $Bz^{(i)}$ ist Eigenvektor von A zum Eigenwert λ_i. Folglich lässt B den zu λ_i gehörenden Eigenraum $E_A(\lambda_i)$ invariant: $BE_A(\lambda_i) \subset E_A(\lambda_i)$. Seien $\mu_{i,j},\ j=1,\dots,m_i$, die Eigenwerte von $B|_{E_A(\lambda_i)}$ und $\{z^{(i,j)},\ j=1,\dots,m_i\}$, ein zugehöriges Orthonormalsystem in $E_A(\lambda_i)$ von Eigenvektoren. Dann sind konstruktionsgemäß alle diese Eigenvektoren $z^{(i,j)} \in E_A(\lambda_i)$ von B auch Eigenvektoren von A. Die Vereinigung $\cup_i \{z^{(i,j)},\ j=1,\dots,m_i\}$ ist dann ein gemeinsames Orthonormalsystem von Eigenvektoren von B und A. Q.E.D.

Anwendung 2.2.3: Wir stellen uns die Aufgabe, für eine Matrix $A \in \mathbb{K}^{n\times n}$ mit Norm $\rho := \|A\|_2 < 1$ die Inverse $(I-A)^{-1}$ näherungsweise zu berechnen. Nach Lemma 1.16 ist die Matrix $I - A$ regulär. Wir betrachten die Potenzreihenentwicklung

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k,$$

welche für alle $x \in \mathbb{R}$ mit $|x| < 1$ absolut konvergiert. Dann gilt auch

$$(I-A)^{-1} = s_\infty(A) := \sum_{k=0}^{\infty} A^k,$$

wobei die Reihe rechts im oben definierten Sinne in $\mathbb{K}^{n\times n}$ konvergiert. Diese wird „Neumannsche[2] Reihe“ genannt. Zur Approximation von $(I-A)^{-1}$ haben wir also Partialsummen $s_n(A)$ dieser Reihe auszuwerten, was lediglich Matrixmultiplikationen und Additionen erfordert. Die Frage ist nun, wie groß muß n gewählt werden, um eine vorgegebene

[2] John von Neumann (1903–1957): US-amerikanischer Mathematiker ungarischer Abstammung; wirkte hauptsächlich am Institute for Advanced Studies in Princeton (zus. mit A. Einstein u. a.) und gilt als mathematisches Genie; lieferte fundamentale Beiträge zu den mathematischen Grundlagen der Quantenmechanik, zur Operatortheorie, zur Spieltheorie, zur Gruppentheorie und zur Theorie der partiellen Differentialgleichungen; Pionier der Automatentheorie und „Theoretischen Informatik“.

Genauigkeit ε zu erreichen. Dazu betrachten wir den Fehler:

$$\begin{aligned}\Big\|(I-A)^{-1}-\sum_{k=0}^{m}A^k\Big\|_2 &= \Big\|\sum_{k=m+1}^{\infty}A^k\Big\|_2 = \lim_{r\to\infty}\Big\|\sum_{k=m+1}^{r}A^k\Big\|_2\\ &\le \lim_{r\to\infty}\sum_{k=n+1}^{r}\|A\|_2^k = \sum_{k=n+1}^{\infty}\|A\|_2^k\\ &= \|A\|_2^{n+1}\sum_{k=0}^{\infty}\|A\|_2^k = \frac{\rho^{n+1}}{1-\rho}.\end{aligned}$$

Also ist der Fehler in der gewählten Norm im Falle

$$n \ge \frac{\ln(\varepsilon(1-\rho))}{\ln(\rho)} - 1$$

höchstens gleich ε.

2.3 Übungen

Übung 2.1: Für welche $x\in\mathbb{R}^n$ sind die folgenden Funktionen definiert und stetig:

$$a)\quad f(x) := \ln\ln(\|x\|_2), \qquad b)\quad f(x) := \begin{cases}\|x\|_2, & \|x\|_2\le 1\\ 1, & \|x\|_2 > 1\end{cases}.$$

Übung 2.2: Der Produktraum $\mathbb{K}^n\times\mathbb{K}^n$ sei mit der natürlichen Norm $\|\{x,y\}\|_2 := \big(\|x\|_2^2+\|y\|_2^2\big)^{1/2}$ versehen. Man zeige, dass jedes Skalarprodukt $(\cdot,\cdot)$ auf $\mathbb{K}^n$ eine stetige Funktion

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

auf $\mathbb{K}^n\times\mathbb{K}^n$ ist. Ist diese Funktion auch Lipschitz-stetig?

