

Nichtlineare Analysis

Ben Schweizer

Vorlesung an der Universität Basel

Sommersemester 2005

Inhaltsverzeichnis

Prolog: Gewöhnliche Differenzialgleichungen	7
1. Aufgabenstellung bei Differenzialgleichungen	8
2. Phasenraum-Analyse	10
Teil 1. Calculus im Banachraum	15
3. Differenzierbarkeit & Euler-Gleichungen	16
4. Implizite Funktionen & Stetige Abhängigkeit	19
5. Kompakte Operatoren & Schauder'scher Fixpunktsatz	24
6. Fredholmoperatoren & Ljapunov-Schmidt Reduktion	29
Teil 2. Der Abbildungsgrad	35
7. Axiome des Abbildungsgrades	36
8. Konstruktion im Endlichdimensionalen	39
9. Anwendungen des Abbildungsgrades	45
10. Der Abbildungsgrad im Banachraum	51
Teil 3. Verzweigungstheorie	57
11. Lokale Verzweigung	58
12. Stabilität der stationären Zweige	66
13. Globale Verzweigung	72
14. Hopf-Verzweigung	75
Literaturverzeichnis	81

Zur Konzeption der Vorlesung

Inhalte der Vorlesung. In dieser Vorlesung werden wir Methoden kennenlernen, mit denen man nichtlineare Probleme aus der Analysis studieren kann. Die wichtigsten Werkzeuge sind dabei

- Implizite Funktionen
- Fixpunktsätze
- Reduktionsmethoden
- Der Abbildungsgrad
- Verzweigungstheorie

Man kann diese Methoden auch einteilen in *lokale* und *globale* Methoden. Das fundamentale Werkzeug der lokalen Theorie sind die Impliziten Funktionen, sie liefern eine Reduktionsmethode (Ljapunov-Schmidt), und die lokale Verzweigungstheorie basiert größtenteils auf den Impliziten Funktionen. Das fundamentale Werkzeug der globalen Theorie ist der Abbildungsgrad. Mit ihm können Fixpunktsätze und globale Verzweigungsresultate abgeleitet werden. Die Trennung ist allerdings nicht scharf, oft werden verschiedene Methoden kombiniert.

Ein anderer Einteilungsversuch wäre der nach den getroffenen *Aussagen*. Die meisten unserer Aussagen betreffen die Existenz. Für eine nichtlineare Gleichung

$$F(u) = 0$$

mit $F : X \rightarrow Y$, X, Y Banachräume, zeigen wir die Existenz einer Lösung $u \in X$. Dies ist eine typische Aussage in all den obigen Methoden. Für uns liegt dann das Interesse darin, wie wir eine gegebene Anwendung (z.B. eine Differenzialgleichung) in die obige Form bringen, also: Was wählen wir als F , was als u ? Oft noch wichtiger: Welche Räume X, Y sind geeignet? Was für eine Struktur hat F , damit wir etwas aussagen können?

Die Existenzaussage gewinnt an Wichtigkeit, wenn sie auf einer konstruktiven Methode basiert. Dann kann man z.B. numerische Verfahren ableiten. Andere Aussagen betreffen: Struktur der Lösungsmenge, Qualitative Eigenschaften der Lösungen, Stabilität, Regularität oder stetige Abhängigkeit von Daten.

Zum Namen der Vorlesung. Der Titel “Nichtlineare Analysis” ist nicht standard, es ist auch für Experten nicht unbedingt klar, welche Themen in eine solche Vorlesung gehören. Eher standardisiert ist der Name “Nichtlineare Funktionalanalysis”. Diese Vorlesung orientiert sich an einer typischen Stoffauswahl für Nichtlineare Funktionalanalysis, ist aber so konzipiert, dass weder die Vorlesung “Lineare Funktionalanalysis” noch “Partielle Differenzialgleichungen” vorausgesetzt wird. Insbesondere werden als Anwendungen gewöhnliche Differenzialgleichungen betrachtet und keine partiellen Differenzialgleichungen — oder zumindest nur selten ;-)

Beispiele zu den Resultaten.

Der Satz über **Implizite Funktionen** im Banachraum erlaubt folgendes Vorgehen. Wir betrachten die parameterabhängige Differenzialgleichung $\partial_t u = f_\lambda(u)$ auf einem Zeitintervall $[0, T]$ mit festen Startwerten u_0 . Wir können u als Nullstelle einer Gleichung im Banachraum auffassen:

$$X := \{u \in C^1([0, T], \mathbb{R}^n) | u(0) = u_0\},$$

$$F : X \times \mathbb{R} \ni (u, \lambda) \mapsto F(u, \lambda) := \partial_t u - f(u, \lambda) \in Y := C^0([0, T], \mathbb{R}^n).$$

Angenommen, für $\lambda = 0$ haben wir eine Lösung u^0 gegeben, also $F(u^0, 0) = 0$. Dann können wir versuchen, mit dem Satz über Implizite Funktionen aufzulösen. Wir finden dann für kleine $\lambda \in \mathbb{R}$ Nullstellen $u^\lambda \in X$ mit $F(u^\lambda, \lambda) = 0$, also Lösungen der Differenzialgleichung zu $\lambda \neq 0$.

Der **Abbildungsgrad** ordnet einer stetigen Funktion $f : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^n$ und einem Wert $y_0 \in \mathbb{R}^n$ eine ganze Zahl zu,

$$\deg(f, \Omega, y_0) \in \mathbb{Z}.$$

Diese Zahl sieht die *topologischen Eigenschaften* von f , in dem Sinne, dass eine kleine (genauer: eine C^0 -kleine) Sörung von f denselben Abbildungsgrad hat. Weitere Eigenschaften: 1.) der Grad hängt nur von den Werten am Rand ab, also von $f|_{\partial\Omega}$ und 2.) der Grad der Identität ist 1 für $y_0 \in \Omega$. Das *Lösungskriterium* besagt: Falls $\deg(f, \Omega, y_0) \neq 0$, so gibt es einen Punkt $x \in \Omega$ mit $f(x) = y_0$.

Anwendung: Sei $f : \mathbb{R}^n \supset B_1(0) \rightarrow \mathbb{R}^n$ stetig mit $f|_{\partial B_1(0)} = \text{id}|_{\partial B_1(0)}$. Dann hat f eine Nullstelle. Bemerkung: Im Falle $n = 1$ ist dies gerade der Zwischenwertsatz.

Beweis: $\deg(f, B_1(0), 0) = \deg(\text{id}, B_1(0), 0) = 1$ wegen der Unabhängigkeit von den Werten im Inneren und wegen der Normalisierung. Das Lösungskriterium schließlich liefert, dass f eine Nullstelle hat.

Die **Verzweigungstheorie** beschäftigt sich im einfachsten Fall mit dem Problem, die Nullstellenmenge von

$$f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad f(\lambda, x) = x\lambda + g(x, \lambda)$$

zu bestimmen, wo g quadratisch in x ist.

Antwort in diesem Fall: Für $g \equiv 0$ ist die Nullstellenmenge ein Kreuz in \mathbb{R}^2 , d.h. für $\lambda \neq 0$ ist die einzige Nullstelle von $f(., \lambda)$ die Null, für $\lambda = 0$ sind alle $x \in \mathbb{R}$ Nullstellen von $f(., \lambda)$. Der Effekt von g sollte in der Nähe von $(x, \lambda) = (0, 0)$ klein sein. Wir erwarten also (und dies wird in Teil 3 der Vorlesung auch bestätigt), dass die Nullstellen in der Nähe der 0 wieder topologisch die Form eines Kreuzes haben.

Anwendung: Differenzialgleichungen. In dieser Vorlesung betrachten wir Differenzialgleichungen als *die* Anwendung der entwickelten Methoden. Wir messen die Stärke der Theorie daran, wie schnell oder wie elegant sie uns einen Zugang zu Aussagen über Differenzialgleichungen gestattet, und ob sie uns ein tieferes Verständnis von Differenzialgleichungen gewährt.

Prolog: Gewöhnliche Differenzialgleichungen

1. Aufgabenstellung bei Differenzialgleichungen

Gewöhnliche Differenzialgleichungen (DGL) schreiben wir meist als

$$(1.1) \quad \partial_t u = f(u).$$

In physikalischer Interpretation ist für einen Zeitpunkt $t \in \mathbb{R}$ der Vektor $u(t)$ der Zustand des physikalischen Systems zum Zeitpunkt t . Mathematisch ist uns ein Banachraum E gegeben (Zustandsraum), f ist dann eine Abbildung $f : E \rightarrow E$ (eventuell nur definiert auf einer Teilmenge von E). Mit Gleichung (1.1) ist dann folgendes gemeint: Gesucht ist eine Abbildung $u : (T_1, T_2) \rightarrow E$, so dass

$$(1.2) \quad \partial_t u(t) := \frac{\partial}{\partial t} u(t) = \frac{d}{dt} u(t) \stackrel{!}{=} f(u(t))$$

für alle $t \in (T_1, T_2)$. Da man den Zeitparameter auch umbenennen kann ($t' := t - T_1$), reicht es, wenn wir uns ab jetzt immer mit dem Zeitintervall $(0, T)$ beschäftigen.

Leider ist damit noch immer nicht eindeutig definiert, was eine Lösung sein soll, denn es ist keine Klasse von erlaubten Funktionen u vorgegeben, kein *Funktionenraum*. Als *klassische Lösung* bezeichnet man eine stetig differenzierbare Abbildung u , also $u \in C^1((0, T), E)$, für die die Gleichung (1.2) für alle $t \in (0, T)$ gilt. Alternativ könnte man absolutstetige u erlauben und fordern, dass (1.2) für fast alle $t \in (0, T)$ gilt. Oder man erlaubt beliebige $L^1((0, T), E)$ -Funktionen u und fordert, dass (1.1) im Distributionssinn auf $(0, T)$ gilt. Wir werden uns hier mit klassischen Lösungen beschäftigen.

Lösungen der Gleichung (1.1) gibt es meist unendlich viele. Meist ist man an der Lösung interessiert, die zum Zeitpunkt $t = 0$ die richtigen Anfangswerte hat (den richtigen Anfangszustand). Man fordert für gegebenes $u_0 \in E$

$$(1.3) \quad u(0) = u_0.$$

Für klassische Lösungen fordert man zusätzlich zu $u \in C^1((0, T), E)$ noch $u \in C^0([0, T], E)$, also die Stetigkeit bis zu $t = 0$. Dann sagt Gleichung (1.3) wirklich etwas aus (ansonsten könnte man zu einer beliebigen Lösung von (1.1) definieren $u(0) = u_0$, und hätte eine Lösung des *Anfangswertproblems*).

Bei gewöhnlichen Differenzialgleichungen ist man normalerweise an einem endlichdimensionalen Banachraum E interessiert. Nach Wahl einer Basis ist man also im Fall $E = \mathbb{R}^n$. Eine typische Fragestellung ist dann:

Gegeben sei $E = \mathbb{R}^n$, eine Abbildung $f : E \rightarrow E$, ein Anfangswert $u_0 \in E$ und ein Zeitintervall $(0, T)$. Existiert eine klassische Lösung von (1.1), (1.3)?

Die Antwort lautet: Im Allgemeinen, nein. Das heißt, es gibt Abbildungen f und Intervalle $(0, T)$, so dass keine Lösung existiert. Beispiel 1: $E = \mathbb{R}$, $u_0 = 1$, $f(y) = y^2$, $T = 2$. Die eindeutige Lösung auf $(0, 1)$

lautet $u(t) = (1-t)^{-1}$. Diese ist klassisch nicht über $t = 1$ hinaus fortsetzbar.

Man kann etwas bescheidener nach lokaler Existenz von Lösungen fragen:

Gegeben sei $E = \mathbb{R}^n$, eine Abbildung $f : E \rightarrow E$ und ein Anfangswert $u_0 \in E$. Existiert dazu ein Zeitintervall $(0, T) \neq \emptyset$ und eine klassische Lösung von (1.1), (1.3)?

Die Antwort ist wieder: Im Allgemeinen, nein. Beispiel 2: $E = \mathbb{R}$, $u_0 = 0$, $f(y) = -1$ für $y \geq 0$ und $f(y) = 1$ für $y < 0$. Die Lösung kann nicht von der 0 weglauen, denn f treibt u mit Geschwindigkeit 1 zurück zur 0, aber die Funktion $u \equiv 0$ erfüllt $\partial_t u \equiv 0 \neq 1 = f(0)$.

Wir werden noch bescheidener und schränken die Klasse der zulässigen Funktionen f ein. Dann gibt es ein positives Resultat, das hier als bekannt vorausgesetzt wird.

SATZ 1.1 (Existenzsatz von Picard-Lindelöf). *Sei $u_0 \in D \subset E = \mathbb{R}^n$, D offen, und $f : D \rightarrow E$ eine lokal Lipschitz-stetige Funktion. Dann existiert $T > 0$ und eine klassische Lösung $u : (0, T) \rightarrow E$ von*

$$\partial_t u = f(u), \quad u(0) = u_0.$$

Auf $(0, T)$ ist die Lösung eindeutig bestimmt.

BEWEIS. Beweis mit einer Fixpunktiteration und dem Banach'schen Fixpunktsatz. Siehe z.B. Forster, Analysis 2. \square

SATZ 1.2 (Banach'scher Fixpunktsatz). *Sei (X, d) ein metrischer Raum und $g : X \rightarrow X$ kontraktiv, d.h. für ein $\Theta < 1$ gilt*

$$d(g(x), g(y)) \leq \Theta d(x, y) \quad \forall x, y \in X.$$

Dann hat g einen eindeutig bestimmten Fixpunkt.

BEWEIS. Für den Beweis definiert man eine Folge $x_n \in X$ rekursiv durch $x_{n+1} = g(x_n)$ und zeigt, dass dies eine Cauchy-Folge definiert. Im Banachraum existiert dann ein Grenzwert x , welcher $g(x) = x$ erfüllt. \square

In dieser Vorlesung werden wir auch den Existenzsatz von Peano beweisen. Er liefert die Existenz einer Lösung für stetiges f . Allerdings gilt keine Eindeutigkeit, falls f nicht Lipschitz-stetig ist, sh. Übung.

Wir wollen noch darauf hinweisen, dass mit den obigen Konzepten auch Gleichungen zweiter Ordnung behandelt werden können. Die Gleichung

$$(1.4) \quad \partial_t^2 u + g(u, \partial_t u) = 0$$

für $u(t) \in E$, wird mit Hilfe von $X := E \times E$, $v(t) = \partial_t u(t)$, und $U(t) = (u(t), v(t))$ transformiert zu

$$(1.5) \quad \partial_t U(t) = \begin{pmatrix} v(t) \\ -g(u(t), v(t)) \end{pmatrix} =: F(U(t)).$$

Es müssen dann als Startwerte gegeben sein:

$$u(0) = u_0, \quad \partial_t u(0) = u_1.$$

Allerdings hat man oft auch *Randwertprobleme* zu lösen. Dann ist einem nicht Wert und Ableitung an einem Rand gegeben, sondern es sind einem zum Beispiel die Werte an den beiden Rändern gegeben,

$$u(0) = u_0, \quad u(T) = u_T.$$

Beispiel: Die Kettenlinie. Hier ist $t = x \in (0, x_1)$ die horizontale Koordinate, u die vertikale Koordinate. Man gibt die Endpunkte der Kette (im Zweidimensionalen Raum) vor als $(0, u_0)$ und (x_1, u_1) und die Länge als $L > 0$. Die Gleichung lautet dann

$$\begin{aligned} \partial_x^2 u &= k \sqrt{1 + |\partial_x u|^2} \text{ in } (0, x_1), \\ u(0) &= u_0, \quad u(x_1) = u_1. \end{aligned}$$

Die Konstante $k \in \mathbb{R}$ muss über

$$L = \int_0^{x_1} \sqrt{1 + |\partial_x u|^2}$$

bestimmt werden.

Im Fall des Randwertproblems lässt sich die Gleichung nicht mit Picard-Lindelöf lösen. Einen (naiven?) Versuch stellt die shooting-Methode dar: Man löst das Anfangswertproblem zu geratenem u_1 und prüft das Ergebnis $u(T)$. Durch Anpassen von u_1 kann man eventuell einen Treffer erzielen, also erreichen, dass die Lösung tatsächlich $u(T) = u_T$ erfüllt. Dann hat man das Ausgangsproblem gelöst.

2. Phasenraum-Analysis

Vieles lässt sich über die Gleichung $\partial_t u = f(u)$ in $E = \mathbb{R}^n$ sagen, wenn man sich im Phasenraum E die Geschwindigkeitsvektoren f als Pfeile aufzeichnet (oder vorstellt). Für eine Lösung u ist dann jedes $u(t)$ ein Punkt im Phasenraum, der sich in die durch den Pfeil angegebenen Richtung bewegt. Damit lassen sich Lösungskurven bestimmen. Eine Lösungskurve $u : (0, T) \rightarrow E$ im Phasenraum heißt auch Orbit.

Stationäre Punkte. Ein Punkt im Phasenraum $\bar{u} \in E$ mit $f(\bar{u}) = 0$ heißt stationärer Punkt. Tatsächlich ist die triviale Funktion $u(t) = \bar{u} \forall t$ eine Lösung der Gleichung, denn sie erfüllt $\partial_t u \equiv 0 \equiv f(u)$. Die Funktion u heißt auch stationäre Lösung der Gleichung. Manchmal spricht man auch von einem Fixpunkt. Dann denkt man an die Zeit-t-Abbildung $E \ni u_0 \mapsto u(t) \in E$. Ein stationärer Punkt ist eine Punkt im Phasenraum, der Fixpunkt für alle Zeit-t-Abbildungen ist.

Lokale Analysis in der Umgebung eines stationären Punktes. Das Aussehen von f in der Umgebung einer Nullstelle $y_0 = \bar{u}$ ist

deswegen interessant, weil f hier 'sein Vorzeichen ändern kann'. Die Taylor-Reihe für f lautet (wir setzen Differenzierbarkeit voraus)

$$(2.1) \quad \begin{aligned} f(y) &= f(y_0) + Df(y_0) \cdot (y - y_0) + o(|y - y_0|) \\ &= A \cdot (y - y_0) + o(|y - y_0|) \end{aligned}$$

für $y \rightarrow y_0$. Dabei haben wir $A := Df(y_0)$ gesetzt. Ohne Einschränkung sei $y_0 = 0$. Nehmen wir zusätzlich zunächst an, dass f tatsächlich linear ist, also $f(y) = A \cdot y$, $Df(0) = A$. Das Verhalten von Lösungen in der Nähe der Null ist dann durch die Eigenwerte von A bestimmt. Nehmen wir weiterhin an, dass A diagonal ist, also

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Man kann sich die Situation mit Pfeilchenbildern illustrieren, siehe Abbildung 1.

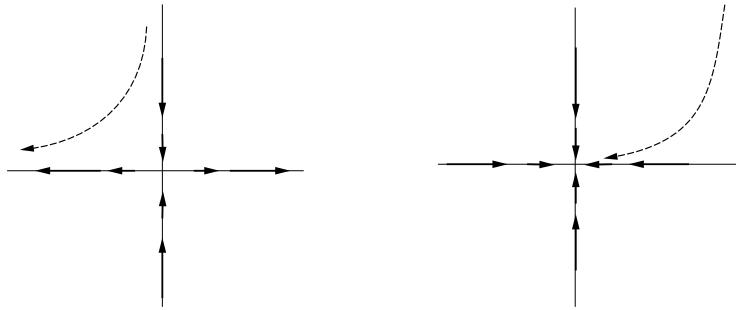


ABBILDUNG 1. 'Pfeilchenbilder' und typische Orbits in der Nähe eines stationären Punktes

Im Falle $\lambda_1, \lambda_2 > 0$ zeigen alle Pfeile von der Null weg, alle Lösungen (außer der Nulllösung) laufen vom Punkt $y_0 = 0$ weg. Wir können die Lösungen angeben. Für $u_0 = (a_1, a_2)$ lautet die Lösung auf $(0, \infty)$

$$(2.2) \quad u(t) = \begin{pmatrix} a_1 e^{\lambda_1 t} \\ a_2 e^{\lambda_2 t} \end{pmatrix}.$$

Im Falle $0 > \lambda_1, \lambda_2$ sind die Lösungen durch dieselbe Formel gegeben, diesmal laufen die Lösungen zum Punkt $y_0 = 0$ hin (rechtes Bild). Man sagt, dass y_0 stabil ist, denn

$$\forall \varepsilon > 0 \exists \delta > 0 : \quad \forall u_0 = (a_1, a_2) \in B_\delta(y_0) :$$

Die zugehörige Lösung u erfüllt $|u(t) - y_0| < \varepsilon \forall t \in (0, \infty)$.

Es gilt sogar mehr: Die stationäre Lösung y_0 ist asymptotisch stabil, denn

$$\exists \delta > 0 : \quad \forall u_0 = (a_1, a_2) \in B_\delta(y_0) :$$

Die zugehörige Lösung u erfüllt $u(t) \rightarrow y_0$ für $t \rightarrow \infty$.

Im Falle $\lambda_1 > 0 > \lambda_2$ (linkes Bild) heißt der Punkt $y_0 = 0$ ein Sattelpunkt. Zum Namen kommen wir später. Zunächst stellen wir fest, dass Pfeile in der Nähe der e_1 -Achse von y_0 wegzeigen, Pfeile in der Nähe der e_2 -Achse eher zu y_0 hinzeigen. Wenn man die Lösungen aus (2.2) betrachtet, so stellt man allerdings fest, dass alle Lösungen letztlich von y_0 wegläufen, außer denen mit $a_1 = 0$, also Lösungen auf der Achse $\mathbb{R}e_2$. In diesem Fall ist $\mathbb{R}e_2$ die stabile Mannigfaltigkeit, denn es gilt

$$(2.3) \quad \mathbb{R}e_2 = \{u_0 \in E \mid u(t) \rightarrow y_0 \text{ für } t \rightarrow \infty, u \text{ die Lösung zu } u_0\}.$$

Man sollte nicht versuchen, die instabile Mannigfaltigkeit als die Startpunkte zu definieren, für die die Lösung wegläuft, denn diese Punkte bilden keine Mannigfaltigkeit. Vielmehr ist hier $\mathbb{R}e_1$ die instabile Mannigfaltigkeit, denn es gilt

$$(2.4) \quad \mathbb{R}e_1 = \{u_0 \in E \mid u(t) \rightarrow y_0 \text{ für } t \rightarrow -\infty, u \text{ die Lösung zu } u_0\}.$$

Der Punkt y_0 heißt instabil, denn

$$(2.5)$$

$$\exists \varepsilon > 0 \forall \delta > 0 \exists u_0 \in B_\delta(y_0), t \in (0, \infty) : |u(t) - y_0| > \varepsilon, u \text{ Lösung zu } u_0.$$

Es genügt, u_0 außerhalb der stabilen Mannigfaltigkeit zu wählen.

Wir wollen nun den Namen 'Sattelpunkt' erklären: Man kann sich eine Berglandschaft über \mathbb{R}^2 vorstellen, zu jedem Punkt $x = (x_1, x_2)$ gehört eine Höhe $h(x_1, x_2) \in \mathbb{R}$. Wenn wir setzen

$$h(x) := -\frac{1}{2}(\lambda_1 x_1^2 + \lambda_2 x_2^2),$$

so gilt für die Anwendung der Ableitung in x auf eine Richtung $y \in \mathbb{R}^2$ die Gleichung

$$Dh(x) \langle y \rangle = -(\lambda_1 x_1 y_1 + \lambda_2 x_2 y_2) = - \left[\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

also

$$\nabla h(x) = - \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Zur Bezeichnung: $Dh(x)$ ist eine Abbildung, die einer Richtung y im Urbildraum einen Richtungsvektor $Dh(x) \langle y \rangle$ im Bildraum zuordnet. Falls (wie hier) h eine skalare Größe ist (nach \mathbb{R}^1 abbildet), so kann man $Dh(x) \langle y \rangle$ mit einer Zahl identifizieren. Falls man zusätzlich im Urbildraum ein Skalarprodukt hat und die Abbildung $Dh(x) : y \mapsto Dh(x) \langle y \rangle$ mit einem Vektor darstellen kann, so nennt man diesen Vektor $\nabla h(x)$, also

$$Dh(x) \langle y \rangle = \langle \nabla h(x), y \rangle = \nabla h(x) \cdot y.$$

Unsere Gleichung lautete also

$$\partial_t u(t) = f(u(t)) = A \cdot u(t) = -\nabla h(u(t)).$$

Eine Gleichung von dieser Form heißt Gradientenfluss. Die Lösung u bewegt sich im Phasenraum immer in die Richtung, in die h am steilsten abfällt. Der Name 'Sattelpunkt' kommt daher, dass im Fall $\lambda_1 > 0 > \lambda_2$ die Höhenfunktion wie ein Sattel aussieht.

Für jeden Gradientenfluß gilt

$$\frac{d}{dt}h(u(t)) = Dh(u(t)) \cdot \partial_t u(t) \equiv \nabla h(u(t)) \cdot \partial_t u(t) = -|\nabla h(u(t))|^2 \leq 0,$$

wobei die letzte Ungleichung strikt ist, außer, die Lösung durchläuft nur kritische Punkte von h . Jede Funktion $H : E \rightarrow \mathbb{R}$, die entlang Lösungen höchstens abnehmen kann ($t \mapsto H(u(t))$ ist monoton nichtwachsend), heißt Ljapunovfunktion.

Ein *nichtdiagonales* A wird mit einem Koordinatenwechsel analysiert. Falls es zu

$$f(y) = A \cdot y, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

eine Basis $s_1, s_2 \in \mathbb{R}^2$ gibt, so dass mit $S = (s_1 \ s_2)$ gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix},$$

so müssen wir nur die Lösung $u(t)$ bezüglich der neuen Basis ausdrücken. Wir schreiben $u(t) = u_1(t)e_1 + u_2(t)e_2 = \tilde{u}_1(t)s_1 + \tilde{u}_2(t)s_2$ und bemerken, dass wir die neuen Koordinaten mit S^{-1} berechnen können, $(\tilde{u}_1, \tilde{u}_2) = S^{-1}(u_1, u_2)$. Wir rechnen

$$\partial_t(\tilde{u}_1, \tilde{u}_2) = S^{-1}\partial_t(u_1, u_2) = S^{-1}A \cdot (u_1, u_2) = S^{-1}AS(\tilde{u}_1, \tilde{u}_2).$$

Die tilde-Koordinatenfunktionen erfüllen also eine Gleichung mit Diagonalmatrix. Diese Methode funktioniert zum Beispiel, wenn die Matrix A symmetrisch ist. Im allgemeinen Fall kommt man nur auf eine Jordan-Matrix, nicht auf eine Diagonalmatrix. Aber auch in diesem Fall kann man explizit lösen, siehe Übungsaufgabe.

Falls die rechte Seite $y \mapsto f(y)$ *nichtlinear* ist, also nicht von der Form $f(y) = A \cdot y$, so betrachtet man zunächst die Linearisierung von f . In der Taylor-Entwicklung von f um einen stationären Punkt $y_0 = 0$ mit $A := Df(y_0)$ gilt

$$f(y) = A \cdot y + g(y).$$

Wegen $g(y) = o(|y|)$ ist das qualitative Aussehen des Pfeilchenbildes für f dasselbe wie für A . Tatsächlich gilt: Falls die Eigenwerte alle einene von 0 verschiedenen Realteil haben, so sieht das Pfeilchenbild für die Linearisierung qualitativ wie das Pfeilchenbild zu A aus.

Eine Aussage, die dies ausnutzt ist die linearisierte Stabilität: Falls alle Eigenwerte von $A = DF(y_0)$ negativen Realteil haben (d.h. y_0 ist stabil für die linearisierte Gleichung), so ist y_0 auch stabil für die nichtlineare Gleichung.

Entsprechend gilt auch: Falls ein Eigenwert von $A = DF(y_0)$ positiven Realteil hat (d.h. y_0 ist instabil für die linearisierte Gleichung), so ist y_0 auch instabil für die nichtlineare Gleichung.