Übung 2.3: Man untersuche, ob die folgenden Mengen im normierten Raum $(\mathbb{R}^2,\|\cdot\|_2)$ zusammenhängend sind:

$$a)\quad M := \partial\{K_1(0)\setminus\{0\}\}, \qquad b)\quad M := \overline{K_1(0)\cap K_1(2)},$$

$$c)\quad M := \cup_{a\in\mathbb{R}^2, a_i\in\mathbb{Z}}\overline{K_{1/2}(a)}, \qquad d)\quad M := \big\{x\in\mathbb{R}^2\,|\,x_1\in\mathbb{R}_+, x_2=\sin(1/x_1)\big\}\cup\{0\}.$$

Übung 2.4: Für eine nichtleere Teilmenge $M\subset\mathbb{R}^n$ sei die „Abstandsfunktion“ $d_M:\mathbb{R}^n\to\mathbb{R}$ definiert durch

$$d_M(x) = \mathrm{dist}(x,M) := \inf_{y\in M}\|x-y\|.$$

a) Man zeige, dass $d_M(\cdot)$ Lipschitz-stetig ist. Wie groß ist die Lipschitz-Konstante?

b) Sei $M \subset \mathbb{R}^n$ ein Untervektorraum und $d_M(\cdot)$ der euklidische Abstand. Man zeige, dass zu jedem $x \in \mathbb{R}^n$ eine eindeutig bestimmte sog. „Bestapproximation“ $x_M \in M$ existiert mit den Eigenschaften

$$d_M(x) = \|x - x_M\|_2, \qquad (x - x_M, y)_2 = 0 \quad \forall y \in M.$$

Übung 2.5: Seien $K_1, K_2 \subset \mathbb{K}^n$ (nichtleere) kompakte Mengen. Man beweise:

a) Die Menge $K_1 \times K_2$ ist kompakt im Produktraum $\mathbb{K}^n \times \mathbb{K}^n$ und die Funktion $f(x,y) := \|x - y\|$ ist stetig auf $K_1 \times K_2$.

b) Es gibt Punkte $a \in K_1$ und $b \in K_2$ mit

$$\|a - b\| = \inf_{x \in K_1, y \in K_2} \|x - y\|,$$

d. h.: $\|a - b\|$ ist der „Abstand“ der Mengen K_1 und K_2.

c) Im Fall $K_1 \cap K_2 = \emptyset$ ist $\inf_{x \in K_1, y \in K_2} \|x - y\| > 0$. Insbesondere für eine einpunktige Menge $K_1 = \{a\} \not\subset K_2$ ist dann $b \in K_2$ eine sog. „Projektion“ des Punktes a auf die Menge K_2. Man mache sich durch eine geometrische Überlegung klar, dass diese Projektion i. Allg. nicht eindeutig bestimmt ist.

d) Zusatzaufgabe für Anspruchsvolle: Welche Zusatzbedingung an die Menge $K_2 \subset \mathbb{K}^n$ würde die Eindeutigkeit der in c) definierten „Projektion“ von $a \notin K_2$ auf K_2 garantieren?

Übung 2.6: a) Sei $(V, \|\cdot\|)$ ein allgemeiner Banach-Raum, d. h. ein *vollständiger* normierter Raum. Man formuliere in diesem Kontext den Banachschen Fixpunktsatz für eine Lipschitz-stetige Abbildung $g : D \subset V \to V$ und die Fehlerabschätzung für die zugehörige Fixpunktiteration. Gilt der Banachsche Fixpunktsatz auch im normierten Funktionenraum $(C[a,b], \|\cdot\|_2)$ mit der L^2-Norm $\|\cdot\|_2$?

b) Ist die durch

$$g(x) := Ax + b, \qquad A := \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & -2/3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

definierte lineare Abbildung eine Kontraktion auf $\mathbb{R}^2$? (Hinweis: Man bestimme die zugehörigen Lipschitz-Konstanten bzgl. geeignet erscheinender Matrixnormen; z. B.: Maximumnorm, l_1-Norm, Spektralnorm, ...)

c) Zusatzaufgabe für Anspruchsvolle: Man gebe eine Version des Banachschen Fixpunktsatzes in einem allgemeinen metrischen Raum $(X, d(\cdot,\cdot))$ an und übertrage den Beweis für den $\mathbb{K}^n$ aus dem Text auf diese Situation.