Die genaueste Aussage liefert das Theorem von Grobman-Hartman: Falls alle Eigenwerte Realteil ungleich 0 haben, so besitzt das nichtlineare System stabile und instabile Mannigfaltigkeiten die dieselbe Dimension haben wie die des linearisierten Systems. Im stationären Punkt sind jeweils die Mannigfaltigkeiten des nichtlinearen Systems tangential an den Mannigfaltigkeiten des linearen Systems.

Globale Analysis: Um sich ein globales Bild von den Lösungskurven zu machen, versucht man, zusätzlich zu den stationären Lösungen noch (weitere) Orbits zu finden, die die globale Struktur determinieren. Dies sind periodische Orbits, homokline Orbits und heterokline Orbits. Ein periodische Orbit ist eine Kurve, die durch eine Lösung $u : [0, T] \rightarrow E$ gegeben ist mit $u(T) = u(0)$. Die Lösung kann dann fortgesetzt werden und durchläuft immer wieder den Orbit. Ein heterokliner Orbit ist eine Lösungskurve $u : (-\infty, \infty) \rightarrow E$, die zwei stationäre Punkte verbindet: Für u_- , $u_+ \in E$ mit $f(u_\pm) = 0$ gilt $u(t) \rightarrow u_-$ für $t \rightarrow -\infty$ und $u(t) \rightarrow u_+$ für $t \rightarrow +\infty$. In anderen Worten: Ein heterokliner Orbit ist eine Lösungskurve im Schnitt der instabilen Mannigfaltigkeit von u_- mit der stabilen Mannigfaltigkeit von u_+ . Dabei fordert man beim heteroklinen Orbit zusätzlich $u_- \neq u_+$, während man beim homoklinen Orbit $u_- = u_+$ fordert.

Teil 1

Calculus im Banachraum

3. Differenzierbarkeit & Euler-Gleichungen

In diesem Abschnitt sind X und Y immer Banachräume. Wir wollen uns mit nichtlinearen Abbildungen f von X nach Y beschäftigen. Da f nicht unbedingt auf ganz X definiert sein muss, betrachten wir $f : X \supset G \rightarrow Y$, wobei G eine offene Teilmenge von X ist. Wir werden immer die Stetigkeit von f fordern, in Formeln $f \in C(G, Y)$. Die Definitionen sind wie im Endlichdimensionalen. Man sagt, dass f stetig ist in einem Punkt $x_0 \in G$, falls

$$\forall \delta > 0 \exists \varepsilon > 0 : f(B_\varepsilon(x_0)) \subset B_\delta(f(x_0)).$$

Dabei sind die Kugeln bezüglich der jeweiligen Norm zu sehen; als Topologie in G wird die Topologie von X gewählt. Stetigkeit wird definiert als Stetigkeit in jedem Punkt. Die stetigen und beschränkten Funktionen bilden mit der Supremumsnorm $\|\cdot\|_\infty$ den Banachraum $C_b(G, Y)$.

Die Menge $\mathcal{L}(X, Y)$ sei der Raum der beschränkten linearen Abbildungen $X \rightarrow Y$. Mit der Norm

$$\|T\| := \|T\|_{\mathcal{L}(X, Y)} := \sup\{\|Tx\|_Y : x \in X, \|x\|_X \leq 1\}$$

ist $\mathcal{L}(X, Y)$ wieder ein Banachraum.

3.1. Ableitung und Richtungsableitung. Idee: Die Ableitung einer nichtlinearen Funktion $f : X \rightarrow Y$ in einem Punkt x_0 ist die lineare Abbildung $X \rightarrow Y$, die f bestmöglich approximiert.

Dies ist das Konzept der Fréchet-Ableitung, für uns *die* Ableitung.

DEFINITION 3.1 (Ableitung). f heißt differenzierbar in $x_0 \in X$, falls für ein $A \in \mathcal{L}(X, Y)$ und den 'Fehler'

$$g(z) := f(x_0 + z) - f(x_0) - Az$$

gilt:

$$\frac{\|g(z)\|_Y}{\|z\|_X} \rightarrow 0$$

für $z \rightarrow 0$ in X . Wir schreiben dann $Df(x_0) := A$.

Bemerkungen: Ein solches A ist (falls es existiert) eindeutig. Wichtig an dem Konzept ist, dass $Df(x_0)$ eine Abbildung $X \rightarrow Y$ ist. Für die Auswertung in eine 'Richtung' $v \in X$ schreiben wir meist

$$Df(x_0) : v \mapsto Df(x_0) \langle v \rangle.$$

In der Differentialgeometrie sind die Konzepte in gewisser Weise klarer. Zu Mannigfaltigkeiten M, N mit Punkten $x \in M, y \in N$ und Tangentialräumen $\{T_x M : x \in M\}$ und $\{T_y N : y \in N\}$ gilt: Für $f : M \rightarrow N$ ist $Df(x)$ eine Abbildung $Df(x) : T_x M \rightarrow T_{f(x)} N$. In unserem Fall sind M und N offene Teilmengen von Banachräumen, und damit alle Tangentialräume identisch zum Banachraum. Höhere Ableitungen

sind durch obige Definition abgedeckt: Die (erste) Ableitung ist eine Abbildung $Df : G \rightarrow \mathcal{L}(X, Y)$. Wir können wieder ableiten und finden

$$\begin{aligned} D^2 f : G &\rightarrow \mathcal{L}(X, \mathcal{L}(X, Y)) \equiv \mathcal{L}(X \times X, Y), \\ D^2 f(x) : (v, w) &\mapsto D^2 f(x) \langle v, w \rangle. \end{aligned}$$

Dabei ist natürlich vorausgesetzt, dass wir ein zweites mal differenzieren können.

Falls $Df(x_0) \in \mathcal{L}(X, Y)$ stetig von $x_0 \in X$ abhängt, so schreiben wir $f \in C^1(X, Y)$. Analog wird mit höheren Ableitungen $C^k(X, Y)$ definiert.

Es gilt die **Kettenregel** (Beweis wie in der Infinitesimalrechnung): Falls $f : X \supset G \rightarrow Y$ und $g : Y \supset V \rightarrow Z$ differenzierbar sind und $g \circ f$ in einer Umgebung von x_0 definiert ist, so gilt

$$D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0) : v \mapsto Dg(f(x_0)) \langle Df(x_0) \langle v \rangle \rangle.$$

Kompositionen von C^k -Abbildungen sind wieder C^k .

Ein schwächeres Konzept ist die **Richtungsableitung**. Hier betrachtet man zu gegebenem $f : X \supset G \rightarrow Y$ und $x_0 \in G$ Wege $\gamma : (-\varepsilon, \varepsilon) \rightarrow X$, die durch x_0 laufen, also

$$\gamma \in C^1((- \varepsilon, \varepsilon), X), \quad \gamma : s \mapsto \gamma(s), \quad \gamma(0) = x_0.$$

Die 'Geschwindigkeit' in 0 (also im Bildpunkt x_0) ist $v := \gamma'(0)$. Wir können nun fragen, wie sich die Werte von f verändern, wenn wir den Weg γ entlanglaufen. In der 0 bestimmen wir die Richtungsableitung von f in Richtung v nach der Kettenregel als

$$\frac{d}{ds}(f \circ \gamma)(0) = Df(x_0) \langle v \rangle.$$

Wir sehen, dass eine differenzierbare Funktion Richtungsableitungen in jede Richtung hat; diese werden durch die Auswertung von $Df(x)$ bestimmt. Umgekehrt gibt es allerdings Funktionen, die zwar in jede Richtung Richtungs-differenzierbar sind, aber nicht differenzierbar.

Aus der Kettenregel und dem Hauptsatz der Differenzial und Integralrechnung folgt sofort der **Mittelwertsatz**: Falls $f \in C^1(G)$ auf einer offenen, konvexen Menge G , dann

$$\begin{aligned} f(x_1) - f(x_0) &= \int_0^1 \frac{d}{dt} f(x_0 + t(x_1 - x_0)) \, dt \\ &= \int_0^1 Df(x_0 + t(x_1 - x_0)) \langle x_1 - x_0 \rangle \, dt \\ &= \left[\int_0^1 Df(x_0 + t(x_1 - x_0)) \, dt \right] \langle x_1 - x_0 \rangle. \end{aligned}$$

3.2. Euler-Lagrange Gleichungen. An einem Beispiel wollen wir demonstrieren, wie man Ableitungen im Banachraum für die Herleitung von Differenzialgleichungen einsetzen kann. Wir wollen die isoperimetrische Ungleichung betrachten: Im zweidimensionalen Raum wollen wir mit einer Linie minimaler Länge ein vorgegebenes Volumen einschließen.

Variablen: Die Höhenfunktion sei $u : (0, 1) \rightarrow \mathbb{R}$. Die Höhe in den Randpunkten sei 0, also $u(0) = 0 = u(1)$.

Energien: Die Fläche unter u sei durch F_0 fest gegeben, die Länge des Graphen von u soll minimiert werden. Wir nehmen also im Folgenden an, dass für u gilt

$$\begin{aligned} u \in X &:= \{u \in C^2([0, 1], \mathbb{R}) | u(0) = 0 = u(1)\}, \\ F(u) &:= \int_0^1 u(x) dx = F_0, \\ L(u) &:= \int_0^1 \sqrt{1 + |\partial_x u(x)|^2} dx = \min \{L(v) | v \in X, F(v) = F_0\}. \end{aligned}$$

Analysis: Wir versuchen, u durch eine Gleichung zu beschreiben. Zunächst stellen wir fest, dass $F : X \rightarrow \mathbb{R}$ und $L : X \rightarrow \mathbb{R}$ differenzierbar sind,

$$\begin{aligned} DF(u) \langle v \rangle &= \int_0^1 v(x) dx, \\ DL(u) \langle v \rangle &= \int_0^1 \frac{1}{\sqrt{1 + |\partial_x u(x)|^2}} \partial_x u(x) \partial_x v(x) dx. \end{aligned}$$

Für F folgt dies aus der Linearität von F , die Ableitung (beste lineare Approximation) ist die Funktion selbst. Die Formel für L gewinnt man durch die Kettenregel, denn $L = I \circ g \circ d$ ist die Verkettung vom (linearen) Integral $I : C^1 \rightarrow \mathbb{R}$ mit der Funktion $g : C^1 \ni \xi(\cdot) \mapsto \sqrt{1 + |\xi(\cdot)|^2} \in C^1$ mit der linearen Funktion d 'Differenziation', $d : C^2 \ni u \mapsto \partial_x u \in C^1$.

Nun betrachten wir Wege

$$\gamma \in C^1((-\varepsilon, \varepsilon), X), \quad \gamma : s \mapsto \gamma(s), \quad \gamma(0) = u.$$

Die physikalische Aussage war, dass für jeden Weg γ mit $F(\gamma(\cdot)) \equiv F_0$ gilt, dass $L(\gamma(\cdot))$ in $s = 0$ ein Minimum hat. Zunächst wollen wir überprüfen, dass es zu jeder Richtung v mit

$$\int_0^1 v(x) dx = DF(u) \langle v \rangle = 0,$$

tatsächlich einen Weg γ gibt mit $F(\gamma(\cdot)) \equiv F_0$. Im vorliegenden (linearen) Fall genügt es, den Weg $\gamma(s) := u + sv$ zu betrachten. Entlang des

Weges γ ist L im Nullpunkt minimal, also

$$\begin{aligned} 0 &= \frac{d}{ds} L(\gamma(s))|_0 = DL(u) \langle v \rangle \\ &= \int_0^1 \frac{1}{\sqrt{1 + |\partial_x u(x)|^2}} \partial_x u(x) \partial_x v(x) dx \\ &\stackrel{P.I.}{=} - \int_0^1 \partial_x \left(\frac{\partial_x u}{\sqrt{1 + |\partial_x u|^2}} \right) (x) \cdot v(x) dx \end{aligned}$$

Für jedes $v \in X$ mit Mittelwert 0 verschwindet also das letzte Integral. Dies ist nur möglich, falls der Ausdruck

$$K(x) := \partial_x \left(\frac{\partial_x u}{\sqrt{1 + |\partial_x u|^2}} \right) (x) = K_0$$

unabhängig von x ist. Zum Beweis: Angenommen, K wäre an zwei Stellen x_1 und x_2 unterschiedlich groß. Dann wählen wir v so, dass $v \neq 0$ nur in der Nähe der zwei Punkte, positiv in der Nähe von x_1 und negativ in der Nähe von x_2 mit Mittelwert 0. Das Integral über $K(x) \cdot v(x)$ wäre dann ungleich 0.

Wir haben erhalten, dass die *Krümmung* K entlang der Kurve konstant ist. Die Kurve ist also ein Kreissegment. Die Argumentation bricht zusammen, sobald die Kurve u genau einen Halbkreis beschreibt. Danach ist die Lösungskurve nicht mehr als Graph zu schreiben, die Variablen des Problems sind dann ungeeignet gewählt.

4. Implizite Funktionen & Stetige Abhängigkeit

4.1. Der Satz über Implizite Funktionen. Wir wiederholen den Satz über implizite Funktionen. Er gilt im Banachraum ebenso wie im Endlichdimensionalen. Ziel ist das Auflösen der Gleichung $f(x, y) = 0$ nach x .

SATZ 4.1. Sei

$$f : X \times Y \supset U \rightarrow Z$$

stetig für eine Umgebung U der Lösung (x_0, y_0) von $f(x, y) = 0$. Die Ableitung $D_x f$ in U existiere und sei stetig in U . Falls

$$A := D_x f(x_0, y_0) : X \rightarrow Z \quad \text{Isomorphismus},$$

dann kann man lokal eindeutig auflösen:

- (i) Es gibt $r > 0$, so dass für eine eindeutig bestimmte stetige Funktion $u : Y \supset B_r(y_0) \rightarrow X$ gilt

$$f(u(y), y) = 0 \quad \forall y \in B_r(y_0), \quad u(y_0) = x_0.$$

- (ii) $f \in C^1$ impliziert $u \in C^1(B_r(y_0), X)$ und es gilt dann

$$D_y u(y) = -[D_x f(u(y), y)]^{-1} D_y f(u(y), y).$$

(iii) Für alle $p \geq 1$ gilt: $f \in C^p$ impliziert $u \in C^p$.

BEWEIS. Ohne Einschränkung sei $x_0 = 0$ und $y_0 = 0$. Grundidee des Beweises ist die Äquivalenz

$$\begin{aligned} f(x, y) = 0 &\iff Ax = -f(x, y) + Ax =: R(x, y) \\ &\iff x = A^{-1}R(x, y) =: g(x, y). \end{aligned}$$

Man zeigt die Existenz eines solchen Fixpunktes x (für jedes y) mit Hilfe des Banach'schen Fixpunktsatzes.

Zeige: Für geeignetes r und δ ist

$$g(., y) : B_\delta(0) \rightarrow B_\delta(0) \subset X \text{ eine Kontraktion}$$

für alle $y \in B_r(0) \subset Y$. Der eindeutige Fixpunkt ist $x =: u(y)$.

Sei $\varepsilon > 0$ mit $\varepsilon \|A^{-1}\| \leq 1/2$. Wir rechnen für R :

$$\begin{aligned} R(x_1, y) - R(x_2, y) &= Ax_1 - Ax_2 - (f(x_1, y) - f(x_2, y)) \\ &= A(x_1 - x_2) - \left[\int_0^1 D_x f(tx_1 + (1-t)x_2) dt \right] (x_1 - x_2) \\ &= \left[A - \int_0^1 D_x f(tx_1 + (1-t)x_2) dt \right] (x_1 - x_2) \\ &= \left[\int_0^1 (D_x f(0, 0) - D_x f(tx_1 + (1-t)x_2)) dt \right] (x_1 - x_2). \end{aligned}$$

Wegen der Stetigkeit von $D_x f$ können wir r und δ wählen mit $\|D_x f(0, 0) - D_x f\| \leq \varepsilon$ auf $B_\delta \times B_r \subset X \times Y$. Dann gilt

$$\|R(x_1, y) - R(x_2, y)\| \leq \varepsilon \|x_1 - x_2\| \quad \forall y \in B_r$$

und

$$\|g(x_1, y) - g(x_2, y)\| \leq \frac{1}{2} \|x_1 - x_2\| \quad \forall y \in B_r.$$

Wir müssen noch nachweisen, dass $g(., y)$ eine Selbstabbildung von B_δ ist. Nach evtl. Verkleinern von r gilt $\|g(0, y)\| \leq \frac{1}{2}\delta$ für alle $y \in B_r$.

Wegen $\|g(x, y) - g(0, y)\| \leq \frac{1}{2}\|x - 0\|$ gilt dann auch

$$\|g(x, y)\| \leq \frac{1}{2}\delta + \|g(0, y)\| \leq \delta.$$

Dies liefert die eindeutige Lösung $x =: u(y)$.

Stetigkeit von u : Seien $y_1, y_2 \in B_r$. Wegen der Lösungseigenschaft gilt

$$\begin{aligned} \|u(y_1) - u(y_2)\| &\leq \|g(u(y_1), y_1) - g(u(y_2), y_2)\| \\ &\leq \|g(u(y_1), y_1) - g(u(y_2), y_1)\| \\ &\quad + \|g(u(y_2), y_1) - g(u(y_2), y_2)\| \\ &\leq \frac{1}{2}\|u(y_1) - u(y_2)\| + \|g(u(y_2), y_1) - g(u(y_2), y_2)\|. \end{aligned}$$

Der zweite Term konvergiert gegen 0 für $y_1 \rightarrow y_2$, weil g stetig ist. Damit ist auch u stetig und (i) ist gezeigt.

(ii) *Differenzierbarkeit von u :* Für zwei Punkte y und $y + \delta_y$ in B_r betrachten wir $\delta_u := u(y + \delta_y) - u(y)$. Für

$$B(y) := -[D_x f(u(y), y)]^{-1} D_y f(u(y), y) : Y \rightarrow X$$

müssen wir zeigen

$$\|\delta_u - B(y)\delta_y\| = o(\|\delta_y\|).$$

Die Differenzierbarkeit von f impliziert, dass für beliebiges $\varepsilon > 0$ eine Schranke $\delta_0(\varepsilon) > 0$ existiert, so dass für alle δ_y mit $\|\delta_y\| < \delta_0$ gilt (verwende $\delta_u \rightarrow 0$ für $\delta_y \rightarrow 0$ wegen Stetigkeit von u):

$$\begin{aligned} \|f(u(y + \delta_y), y + \delta_y) - f(u(y), y) - D_x f(u(y), y)\delta_u - D_y f(u(y), y)\delta_y\| \\ \leq \varepsilon(\|\delta_y\| + \|\delta_u\|). \end{aligned}$$

Bemerke, dass die beiden ersten Terme verschwinden, es gilt also

$$\|D_x f(u(y), y)\delta_u + D_y f(u(y), y)\delta_y\| \leq \varepsilon(\|\delta_y\| + \|\delta_u\|).$$

Wegen $D_x f(u(y), y) \rightarrow D_x f(0, 0)$ für $y \rightarrow 0$ (Stetigkeit von $D_x f$ und von u) und da $A = D_x f(0, 0)$ invertierbar, ist auch $D_x f(u(y), y)$ invertierbar mit beschränkter Inversen (für kleines r , $y \in B_r$). Dann gilt

$$\begin{aligned} \|\delta_u - B(y)\delta_y\| &\leq C\varepsilon(\|\delta_y\| + \|\delta_u\|) \\ &\leq C\varepsilon(\|\delta_y\| + \|\delta_u - B(y)\delta_y\| + \|B(y)\delta_y\|). \end{aligned}$$

Für $\varepsilon > 0$ klein genug kann der zweite Term der rechten Seite absorbiert werden. Die Operatoren $B(y)$ sind beschränkt und es folgt

$$\|\delta_u - B(y)\delta_y\| \leq 2C\varepsilon\|\delta_y\|,$$

also die Differenzierbarkeit von u im Punkt y . Da die Formel für die Ableitung stetig von y abhängt, ist u differenzierbar. (iii) folgt aus der Formel für die Ableitung. \square

SATZ 4.2 (Lokale Inverse). Sei $f : X \supset U \rightarrow Y$ eine C^p -Funktion, $p \geq 1$, $x_0 \in U$ und $f(x_0) = y_0$. Falls $D_x f(x_0) : X \rightarrow Y$ ein Isomorphismus, so gilt für ein $r > 0$:

In $B_r(y_0)$ gibt es eine eindeutige 'Inverse', d.h. $u : B_r(y_0) \rightarrow X$ stetig mit $u(y_0) = x_0$ und $f(u(y)) = y \forall y \in B_r(y_0)$. Die Inverse u ist von der Klasse C^p .

BEWEIS. Löse $F(x, y) = f(x) - y = 0 \in Y$ nach $x \in X$ auf. $D_x F(x_0, y_0) = D_x f(x_0)$ und der Satz über implizite Funktionen ist anwendbar. \square

4.2. Anwendung auf DGL: Stetige Abhängigkeit. In diesem Abschnitt untersuchen wir die Abhängigkeit von Lösungen von gewöhnlichen DGL von Ihren Daten. Dabei sehen wir, wie der Satz über implizite Funktionen auf DGLs angewendet werden kann.

Wir betrachten das Anfangswertproblem

$$\left. \begin{array}{l} \frac{d}{dt}u(t) = f(t, u(t), \lambda) \\ u(\sigma) = \xi. \end{array} \right\} (*)$$

für $\xi \in X$, $\sigma \in \mathbb{R}$, $\lambda \in \Lambda$, X und Λ sind Banachräume.

Frage: Wie hängt die Lösung u von den Daten (σ, ξ, λ) ab?

SATZ 4.3. Sei $f : \mathbb{R} \times X \times \Lambda \rightarrow X$ stetig, $D_x f$ existiere und sei stetig in einer Umgebung von $(\sigma_0, \xi_0, \lambda_0)$. Dann existiert $T > 0$ und eine Umgebung U von $(\sigma_0, \xi_0, \lambda_0)$, so dass für alle $(\sigma, \xi, \lambda) \in U$ eine Lösung $u(\cdot; \sigma, \xi, \lambda)$ auf $(\sigma - T, \sigma + T)$ existiert. Falls $f \in C^k$, so gilt

$$u \in C^k((\sigma - T, \sigma + T) \times U, X).$$

BEWEIS. Wir betrachten ein kleines Existenzintervall; für kleines $T > 0$ seien alle t mit $|t - \sigma| \leq T$. Wir normalisieren die Zeit und die Lösung durch die Umparametrisierung

$$t = \sigma + \tau T, \quad u(t) = \xi + \Phi\left(\frac{t - \sigma}{T}\right).$$

Dann lautet die Gleichung für $\Phi(\tau)$

$$\begin{cases} \frac{d}{d\tau} \Phi(\tau) = Tf(\sigma + T\tau, \xi + \Phi(\tau), \lambda) \\ \Phi(0) = 0. \end{cases} \quad (*)'$$

Wir definieren die Räume

$$\begin{aligned} A &:= \{\Phi \in C^1([-1, 1], X) | \Phi(0) = 0\}, \\ Y &:= \mathbb{R} \times \mathbb{R} \times X \times \Lambda, \\ B &:= C^0([-1, 1], X), \end{aligned}$$

und die Abbildung (lokal definiert)

$$\begin{aligned} F : A \times Y &\rightarrow Z \\ (\Phi, (T, \sigma, \xi, \lambda)) &\mapsto \partial_\tau \Phi(\cdot) - Tf(\sigma + T\cdot, \xi + \Phi(\cdot), \lambda). \end{aligned}$$

Wir wollen $F(\Phi, (T, \sigma, \xi, \lambda)) = 0$ nach Φ auflösen. Wir berechnen

$$D_\Phi F(0, (0, \sigma_0, \xi_0, \lambda_0)) : A \rightarrow B, \quad \varphi \mapsto \partial_\tau \varphi,$$

wobei wir benutzt haben, dass $T = 0$ im fraglichen Punkt. Daher ist $D_\Phi F(0, (0, \sigma_0, \xi_0, \lambda_0)) : A \rightarrow B$ ein Isomorphismus. Mit dem Satz über implizite Funktionen können wir nach Φ auflösen. Wir benutzen die Stetigkeit von $D_\Phi F$, die aus der Stetigkeit von $D_x f$ folgt. Die Lösung Φ erbt die Differenzierbarkeitseigenschaften von F , also die von f . Die Lösung $u(t, \sigma, \xi, \lambda)$ hat die Differenzierbarkeitseigenschaften von Φ . \square

Die differenzierbare Abhängigkeit von Parametern ist global interessant. Der obige Satz impliziert auch die folgende Aussage:

Sei $u : [0, T] \rightarrow X$ eine Lösung von $(*)$ zu den Daten (σ, ξ, λ) . Falls f von der Klasse C^k ist, so ist die Abhängigkeit von $u(t)$ von den Daten (σ, ξ, λ) auch C^k .

Zum Beweis überdeckt man den Weg $u([0, T])$ mit endlich vielen Kugeln, in denen der obige Satz gilt. Dann ist die Abbildung $(\sigma, \xi, \lambda) \mapsto u(t)$ eine Verkettung von endlich vielen C^k -Abbildungen und also auch C^k .

4.3. Lagrange-Multiplikatoren. In der Ableitung der Euler Gleichungen für ein Extremalproblem mussten wir einen Weg definieren, der mit vorgegebener Geschwindigkeit v durch den kritischen Punkt u läuft. Falls die Vergleichsmenge ein linearer Raum ist, so kann man den Weg $\varepsilon \mapsto u + \varepsilon v$ wählen. Ist die Vergleichsmenge allerdings eine Mannigfaltigkeit, so muss man geeignete Wege erst finden. Dies geschieht mit dem Satz über implizite Funktionen.

SATZ 4.4. Seien X, Y Hilberträume, $M \subset X$ offen. Das Energiefunktional sei durch $E : M \rightarrow \mathbb{R}$, eine Nebenbedingung durch $\Phi : M \rightarrow Y$ gegeben, beide stetig differenzierbar. Wir betrachten ein Extremum u der Energie E in der Vergleichsmenge $M_0 := \{x \in M | \Phi(x) = 0\}$.

Falls $D\Phi(u) : X \rightarrow Y$ surjektiv ist, so existiert $\lambda \in Y$ mit

$$(4.1) \quad DE(u)v = \langle \lambda, D\Phi(u)v \rangle \quad \forall v \in X.$$

Im Beweis werden wir folgende Aussagen aus der linearen Funktionalanalysis verwenden:

- (1) Riesz-Darstellung. Im Hilbertraum H gilt: Jedes Funktional $f \in H' = \mathcal{L}(H, \mathbb{R})$ lässt sich mit einem $\lambda \in H$ darstellen: $f(h) = \langle \lambda, h \rangle$ für alle $h \in H$.
- (2) Zerlegung. Sei $H_0 \subset H$ abgeschlossener Unterraum in einem Hilbertraum H . Dann gibt es einen abgeschlossenen Unterraum H_1 mit $H = H_0 \oplus H_1$ (sogar orthogonale Zerlegung möglich).
- (3) Satz von der Inversen: Für Banachräume X, Y gilt: Falls $A \in \mathcal{L}(X, Y)$ bijektiv ist, so existiert $A^{-1} \in \mathcal{L}(Y, X)$.

BEWEIS. Wir betrachten den Tangentialraum an M_0 , $X_0 := \ker(D\Phi(u)) \subset X$ und zerlegen X als $X = X_0 \oplus X_1$. Die Abbildung $A := D\Phi(u)|_{X_1} : X_1 \rightarrow Y$ ist bijektiv, es existiert also $A^{-1} \in \mathcal{L}(Y, X_1)$.

Die Abbildung $f := DE(u) \circ A^{-1} : Y \rightarrow \mathbb{R}$ ist stetig und kann dargestellt werden mit $\lambda \in Y$ als $f(y) = \langle \lambda, y \rangle$. Wir verwenden dies für $y = Ax_1$, also gilt

$$DE(u)x_1 = \langle \lambda, Ax_1 \rangle \quad \forall x_1 \in X_1.$$

Damit ist die Aussage für alle 'Normalenrichtungen' gezeigt.