Übung 2.7: Man beweise die folgende Verallgemeinerung des Banachschen Fixpunktsatzes: Für eine Lipschitz-stetige Selbstabbildung g einer abgeschlossenen Menge $M \subset \mathbb{K}^n$

seien mit L_k die Lipschitz-Konstanten der iterierten Abbildungen g^k bezeichnet. Unter der Bedingung

$$(*) \qquad \sum_{k=1}^{\infty} L_k < \infty$$

besitzt dann g in M genau einen Fixpunkt. Dieser wird als Limes der folgenden sukzessiven Iteration erhalten:

$$x^{(k)} = g(x^{(k-1)}), \quad k \in \mathbb{N}, \quad x^{(0)} \in M.$$

(Die Bedingung $(*)$ ist automatisch erfüllt, wenn g eine Kontraktion ist.)

Übung 2.8: Man zeige, dass jedes Polynom ungeraden Grades auf dem $\mathbb{R}^n$ mindestens eine reelle Nullstelle besitzt.

Übung 2.9: Die Oberfläche der Erdkugel sei als Kugelsphäre $\partial K_R(0) := \{x \in \mathbb{R}^3 : \|x\|_2 = R\}$ angenommen und die (momentane) Oberflächentemperatur $T(x)$ als stetige Funktion

$$T : \partial K_R(0) \to \mathbb{R}.$$

a) Man zeige mit Hilfe des Zwischenwertsatzes, dass es dann zwei „gegenüberliegende“ Punkte $x, -x \in \partial K_R(0)$ gibt mit der Eigenschaft

$$T(x) = T(-x).$$

(Hinweis: Man betrachte auf $\partial K_R(0)$ die Funktion $f(x) := T(x) - T(-x)$.)

b) Zusatzaufgabe für Anspruchsvolle: Wie muss die Aussage von Teil a) modifiziert werden, wenn zur Definition der Kugelsphäre eine andere Norm verwendet wird, z. B. eine gewichtete l_2-Norm, etwa um der tatsächlichen, leicht abgeplatteten Gestalt der Erde gerecht zu werden? Man skizziere die zugehörige Argumentation.

Übung 2.10: Sei $A \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix mit Eigenwerten $\lambda_k \in \mathbb{R}$, ihrer Vielheit entsprechend oft gezählt. Man zeige für rationale Funktionen $r(x) = p(x)/q(x)$ die Abschätzung:

$$\|r(A)\|_2 \le \max_{k=1,\dots,n} |r(\lambda_k)|,$$

vorausgesetzt $q(\lambda_i) \neq 0$, $i = 1,\dots,n$. (Hinweis: Die hermitesche Matrix A besitzt eine Orthonormalbasis von Eigenvektoren.)

Übung 2.11: Sei $A \in \mathbb{K}^{n\times n}$ hermitesch und positiv-definit. Man zeige, dass für die Startmatrix $X^{(0)} = A$ die Folge der Marizen

$$X^{(k)} := \tfrac{1}{2}\big(X^{(k-1)} + (X^{(k-1)})^{-1}A\big), \quad k \in \mathbb{N},$$

gegen die Quadratwurzel $A^{1/2}$ konvergiert.
(Hinweis: Ein analoge Iteration ist zur Berechnung der Quadratwurzel einer Zahl $a \in \mathbb{R}_+$ verwendet worden. Man versuche diesen Beweis für Matrizen zu übertragen. Dazu zeige man, dass alle Iterierten $X^{(k)}$ ein gemeinsames Orthonormalsystem von Eigenvektoren mit A besitzen und folglich kommutieren.)

Übung 2.12: Sei $A \in \mathbb{R}^{n\times n}$ eine reelle symmetrische Matrix.

a) Wie sind dann die Matrizen $e^{iA}, \sin(A), \cos(A)$ definiert? Man gebe Darstellungen mit Hilfe der Eigenwerte und Eigenvektoren von A an.

b) Man zeige für diese Situation die Eulersche Identität

$$e^{iA} = \cos(A) + i\sin(A).$$

c) Bezüglicher welcher Matrixnorm $\|\cdot\|$ auf $\mathbb{R}^2$ gilt dann wie zu erwarten

$$\|\sin(A)\| \le 1, \qquad \|\cos(A)\| \le 1\,?$$

(Hinweis: Man beachte, dass symmetrische Matrizen eine Orthonormalbasis von Eigenvektoren besitzen.)