Für alle 'Tangentialrichtungen' $x_0 \in X_0$ gilt nach Definition von X_0 dass $\langle \lambda, D\Phi(u)x_0 \rangle = 0$. Es bleibt zu zeigen, dass $DE(u)x_0 = 0$. Dafür müssen wir einen Weg $\gamma : (-\varepsilon_0, \varepsilon_0) \rightarrow M_0$ finden mit $\gamma(0) = u$ und $\gamma'(u) = x_0$. Sobald dieses γ gefunden ist, gilt wegen Extremalität von E

$$0 = \frac{d}{d\varepsilon} E(\gamma(\varepsilon))|_0 = DE(u)x_0,$$

und also die Behauptung.

Den Weg finden wir mit dem Satz über implizite Funktionen. Die Gleichung

$$\Psi(x_0, x_1) := \Phi(u + x_0 + x_1)$$

lässt sich wegen Invertierbarkeit von $D_{x_1}\Psi(0) = D\Phi(u)|_{X_1} = A$ lokal lösen durch $x_1 = x_1(x_0)$. Wir definieren für festes $x_0 \in X_0$ den Weg als $\gamma(\varepsilon) := (u + \varepsilon x_0 + x_1(\varepsilon x_0))$. Die Formel für die Ableitung liefert $Dx_1(0) = -A^{-1}D\Phi(u)|_{X_0} = 0$, also auch $\gamma'(0) = x_0$. \square

Um die Aussage zu illustrieren, betrachten wir das Standardbeispiel.

BEISPIEL 4.5. Sei $\Omega \subset \mathbb{R}^n$ offen und glatt und der Banachraum sei $X = \{u \in C^2(\bar{\Omega}, \mathbb{R}) | u = 0 \text{ auf } \partial\Omega\}$. Wir betrachten die Energie $E : X \rightarrow \mathbb{R}$ und die Nebenbedingung $\Phi : X \rightarrow Y := \mathbb{R}$,

$$E(u) := \int_{\Omega} \frac{1}{2} |\nabla u|^2, \quad \Phi(u) := -1 + \int_{\Omega} |u|^2.$$

Ein Minimum von E unter der Nebenbedingung $\Phi = 0$ erfüllt notwendigerweise für ein $\lambda \in \mathbb{R}$

$$\begin{aligned} DE(u)v &= \lambda D\Phi(u)v \forall v \in X \\ \Rightarrow \quad \int_{\Omega} (-\Delta u)v &= \int_{\Omega} \nabla u \cdot \nabla v = \lambda \int_{\Omega} uv \forall v \in X \\ \Rightarrow \quad -\Delta u &= \lambda u. \end{aligned}$$

Für eine nichtlineare Euler-Lagrange Gleichung siehe Übung.

5. Kompakte Operatoren & Schauder'scher Fixpunktsatz

Wir betrachten die folgende Aussage.

Stetige Abbildungen von der abgeschlossenen Einheitskugel in sich haben einen Fixpunkt.

Diese Aussage ist im Endlichdimensionalen richtig und im Banachraum falsch.

BEWEIS. a) Im Endlichdimensionalen ist die Aussage gerade der Satz von Brouwer. Man kann den Satz elementar beweisen, siehe z.B. [9]. Wir werden den Satz von Brouwer später mit Hilfe des Abbildungsgrades beweisen.

b) Betrachte $X := l^2 := \{x : \mathbb{N} \rightarrow \mathbb{R} : \|x\|^2 := \sum_{i=0}^{\infty} |x_i|^2 < \infty\}$. Sei $K := \bar{B}_1 \subset X$ und $f : K \rightarrow K$ definiert durch

$$f(x) := (\sqrt{1 - \|x\|^2}, x_0, x_1, \dots).$$

f ist eine stetige Selbstabbildung der Kugel. Wir werden sehen, dass f keinen Fixpunkt hat.

Ann.: Sei $f(x) = x$. Dann $\|x\| = \|f(x)\| = 1$ nach Definition von f . Andererseits gilt dann $x_0 = \sqrt{1 - \|x\|^2} = 0$. Damit folgt $x_1 = x_0 = 0$, $x_2 = x_1 = 0$ usw. Ein Widerspruch. \square

Das Beispiel zeigt, dass man im Banachraum eine zusätzliche Forderung an f stellen muss. Es hat sich gezeigt, dass eine geeignete Forderung die Kompaktheit ist.

5.1. Definition, Endlichdimensionale Approximationen.

DEFINITION 5.1. Sei $A \subset X$. Eine Abbildung $f : A \rightarrow Y$ heißt kompakt, falls sie stetig ist und $\text{cl}(f(B))$ kompakt ist für alle beschränkten Teilmengen $B \subset A$.

Bemerkungen: Da Y vollständig ist, gilt: \bar{A} kompakt $\iff A$ präkompakt, d.h. es gibt endliche Überdeckungen mit ε -Kugeln $\forall \varepsilon > 0$. In der Definition können wir dann fordern: $f(B)$ präkompakt für alle beschränkten $B \subset A$.

Falls A beschränkt ist, reicht es zu fordern: $f(A)$ präkompakt. Begründung: abgeschlossene Teilmengen kompakter Mengen sind kompakt.

BEISPIEL 5.2. Sei $f : X \rightarrow Y$ stetig, $A \subset X$ abgeschlossen und beschränkt.

- X und Y endlichdimensional: $f(A)$ ist abgeschlossen und beschränkt.
- X unendlichdimensional: $f(A)$ muss weder abgeschlossen noch beschränkt sein.

BEWEIS. Im endlichdimensionalen Fall gilt: A kompakt, also $f(A)$ kompakt (offene Überdeckungen des Bildes liefern offene Überdeckungen des Urbildes), also $f(A)$ ist abgeschlossen und beschränkt.

Im unendlichdimensionalen Fall können wir Punkte $x_n \in \partial B_1(0) \subset X$ finden mit $\|x_n - x_m\| \geq 1$ für alle $n \neq m$. Setze

$$\varphi(x) := \begin{cases} k \cdot (1 - 2\|x - x_k\|) & \text{für } x \in B_{1/2}(x_k), \\ 0 & \text{sonst.} \end{cases}$$

Die Abbildung φ ist stetig aber unbeschränkt auf $\bar{B}_2(0)$.

Ersetze in der Definition von φ den Faktor k durch $1 - 1/k$. Dann ist das Bild $\varphi(\bar{B}_2(0)) = [0, 1]$. \square

Das nachfolgende Theorem sagt, dass kompakte Abbildungen genau die sind, die sich gut durch endlichdimensionale Abbildungen approximieren lassen.

SATZ 5.3. Sei $A \subset X$ abgeschlossen und beschränkt. Dann sind für $f : A \rightarrow Y$ äquivalent:

- (1) f ist kompakt.
- (2) $\exists (f_n)_n : A \rightarrow Y$ mit $f_n \rightarrow f$ in $C^0(A, Y)$ und $\text{span}(f_n(A))$ endlichdimensional und $f_n(A)$ beschränkt.

Die Folge f_n kann so gewählt werden, dass sie in die konvexe Hülle von $f(A)$ abbildet.

BEWEIS. '(2) \Rightarrow (1)': Da im Banachraum jeder Punkt eine abzählbare Umgebungsbasis hat, reicht es, zu zeigen, dass $\text{cl}(f(A))$ folgenkompakt ist. Es reicht also, für eine Folge $f(x_k)$ in $f(A)$ zu zeigen: Es gibt eine Teilfolge, so dass $f(x_k) \rightarrow y \in \text{cl}(f(A))$.

Die Folgen $(f_n(x_k))_k$ sind beschränkte Folgen im Endlichdimensionalen und für Teilfolgen von k gilt $f_n(x_k) \rightarrow y_n$. Die Werte $(y_n)_n$ bilden eine Cauchy-Folge wegen

$$\|y_n - y_m\| \leq \|y_n - f_n(x_k)\| + \|f_n(x_k) - f_m(x_k)\| + \|f_m(x_k) - y_m\|;$$

durch Wahl von $k = k(n, m)$ wird Kleinheit erreicht. Daher konvergieren die y_n gegen ein $y \in Y$. Der Ausdruck

$$\|f(x_k) - y\| \leq \|f(x_k) - f_n(x_k)\| + \|f_n(x_k) - y_n\| + \|y_n - y\|$$

wird beliebig klein für geeignet gewähltes n und $k = k(n)$ gross. Insbesondere gilt die Konvergenz und auch $y \in \text{cl}(f(A))$.

'(1) \Rightarrow (2)': Da f kompakt, ist $\text{cl}(f(A))$ kompakt. Für $\varepsilon = \frac{1}{n} > 0$ gibt es daher eine endliche Überdeckung mit ε -Kugeln,

$$\text{cl}(f(A)) \subset \bigcup_{j=1}^p B_\varepsilon(f(x_j)).$$

Wir wählen eine zugehörige Teilung der 1, $(\psi_j)_{j=1,\dots,p}$ mit $\psi_j = 0$ außerhalb $B_\varepsilon(f(x_j))$. Setze

$$f_n(x) := \sum_{j=1}^p \psi_j(f(x)) f(x_j).$$

Dann hat f_n endlichdimensionales Bild, ist stetig, und erfüllt

$$\|f(x) - f_n(x)\| = \left\| \sum_{j=1}^p \psi_j(f(x)) [f(x) - f(x_j)] \right\|.$$

Für j mit $\psi_j(f(x)) \neq 0$ gilt $f(x) \in B_\varepsilon(f(x_j))$ und daher $\|f(x) - f(x_j)\| \leq \varepsilon$, also

$$\|f(x) - f_n(x)\| \leq \varepsilon.$$

□

LEMMA 5.4. *Sei $f : X \rightarrow Y$ von der Klasse C^1 und kompakt in einer Kugel $B_r(x_0)$. Dann ist $Df(x_0) \in \mathcal{L}(X, Y)$ ebenfalls kompakt.*

BEWEIS. Ohne Einschränkung sei $x_0 = 0$ und $f(x_0) = 0$. Wir setzen $A := Df(x_0)$ und betrachten eine Folge $x_k \in B_1(0) \subset X$. Wir wollen zeigen, dass für eine Teilfolge gilt $Ax_k \rightarrow y$ für ein $y \in Y$.

Wegen der Differenzierbarkeit von f gibt es für jedes $n \in \mathbb{N}$ ein $\delta = \delta(n) > 0$ mit

$$\|f(x) - Ax\| \leq \frac{1}{n} \delta \quad \forall x \in B_\delta(0).$$

Für festes n können wir zu einer Teilfolge übergehen, so dass $f(\delta(n)x_k) \rightarrow y_n$ für $k \rightarrow \infty$ und ein y_n (Kompaktheit von f). Dann gilt für $k, l \rightarrow \infty$

$$0 \leftarrow \|f(\delta(n)x_k) - f(\delta(n)x_l)\| \geq \|\delta(n)A(x_k - x_l)\| - \frac{2}{n} \delta(n),$$

also $\|A(x_k - x_l)\| \leq \frac{3}{n}$ für große k, l . Wir gehen nun successive für jedes n zu einer Teilfolge über und finden insgesamt eine Teilfolge, so dass Ax_k eine Cauchy-Folge ist, also einen Limes besitzt. \square

BEISPIEL 5.5. Sei $X = l^2(\mathbb{N}, \mathbb{R})$ und $A : X \rightarrow X$ die lineare Fortsetzung von $e_k \mapsto \frac{1}{k}e_k$ für Basisvektoren e_k . Dann ist A kompakt.

BEWEIS. Mit dem Satz: Die Operatoren

$$A_n : X \rightarrow X, e_k \mapsto \begin{cases} Ae_k & k \leq n, \\ 0 & \text{sonst,} \end{cases}$$

approximieren A gleichmäßig und sind jeweils endlichdimensional. \square

BEISPIEL 5.6. Für das Intervall $I = (a, b) \subset \mathbb{R}$ ist die Einbettung

$$j : C^\alpha(I, \mathbb{R}) \rightarrow C^0(I, \mathbb{R}),$$

$j(u) = u$, kompakt.

BEWEIS. Für $n \in \mathbb{N}$ unterteilen wir I in n Intervalle I_k der Länge $\Delta x := n(b-a)/n$ mit Stützstellen $x_k := a + \frac{k}{n}(b-a)$. Zu einer Funktion $u \in C^\alpha$ können wir die Punktewerte $u_k \in \mathbb{R}$, $k = 0, \dots, n-1$ bestimmen, $u_k = u(x_k)$. Zu den u_k wiederum betrachten wir die lineare Interpolation u^n , definiert durch

$$u^n(x) := \lambda u_k + (1-\lambda)u_{k+1} \text{ für } x = \lambda x_k + (1-\lambda)x_{k+1} \in I_k.$$

Die Abbildung $\Pi_n : u \mapsto u_n$, $C^\alpha \rightarrow C^0$ ist wohldefiniert, endlichdimensional, und stetig. Wir behaupten, dass $\Pi_n \rightarrow j$ gleichmäßig. Dazu rechnen wir mit geeigneter Stützstelle $x_k = x_k(x)$

$$\begin{aligned} \|j(u) - \Pi_n(u)\|_{C^0} &= \sup_x |u(x) - u^n(x)| \\ &\leq \sup_x (|u(x) - u(x_k)| + |u^n(x) - u^n(x_k)|) \\ &\leq \|u\|_{C^\alpha} \left((\Delta x)^\alpha + (\Delta x) \frac{(\Delta x)^\alpha}{(\Delta x)} \right). \end{aligned}$$

Auf einer Kugel in C^α konvergiert dieser Ausdruck gleichmäßig gegen 0 (die Kleinheit des Ausdrückes ist unabhängig von u). \square

5.2. Der Fixpunktsatz von Schauder.

SATZ 5.7 (Brouwer-Schauder). Sei $A \subset X$ abgeschlossen, konvex und beschränkt, X ein Banachraum. Dann hat jede kompakte Abbildung $f : A \rightarrow A$ einen Fixpunkt.

Wir werden dieses Theorem später mit dem Abbildungsgrad zeigen. Es ist allerdings sehr informativ, sich den unendlichdimensionalen Fall aus dem endlichdimensionalen Fall abzuleiten.

BEWEIS. (*Schließe die Banachraumversion aus der endlichdimensionalen Version*) Seien f_n endlichdimensionale Approximationen aus Satz 5.3. f_n bildet in den endlichdimensionalen Teilraum E_n ab, ausserdem in die konvexe Hülle von $f(A) \subset A$, wegen Konvexität von A also auch nach A . Wir betrachten die Einschränkungen der Abbildungen, $f_n|_{A \cap E_n} : A \cap E_n \rightarrow A \cap E_n$. Dann hat f_n einen Fixpunkt x_n , $f_n(x_n) = x_n$. Wegen der Kompaktheit von $\text{cl}(f(A))$ gibt es eine konvergente Teilfolge $x_n \rightarrow x \in \text{cl}(f(A))$. Es gilt

$$\|x_n - f(x_n)\| = \|f_n(x_n) - f(x_n)\| \leq \frac{1}{n}.$$

Wegen $x_n \rightarrow x$ und $f(x_n) \rightarrow f(x)$ folgt $x = f(x)$. \square

5.3. Beispiel: Inverse Differenzialoperatoren.

SATZ 5.8 (Arzela-Ascoli). *Sei $\bar{\Omega} \subset \mathbb{R}^n$ kompakt und $A \subset C^0(\bar{\Omega}, \mathbb{R})$. Falls A eine gleichgradig stetige Familie von Funktionen ist, d.h. für $C > 0$ gilt*

$$\begin{aligned} \|f\|_{C^0} &\leq C \quad \forall f \in A, \\ \sup_{f \in A} |f(x) - f(y)| &\rightarrow 0 \text{ für } |x - y| \rightarrow 0, \end{aligned}$$

so ist der Abschluss \bar{A} kompakt in C^0 .

SATZ 5.9 (Sobolev-Einbettungen). *Sei $\Omega \subset \mathbb{R}^n$ beschränkt und mit Lipschitz-stetigem Rand. Dann gilt für $m, k \in \mathbb{N}$, $p, q \in (1, \infty)$ und $\alpha \in (0, 1)$*

$$(5.1) \quad k - \frac{n}{p} \geq m - \frac{n}{q} \quad \Rightarrow \quad H^{k,p} \subset H^{m,q},$$

$$(5.2) \quad k - \frac{n}{p} \geq m + \alpha \quad \Rightarrow \quad H^{k,p} \subset C^{m+\alpha}.$$

Die Einbettungen sind kompakt, falls 'echt größer' gilt.

Für Beweise siehe z.B. Alt [1], 2.25, 8.7 und 8.8. Siehe auch die Übungen.

Anwendung auf DGL: Ein Existenzsatz. Wir betrachten wieder die gewöhnliche Differenzialgleichung

$$(5.3) \quad u'(t) = f(t, u(t)), \quad u(t=0) = u_0.$$

SATZ 5.10 (Existenzsatz von Peano). *Sei $f : [0, T] \times X \rightarrow X$ stetig. Dann hat für kleine Zeiten (5.3) eine Lösung.*

BEWEIS. Bemerke zunächst: Für $t_0, r > 0$ und $\bar{B}_r(u_0) \subset X$ ist die Funktion f auf $[0, t_0] \times \bar{B}_r(u_0)$ beschränkt durch eine Konstante $M \in \mathbb{R}$.

Eine Lösung erfüllt nach dem Fundamentalsatz

$$u(t) = u_0 + \int_0^t f(\tau, u(\tau)) d\tau.$$

Wir benutzen

$$Z := \{u \in C^0([0, t_0], X) : u(0) = u_0\}$$

und für $r > 0$ auf $A := \bar{B}_r(u_0) \cap Z \subset C^0([0, t_0], X)$ den Operator S durch

$$Su(t) := u_0 + \int_0^t f(\tau, u(\tau)) d\tau.$$

Wegen $\|Su - u_0\| \leq t_0 \cdot M$ gilt für t_0 klein: $S : A \rightarrow A$. Wenn wir nachweisen, dass S kompakt ist, dann können wir den Satz von Schauder anwenden. Der Fixpunkt ist die gesuchte Lösung.

Beh.: S ist kompakt. Wir benutzen Arzela-Ascoli. Die Menge $S(A) \subset C^0$ ist präkompakt, weil die Funktionen der Menge $S(A)$ gleichgradig stetig sind: Sei $z = Su$. Dann gilt

$$\|z(t_1) - z(t_2)\| \leq \int_{t_1}^{t_2} |f(\tau, u(\tau))| d\tau \leq M |t_2 - t_1|.$$

Dies geht gegen 0, und zwar gleichmässig für alle u . Die zweite Bedingung der gleichgradigen Stetigkeit ist auch erfüllt: $\sup_z \sup_t |z(t)| \leq \|u_0\| + Mt_0$. \square

Eine abstrakte Formulierung:

Der Operator S ist eine Verkettung von drei stetigen Operatoren,

$$F : C^0((0, T), X) \rightarrow C^0((0, T), X), u(\cdot) \mapsto f(\cdot, u(\cdot)),$$

$$I : C^0((0, T), X) \rightarrow C^1((0, T), X), f(\cdot) \mapsto \int_0^\cdot f(s) ds,$$

$$j : C^1((0, T), X) \rightarrow C^0((0, T), X), u \mapsto u.$$

Der Operator S war $S = j \circ I \circ F$. Die Operatoren F und I sind beschränkt auf beschränkten Teilmengen und der Operator j ist kompakt (nach Arzela-Ascoli). Dann ist auch die Verkettung kompakt (folgt direkt aus der Definition).

6. Fredholmoperatoren & Ljapunov-Schmidt Reduktion

6.1. Definition, Stabilität. Fredholmoperatoren sind in gewisser Weise das Gegenstück zu kompakten Operatoren. Während die kompakten Operatoren bis auf endlich viele Komponenten nahe der Null sind, sind die Fredholmoperatoren im Wesentlichen nahe der Identität. Wir betrachten Fredholmoperatoren in zweierlei Kontext. Zum einen konstruieren wir den Abbildungsgrad für nichtlineare Operatoren der Form Identität + kompakt — und dies sind im linearen Fall gerade die Fredholmoperatoren. Zum anderen werden wir in der Verzweigungstheorie annehmen, dass die Ableitung einer nichtlinearen Abbildung ein Fredholmoperator ist. Wenn dies der Fall ist, dann gibt es, grob gesprochen, endlich viele Dimensionen, in denen etwas interessantes passieren kann (nicht-injektive und nicht-surjektive Anteile), auf dem Rest ist der Operator invertierbar.

Der Begriff 'Fredholm-Operator' wird im Folgenden nur für lineare Operatoren verwendet. Wir bezeichnen mit $\ker(L)$ den Kern $\{x \in X | Lx = 0\}$, und mit $R(L) = \{Lx | x \in X\}$ das Bild eines Operators $L : X \rightarrow Y$. Wieder ist $\mathcal{L}(X, Y)$ der Raum der stetigen linearen Abbildungen.

DEFINITION 6.1. Ein Operator $L \in \mathcal{L}(X, Y)$ heißt Fredholm-Operator, falls

- (1) $R(L)$ abgeschlossen,
- (2) $\dim \ker(L) < \infty$ und $\text{codim } R(L) < \infty$.

Die zweite Bedingung heisst, dass es eine Zerlegung $Y = Y_0 \oplus R(L)$ gibt mit endlichdimensionalem Y_0 . Die ganze Zahl $\dim \ker(L) - \text{codim } R(L)$ heißt der Index des Fredholmoperators.

Für eine einheitliche Notation: In dieser Vorlesung werden wir meist so zerlegen: $X = X_0 \oplus X_1$, wobei $X_0 = \ker(L)$ endlichdimensional, und $Y = Y_0 \oplus Y_1$, wobei $Y_1 = R(L)$ und Y_0 endlichdimensional.

Um die Definition zu verstehen, erinnern wir (sh. z.B. [1]) an den

Satz vom abgeschlossenen Komplement. Sei X ein Banachraum, $Y \subset X$ ein abgeschlossener Unterraum, und $Z \subset X$ ein Unterraum mit $X = Y \oplus Z$. Dann sind äquivalent:

- (a) Es gibt eine stetige Projektion $P \in \mathcal{L}(X)$ mit $R(P) = Y$ und $\ker(P) = Z$.
- (b) Z ist abgeschlossen.

Bemerkungen zu Definition 6.1:

(i) Sei X Banachraum und X_0 ein endlichdimensionaler Teilraum. Dann gibt es eine stetige Projektion auf X_0 . Insbesondere hat X_0 ein abgeschlossenes Komplement.

Beweis von (i): Sei $X_0 = \text{span}(e_1, \dots, e_n)$ und $\lambda_k : X_0 \rightarrow \mathbb{R}$ die linearen Funktionale mit $\lambda_k(e_l) = \delta_{kl}$. Setze mit Hahn-Banach λ_k fort zu stetigen linearen Funktionalen auf X . Dann ist durch

$$Px := \sum_{k=1}^n \lambda_k(x) e_k$$

eine stetige lineare Projektion von X auf X_0 definiert.

Insbesondere können wir in der Zerlegung der Räume immer X_1 abgeschlossen wählen.

(ii) Für einen Fredholmoperator L sei $X = X_0 \oplus X_1$ eine Zerlegung von X mit $X_0 = \ker(L)$ und X_1 abgeschlossen. Dann ist der Operator $L_1 := L|_{X_1} : X_1 \rightarrow R(L)$ stetig invertierbar. Dies folgt aus dem Satz über die inverse Abbildung und der Abgeschlossenheit der Räume.

(iii) Mit Hilfe von Bemerkung (i) kann man zeigen: Eigenschaft (1) des Fredholmoperators folgt aus Eigenschaft (2).

Beweis von (iii): Sei $X = X_0 \oplus X_1$ mit $X_0 = \ker L$ und X_1 abgeschlossen. Dann ist $L_1 : X_1 \rightarrow R(L)$ stetig invertierbar nach (ii). Der

erweiterte Operator $\tilde{L} : X_1 \times Y_0 \rightarrow Y$, $(x_1, y_0) \mapsto Lx_1 + y_0$ ist wieder bijektiv, linear und stetig zwischen Banachräumen, insbesondere hat \tilde{L} eine stetige Inverse $B := \tilde{L}^{-1}$. Also ist

$$R(L) = L(X_1) = \tilde{L}(X_1 \times \{0\}) = B^{-1}(X_1 \times \{0\})$$

abgeschlossen.

Der Zusammenhang zu kompakten Operatoren ist wie folgt.

SATZ 6.2. *Falls $K \in \mathcal{L}(X, X)$ kompakt, so ist $L := id + K$ Fredholmoperator mit Index 0.*

BEWEIS. Siehe Alt [1], Nr. 8.15. □

Es gilt sogar auch die Umkehrung:

LEMMA 6.3. *Sei $L : X \rightarrow Y$ Fredholmoperator mit Index 0. Mit einem Banachraumisomorphismus $I : Y \rightarrow X$ ist dann $I \circ L$ von der Form $I \circ L := id + K$ mit $K \in \mathcal{L}(X)$ kompakt.*

BEWEIS. Wir verwenden wieder $\tilde{L} : Y_0 \times X_1 \rightarrow Y$, $(y_0, x_1) \mapsto Lx_1 + y_0$ mit der stetigen Inversen $B := \tilde{L}^{-1}$. Der Operator $B \circ L : X \rightarrow Y_0 \times X_1$ ist die Identität auf dem Anteil X_1 , also überall bis auf einen endlichdimensionalen Teilraum. Die endlichdimensionalen Teilräume X_0 und Y_0 haben dieselbe Dimension und sind daher isomorph. □

Nach dieser Vorbereitung ist das nachfolgende Lemma keine Überraschung.

LEMMA 6.4 (Fredholm + kompakt ist Fredholm). *Sei $L : X \rightarrow Y$ ein Fredholm-Operator und*

$$K : X \rightarrow Y \text{ kompakt.}$$

Beide linear und stetig. Dann ist der Operator

$$L + K : X \rightarrow Y$$

ein Fredholm-Operator mit demselben Index wie L .

BEWEIS. Durch 'triviales' Hinzufügen einer endlichdimensionalen Komponente zu einem der Banachräume X oder Y erreichen wir, dass L den Index 0 hat. Wir müssen zeigen, dass $L + K$ wieder den Index 0 hat. Durch einen Isomorphismus $I : Y \rightarrow X$ erreichen wir nach Lemma 6.3 $I \circ L = id + K_1$ mit $K_1 : X \rightarrow X$ kompakt. Der Operator $I \circ K$ ist als Verkettung stetig mit kompakt wieder kompakt, nach Satz 6.2 gilt also, dass $B \circ (L + K) = id + K_1 + I \circ K$ wieder Fredholm mit Index 0 ist. Der Isomorphismus I ändert diese Eigenschaften nicht, wir erhalten also das Ergebnis für $L + K$. □

Für spätere Anwendungen notieren wir noch das nachfolgende wichtige Corollar zu Satz 6.2.

COROLLAR 6.5 (Fredholm-Alternative). *Sei $\lambda \neq 0$ und $K \in \mathcal{L}(X, X)$ kompakt. Dann gilt:*

Entweder hat $(\lambda - K)x = 0$ eine nichttriviale Lösung oder $(\lambda - K)$ ist invertierbar.

BEWEIS. Betrachte $L = \lambda - K$. Nach Theorem 6.2 ist L ein Fredholm-Operator. Falls $\ker(L) = \{0\}$, so gilt $R(L) = X$ nach der Definition des Index von Fredholm-Operatoren. \square

6.2. Ljapunov-Schmidt Reduktion. Unser Ziel ist die Reduktion eines unendlichdimensionalen Gleichungssystems auf ein endlichdimensionales. Wir wollen auf die 'wesentlichen Koordinaten projizieren'.

Wir betrachten die Gleichung

$$M(x, \lambda) = 0 \quad \text{für } M : X \times \Lambda \rightarrow Y.$$

X , Λ und Y seien Banachräume. Im Folgenden seien M und $D_x M$ immer stetig. Eine triviale Lösung sei gegeben, ohne Einschränkung gelte $M(0, 0) = 0$.

Wäre $D_x M$ invertierbar, so könnten wir mit dem Satz über implizite Funktionen die Gleichung $M(x, \lambda) = 0$ nach x mit $x = x(\lambda)$ auflösen. Häufig hat man jedoch keine Invertierbarkeit. Dann kann man immer noch hoffen, dass man bis auf wenige Koordinaten auflösen kann.

Voraussetzung: Der Operator $L := D_x M(0, 0) : X \rightarrow Y$ ist ein Fredholm Operator.

Wie zuvor können wir dann die Räume zerlegen in $X = X_0 \oplus X_1 = \ker L \oplus X_1$ mit Projektion $P : X \rightarrow X$ auf X_0 mit $\ker P = X_1$. Ebenso $Y = Y_0 \oplus Y_1 = Y_0 \oplus R(L)$ mit Projektion $Q : Y \rightarrow Y$ auf Y_0 mit $\ker Q = Y_1$.

Die Gleichung $M(x, \lambda) = 0$ ist äquivalent zu

$$\begin{aligned} QM(x, \lambda) &= 0, \\ (1 - Q)M(x, \lambda) &= 0. \end{aligned}$$

Wir schreiben diese Gleichungen mit Hilfe der Zerlegung des Raumes X als

$$\begin{aligned} QM(x_0 + x_1, \lambda) &= 0, \\ (1 - Q)M(x_0 + x_1, \lambda) &= 0. \end{aligned}$$

In dieser Formulierung können wir die zweite Gleichung nach (den unendlich vielen Koordinaten) x_1 auflösen, denn die Ableitung

$$D_{x_1}[(1 - Q)M(x_0 + x_1, \lambda)]|_{(0,0)} = (1 - Q)L|_{X_1} : X_1 \rightarrow Y_1$$

ist invertierbar. Wir finden eine Lösungsfunktion $x_1 = v(x_0, \lambda)$. Damit haben wir das System reduziert auf die erste (endlichdimensionale) Y_0 -Gleichung:

$$\Phi(x_0, \lambda) := QM(x_0 + v(x_0, \lambda), \lambda) = 0.$$

Diese Gleichung heißt die reduzierte Gleichung oder die Verzweigungs-gleichung.

Wir fassen die Ergebnisse zusammen.

SATZ 6.6. *Sei $M : X \times \Lambda \rightarrow Y$ mit trivialer Lösung $M(0, 0) = 0$, die Ableitung $D_x M$ sei stetig in einer Umgebung der 0 und $L := D_x M(0, 0)$ sei Fredholm Operator und die Zerlegungen*

$$\begin{aligned} X &= X_0 \oplus X_1 = \ker L \oplus X_1, \quad P : X \rightarrow X \text{ Proj. auf } X_0, \quad \ker P = X_1, \\ Y &= Y_0 \oplus Y_1 = Y_0 \oplus R(L), \quad Q : Y \rightarrow Y \text{ Proj. auf } Y_0, \quad \ker Q = Y_1. \end{aligned}$$

Dann gilt für Umgebungen der 0, $U \subset X_0 = \ker(L)$, $V \subset X_1$, $W \subset \Lambda$: Es gibt eine eindeutig bestimmte Funktion $v : U \times W \rightarrow V$, so dass lokal die Gleichung $(1 - Q)M(x_0 + x_1, \lambda) = 0$ eindeutig durch $x_1 = v(x_0, \lambda)$ gelöst wird, v ist stetig.

Die Gleichung $M(x_0 + x_1, \lambda) = 0$ ist lokal äquivalent zur reduzierten Gleichung für $x_0 \in X_0$,

$$\Phi(x_0, \lambda) := QM(x_0 + v(x_0, \lambda), \lambda) = 0 \in Y_0.$$

Falls $M \in C^1$, so sind v und Φ in C^1 und es gilt

$$D_{x_0}v(0, 0) = 0, \quad D_{x_0}\Phi(0) = 0.$$

Für die Ableitungsformeln rechnen wir im Nullpunkt

$$D_{x_0}v = -[D_{x_1}((1 - Q)M)]^{-1} \circ D_{x_0}((1 - Q)M) = 0,$$

da $D_{x_0}((1 - Q)M) = (1 - Q)L|_{X_0} = 0$. Ähnlich folgt aus $\ker Q = R(L)$

$$D_{x_0}\Phi(0) = Q \circ L \circ (\text{id} + D_{x_0}v) = 0.$$

Teil 2

Der Abbildungsgrad

7. Axiome des Abbildungsgrades

Der Abbildungsgrad soll zu einer Funktion $f : G \rightarrow Y$ und einem Bildpunkt $y_0 \in Y$ eine Zahl $d(f, G, y_0) \in \mathbb{Z}$ liefern. Diese Zahl soll etwas über die Existenz von Lösungen $x \in G$ der Gleichung $f(x) = y_0$ aussagen.

In einer Dimension: Der Zwischenwertsatz

Betrachte stetige $f : [a, b] \rightarrow \mathbb{R}$. Falls $\operatorname{sgn}(f(a)) \neq \operatorname{sgn}(f(b))$ gilt, dann hat die Gleichung

$$f(x) = 0$$

eine Lösung $x \in (a, b)$.

Man könnte so formulieren: Bilde die Zahl

$$d := \frac{1}{2}\operatorname{sgn}(f(b)) - \frac{1}{2}\operatorname{sgn}(f(a)).$$

Wenn $d \neq 0$, dann hat die Gleichung $f(x) = 0$ eine Lösung.

Obiges d ist genau der Abbildungsgrad im Eindimensionalen.

- (1) d hängt nur von den Werten von f am Rand $\partial(a, b)$ ab
- (2) d ist ganzzahlig
- (3) Lösungskriterium: Falls $d \neq 0$ hat *jedes* f mit diesen Randwerten eine Lösung
- (4) d zählt die Nullstellen mit Vorzeichen

Wichtiger Punkt: f mit $d = 0$ kann sehr wohl Nullstellen haben.

In zwei Dimensionen: Die Windungszahl

Zu einem geschlossenen Weg $g : [0, 2\pi] \rightarrow \mathbb{R}^2 \setminus \{0\}$ können wir die *Windungszahl* wie folgt definieren: Wir schreiben $g(t) = |g(t)|(\cos(\Theta(t)), \sin(\Theta(t)))$. Damit ist $\Theta(t)$ bis auf Vielfache von 2π bestimmt. Fordern wir Stetigkeit von Θ , so ist Θ auf $[0, 2\pi]$ bis auf Vielfache von 2π bestimmt. Wir definieren die *Windungszahl* als

$$w(g, 0) := \frac{\Theta(2\pi) - \Theta(0)}{2\pi} \in \mathbb{Z}.$$

Mit Hilfe der komplexen Schreibweise finden wir eine einfache Formel: Wir setzen $S := S^1 = \partial B_1 \subset \mathbb{R}^2 = \mathbb{C}$. Wegen $d \log(z) = \frac{dz}{z}$ gilt

$$w(g, 0) = \frac{1}{2\pi i} \int_{g(S)} \frac{dz}{z}.$$

Für jede holomorphe Fortsetzung f der Randwerte g gilt wegen Transformationsformel und Residuensatz

$$w(f|_S, 0) = \frac{1}{2\pi i} \int_S \frac{f'}{f} = \sum_z v_f(z).$$

Dabei geht die Summe über alle z mit $f(z) = 0$ und $v_f(z)$ ist die Vielfachheit der Nullstelle. Insbesondere gilt: Für $w(f|_S, 0) \neq 0$ hat jede Fortsetzung der Randwerte eine Nullstelle.

Die Windungszahl hat die obigen Eigenschaften (1)-(4). Im Zweidimensionalen haben wir mit der Windungszahl genau den Abbildungsgrad gefunden:

$$d(f, B_1, 0) = w(f|_S, 0).$$

In Dimension n : Es gibt topologische Kriterien

DEFINITION 7.1. Seien M, N topologische Räume und $f, g \in C^0(M, N)$. Dann ist f homotop zu g , falls es ein $h \in C^0(M \times [0, 1], N)$ gibt mit

$$h(., 0) = f, \quad h(., 1) = g.$$

Wir schreiben $f \sim g$, h heisst eine Homotopie. Wir nennen f homotop zur 0, falls es $y_0 \in N$ gibt mit $f \sim y_0$ ($g(x) = y_0 \forall x$).

DEFINITION 7.2. $\varphi : S^{n-1} \rightarrow \mathbb{R}^m \setminus \{0\}$ heisst wesentlich wenn gilt: Jede stetige Fortsetzung $f : B^n \rightarrow \mathbb{R}^m$ von φ hat eine Nullstelle.

SATZ 7.3. $\varphi : S^{n-1} \rightarrow \mathbb{R}^m \setminus \{0\}$, setze $\psi : S^{n-1} \rightarrow S^{m-1}$, $\psi := \frac{\varphi}{|\varphi|}$. Dann

$$\varphi \text{ wesentlich} \iff \psi \text{ nicht homotop zu } 0.$$

Der Satz sagt: Die Randwerte können topologische Eigenschaften haben, so dass es unmöglich ist, ohne eine Nullstelle fortzusetzen.

BEWEIS. ' \Rightarrow' Sei $\psi \sim 0$. Zeige: φ hat eine Fortsetzung ohne Nullstelle.

Sei $h \in C^0(S^{n-1} \times [0, 1], S^{m-1})$ eine Homotopie von ψ nach y_0 . Setze

$$f(x) := \begin{cases} \left(1 - |x|(1 - |\varphi(\frac{x}{|x|})|)\right) \cdot h(\frac{x}{|x|}, 1 - |x|) & \text{für } x \neq 0, \\ y_0 & \text{für } x = 0. \end{cases}$$

f erfüllt $f = \varphi$ auf S^{n-1} , f ist stetig, und $f(x) \neq 0 \forall x$. Letzteres wegen:

$$f(x) = 0 \iff 1 - |x|(1 - |\varphi(\frac{x}{|x|})|) = 0,$$

und es gilt immer

$$|x|(1 - |\varphi(\frac{x}{|x|})|) < |x| \leq 1.$$

' \Leftarrow ': Ann: φ nicht wesentlich. Sei f eine stetige Fortsetzung ohne Nullstelle. Definiere

$$h(x, t) := \frac{f(tx)}{|f(tx)|}.$$

Dann ist h ist eine Homotopie von 0 nach ψ . \square

Die Axiome des Abbildungsgrades.

- (d1) (Ganzzahligkeit) X sei ein Banachraum, $G \subset X$ offen, beschränkt, $f : \bar{G} \rightarrow X$ stetig, $y \notin f(\partial G)$. Dann ist der Abbildungsgrad eine ganze Zahl

$$d(f, G, y_0) \in \mathbb{Z}.$$

Im Unendlichdimensionalen: Fordere $f = \text{id} + g$ mit g kompakt.

- (d2) (Lösungskriterium)

$$d(f, G, y_0) \neq 0 \Rightarrow \exists x \in G : f(x) = y_0.$$

- (d3) (Normierung)

$$d(\text{id}, G, y_0) = \begin{cases} 1 & y_0 \in G \\ 0 & y_0 \notin \bar{G}. \end{cases}$$

- (d4) (Homotopieinvarianz) $h : \bar{G} \times [0, 1] \rightarrow X$ stetig, $y : [0, 1] \rightarrow X$ stetig, $y(t) \notin h(\partial G, t) \forall t$. Dann gilt

$$t \mapsto d(h(., t), G, y(t)) \quad \text{konstant.}$$

Im Fall $\dim(X) = \infty$ fordern wir, dass $h(x, t) = x + g(x, t)$ mit g kompakt auf $G \times [0, 1]$.

- (d5) (Ausschneiden) $G_1, G_2 \subset G$ offen, disjunkt, $f : \bar{G} \rightarrow X$ stetig, $y_0 \notin f(\bar{G} \setminus (\partial G_1 \cup \partial G_2))$. Dann gilt

$$d(f, G, y_0) = d(f, G_1, y_0) + d(f, G_2, y_0).$$

SATZ 7.4. (d1), (d3)-(d5) bestimmen den Grad eindeutig.

Wir werden dieses Theorem im Folgenden nicht verwenden. Zum Beweis verweisen wir auf Deimling [4], Chapter 1, §1 (und die Übungen).

Als eine typische Anwendung des Abbildungsgrades zeigen wir schon hier den Satz von Schauder:

SATZ 7.5. Sei X ein Banachraum, $g : \bar{B} \rightarrow \bar{B}$ stetig und kompakt. Dann hat g einen Fixpunkt.

BEWEIS. Setze $f := \text{id} - g$, $y_0 := 0$, $G := \bar{B}$. Wir suchen Nullstellen von f . Daher können wir o.E. annehmen, dass f keine Nullstelle auf ∂B hat. Wegen (d2) reicht es zu zeigen:

$$d(f, G, 0) \neq 0.$$

Wir benutzen die Homotopie $h(x, t) = x - t \cdot g(x)$ zwischen f ($t = 1$) und id ($t = 0$). Falls f keine Nullstelle hat ist die Homotopie zulässig, da

$$|h(x, t)| \geq 1 - t|g(x)| > 0 \quad \text{für } t < 1$$

und

$$|h(x, t)| = |f(x)| > 0 \quad \text{für } t = 1.$$

Wegen (d4) und (d3) gilt $d(f, G, 0) = d(\text{id}, G, 0) = 1$. \square

Im Folgenden benötigen wir das Lemma von Sard. Wir geben den Beweis für eine sehr einfache Version, die allerdings für unsere Zwecke ausreicht.

DEFINITION 7.6. *Wir betrachten $X \subset \mathbb{R}^n$ und $Y = \mathbb{R}^k$ und Funktionen $F \in C^1(X, Y)$.*

- (1) *Ein Punkt $x_0 \in X$ heisst regulärer Punkt von F , wenn $DF(x_0)$ maximalen Rang hat (d.h. $\min\{n, k\}$).*
- (2) *Ein Punkt, der nicht regulär ist, heisst kritischer Punkt.*
- (3) *Ein Punkt y_0 in Y heisst kritischer Wert, falls $F^{-1}(y_0)$ einen kritischen Punkt enthält.*

SATZ 7.7 (Lemma von Sard, einfache Version). *Sei $G = [0, 1]^n \subset \mathbb{R}^n$ und $F \in C^1(G, \mathbb{R}^n)$ mit beschränkter Ableitung. Dann hat die Menge der kritischen Werte von F das (Lebesgue-) Maß 0 in \mathbb{R}^n .*

BEWEIS. Durch Aufteilen jeder Seite in N gleich lange Abschnitte wird der Würfel in N^n Würfel der Kantenlänge N^{-1} geteilt. Für zwei Punkte x, x_0 in demselben Teilwürfel gilt

$$\begin{aligned} F(x) &= F(x_0) + \int_0^1 DF(x_0 + t(x - x_0)) \cdot (x - x_0) \, dt \\ &= F(x_0) + DF(x_0) \cdot (x - x_0) + o\left(\frac{1}{N}\right). \end{aligned}$$

Dabei gilt Gleichmäßigkeit: $No(\frac{1}{N}) \rightarrow 0$ für $N \rightarrow \infty$ gleichmäßig in x_0 .

Angenommen, in einem Teilwürfel W liegt ein kritischer Punkt x_0 von F . Wegen $\det DF(x_0) = 0$ sind die Werte $F_{app}(x) = F(x_0) + D_x F(x_0) \cdot (x - x_0)$ in einer $n - 1$ -dimensionalen Hyperebene des \mathbb{R}^n enthalten. Für eine Konstante C liegen alle Bildpunkte $F(x)$ des Teilwürfels in einem Quader mit Volumen $C(\frac{1}{N})^{n-1} \cdot o(\frac{1}{N})$.

Wir berechnen das Bildvolumen aller Würfel W , die kritische Punkte enthalten. Es gibt höchstens N^n Stück es gilt

$$\bigcup_W |F(W)| \leq CN^n \left(\frac{1}{N}\right)^{n-1} \cdot o\left(\frac{1}{N}\right) = CNo\left(\frac{1}{N}\right) \rightarrow 0$$

für $N \rightarrow \infty$. Da N beliebig ist, sind die kritischen Werte in einer Nullmenge enthalten. \square

8. Konstruktion im Endlichdimensionalen

8.1. Definition des Grades für C^2 -Abbildungen. In diesem Abschnitt definieren wir den Abbildungsgrad explizit mit Hilfe eines Integrals. Im Folgenden ist $G \subset \mathbb{R}^n$ immer offen und beschränkt.

Wir führen den Raum Ω_ε der Glättungsfunktionen mit kleinem Träger ein:

$$\Omega_\varepsilon := \left\{ \omega \in C^1(\mathbb{R}^n, \mathbb{R}) : \int_{\mathbb{R}^n} \omega = 1, \text{ supp}(\omega) \subset B_\varepsilon^n(0) \right\}.$$

Für die folgende Definition bemerken wir, dass für $f \in C^0(\bar{G}, \mathbb{R}^n)$ und $y_0 \notin f(\partial G)$ mit

$$\varepsilon := \frac{1}{2} \text{dist}(f(\partial G), y_0) > 0$$

eine positive Zahl definiert ist.

DEFINITION 8.1. Zu $f \in C^0(\bar{G}, \mathbb{R}^n) \cap C^2(G, \mathbb{R}^n)$ und $y_0 \notin f(\partial G)$ wählen wir obiges ε und eine Glättungsfunktion $\omega \in \Omega_\varepsilon$. Mit der Funktionaldeterminante $\mathcal{J}_f : G \rightarrow \mathbb{R}$ von f definieren wir

$$(8.1) \quad d(f, G, y_0) := \int_G \omega(f(x) - y_0) \mathcal{J}_f(x) dx.$$

Es bleibt zu zeigen, dass diese Zahl unabhängig von der Wahl von ω ist. Dies geschieht in Lemma 8.3. Weiterhin sind (d1)-(d5) nachzuweisen.

LEMMA 8.2. Sei $K := K^n := (-\varepsilon, \varepsilon)^n$ ein Würfel in \mathbb{R}^n . Zu jeder Funktion $q \in C_0^1(K, \mathbb{R})$ mit $\int_K q = 0$ existiert ein $w \in C_0^1(K, \mathbb{R}^n)$, so dass $q = \text{div}(w)$.

BEWEIS. Induktion über die Dimension n : Für $n = 1$ setze $w(t) := \int_{-\varepsilon}^t q(\xi) d\xi$. Dieses w hat kompakten Träger in K^1 .

Das Lemma sei in n Dimensionen bewiesen, wir wollen es in $n + 1$ Dimensionen beweisen. Die Koordinaten seien $(y, t) = (y_1, \dots, y_n, t)$. Setze

$$m(y) := \int_{-\infty}^{\infty} q(y, t) dt.$$

m ist C^1 und hat kompakten Träger im Würfel K^n . Wir können es also darstellen als $m(y) = \text{div}_n(g_1, \dots, g_n)(y)$. Sei nun $\tau \in C_0^1((-\varepsilon, \varepsilon), \mathbb{R})$ mit $\int \tau = 1$. Setze

$$g_{n+1}(y, t) := \int_{-\infty}^t (q(y, s) - \tau(s)m(y)) ds.$$

Dann hat g_{n+1} seinen Träger in K . Die Funktion

$$w(y, t) := (\tau(t)g_1(y), \dots, \tau(t)g_n(y), g_{n+1}(y, t))$$

erfüllt

$$\begin{aligned} \text{div } w(y, t) &= \tau(t)\text{div}_n(g_1, \dots, g_n)(y) + \partial_t g_{n+1}(y, t) \\ &= \tau(t)m(y) + q(y, t) - \tau(t)m(y) = q(y, t). \end{aligned}$$

□

LEMMA 8.3. Zu $\omega_1, \omega_2 \in \Omega_\varepsilon$ und $f \in C^2$ gibt es $v \in C^1(G, \mathbb{R}^n)$ mit Träger enthalten in G und

$$\omega_1(f(x) - y_0) \mathcal{J}_f(x) - \omega_2(f(x) - y_0) \mathcal{J}_f(x) = \text{div } v(x) \quad \forall x \in G.$$

BEWEIS. 1. *Beweis:* Berechne die Kofaktor-Matrix A von Df : Der Eintrag A_{jk} ist $(-1)^{j+k}$ mal der Determinante der Matrix, die man aus Df durch streichen der j -ten Spalte und der k -ten Zeile erhält. Man rechnet nach:

$$\sum_{j=1}^n \frac{\partial A_{jk}(x)}{\partial x_j} = 0 \quad \forall k.$$

Schreibe nun $\omega_1(y) - \omega_2(y) = \operatorname{div} w(y)$ und setze

$$v_j(x) = \sum_k w_k(f(x) - y_0) \cdot A_{jk}.$$

Man berechnet nun

$$\begin{aligned} \operatorname{div} v &= \sum_{k,j} \partial_j [w_k(f(x) - y_0)] \cdot A_{jk} + 0 \\ &= \sum_{k,j,i} \partial_i w_k(f(x) - y_0) \cdot \partial_j f_i(x) \cdot A_{jk} \\ &= \operatorname{div} w(f(x) - y_0) \cdot \mathcal{J}_f(x). \end{aligned}$$

2. *Beweis:* Wieder schreiben wir $\omega_1(y) - \omega_2(y) = \operatorname{div} w(y)$. Dann gilt mit den Differenzialformen $dV = dy_1 \wedge \dots \wedge dy_n$ und

$$w \, dS = \sum_j w_j(y) (-1)^j dy_1 \wedge \dots \wedge dy_{j-1} \wedge dy_{j+1} \wedge \dots \wedge dy_n$$

die Rechnung (für $y_0 = 0$)

$$\begin{aligned} (\omega_1(f(x)) - \omega_2(f(x))) \, \mathcal{J}_f(x) \, dx_1 \wedge \dots \wedge dx_n \\ = \operatorname{div} w(f(x)) \, df_1 \wedge \dots \wedge df_n \\ = f^*(\operatorname{div} w \, dV) = f^*d(w \, dS) = d(f^*(w \, dS)). \end{aligned}$$

Also ist die linke Seite eine Divergenz. \square

Mit obigem Lemma ist nachgewiesen, dass der Grad für C^2 -Abbildungen wohldefiniert ist.

PROPOSITION 8.4. *Der Abbildungsgrad aus Definition 8.1 erfüllt (d3) und (d5). Bei Homotopien wie in (d4) mit $h \in C^2(G \times [0, 1], \mathbb{R}^n)$ gilt:*

$$t \mapsto d(h(., t), G, y(t)) \text{ ist stetig.}$$

BEWEIS. a) Eigenschaft (d3) folgt sofort aus der Definition.

b) Wegen $y_0 \notin f(\partial G_1 \cup \partial G_2)$ und Abgeschlossenheit der Ränder gilt $\varepsilon := \operatorname{dist}(y_0, f(\partial G_1 \cup \partial G_2)) > 0$. Wir benutzen dieses ε , um die Grade $d(f, G_1 \cup G_2, y_0)$, $d(f, G_1, y_0)$ und $d(f, G_2, y_0)$ zu berechnen. Aussage (d5) folgt aus der Formel (8.1).

c) *Homotopie:* Die Mengen $R := \{(t, h(x, t)) : x \in \partial G, t \in [0, 1]\}$ und $Q := \{(t, y(t)) : t \in [0, 1]\}$ sind kompakt und haben daher einen positiven Abstand ε . Für $\omega \in \Omega_\varepsilon$ hängt der Ausdruck

$$d(h(., t), G, y(t)) = \int_G \omega(h(x, t) - y(t)) \, \mathcal{J}_{h(., t)}(x) \, dx$$

stetig von t ab. \square

8.2. Äquivalente Darstellungen. Notation: Summen über die leere Menge werden gleich 0 gesetzt.

SATZ 8.5 (Gezählte Nullstellen). *Falls y_0 ein regulärer Wert von f ist, gilt*

$$(8.2) \quad d(f, G, y_0) = \sum_{x \in f^{-1}(y_0)} \operatorname{sgn} \mathcal{J}_f(x).$$

BEWEIS. Für $f^{-1}(y_0) = \emptyset$ sind beide Seiten Null.
 $f^{-1}(y_0)$ ist kompakt. Da y_0 regulärer Wert ist, gilt $f^{-1}(y_0) = \bigcup_k \{x_k\}$ mit endlich vielen $x_k \in G$, und lokal gilt:

$f|_{B_\rho(x_k)} : B_\rho(x_k) \rightarrow f(B_\rho(x_k))$ ist ein Diffeomorphismus.

Insbesondere ist $\operatorname{sgn} \mathcal{J}_f$ konstant in der Kugel. Wir können annehmen, dass die Kugeln $B_\rho(x_k)$ disjunkt sind.

$\varepsilon = \operatorname{dist}(f(\partial G), y_0)$ wird eventuell verkleinert, um auch

$$\varepsilon \leq \operatorname{dist}(f(\bar{G} \setminus \bigcup_k B_\rho(x_k)), y_0)\}$$

zu erfüllen. Wir berechnen für $\omega \in \Omega_\varepsilon$:

$$\begin{aligned} d(f, G, y_0) &= \int_G \omega(f(x) - y_0) \mathcal{J}_f(x) dx \\ &= \sum_k \int_{B_\rho(x_k)} \omega(f(x) - y_0) \mathcal{J}_f(x) dx \\ &= \sum_k \operatorname{sgn} \mathcal{J}_f(x_k) \int_{B_\rho(x_k)} \omega(f(x) - y_0) |\mathcal{J}_f(x)| dx \\ &= \sum_k \operatorname{sgn} \mathcal{J}_f(x_k) \int_{f(B_\rho(x_k))} \omega(y - y_0) dy \\ &= \sum_k \operatorname{sgn} \mathcal{J}_f(x_k). \end{aligned}$$

\square

COROLLAR 8.6. *Der Abbildungsgrad aus Definition 8.1 erfüllt (d1)-(d5) für $f \in C^2$ und glatte Homotopien $h \in C^2(G \times [0, 1], \mathbb{R}^n)$.*

BEWEIS. (d3) und (d5) wurden in Proposition 8.4 nachgewiesen.

Falls y_0 regulärer Wert ist, wenden wir Gleichung (8.2) an. Ganzzahligkeit und Lösungskriterium folgen.

Sei y_0 ein kritischer Wert (nicht regulär). Dann gibt es nach dem Lemma von Sard 7.7 beliebig nahe an y_0 einen regulären Wert y_1 . Der Grad $d(f, G, y_1)$ ist ganzzahlig. Da der Grad stetig von y abhängt (Proposition 8.4), ist auch $d(f, G, y_0)$ ganzzahlig.

Es gilt:

$$d(f, G, y_0) \neq 0 \Rightarrow d(f, G, y_1) \neq 0 \Rightarrow \exists x_1 : f(x_1) = y_1.$$

Wir finden eine Folge x_n von approximativen Lösungen und im Limes ein $x \in \bar{G}$ mit $f(x) = y_0$ (verwende Kompaktheit von \bar{G} und Stetigkeit von f).

Für Homotopien gilt: Der Abbildungsgrad $d(h(., t), G, y_0(t))$ ist ganzzahlig und (nach Proposition 8.4) stetig. Dann ist der Grad konstant. \square

SATZ 8.7 (Randintegral). *Für $f \in C^2(G, \mathbb{R}^n) \cap C^1(\bar{G}, \mathbb{R}^n)$ gilt*

$$(8.3) \quad d(f, G, y_0) = \frac{1}{|S^{n-1}|} \int_{\partial G} \frac{\langle f - y_0, \nu_f \rangle}{|f - y_0|^n} d\sigma.$$

Dabei $\nu_f(x) := \mathcal{A}_f(x) \cdot \nu(x)$, ν die Normale an ∂G und $\mathcal{A}_f := (A_{jk})_{jk}$ die Kofaktormatrix $A_{jk} = (-1)^{j+k} \det(\partial_l f^m)_{l \neq k, m \neq j}$.

BEWEIS. Wir verwenden zur Berechnung des Grades eine rotationssymmetrische Funktion $\omega(y) = w(|y|)$, also

$$d(f, G, y_0) = \int_G w(|f(x) - y_0|) \mathcal{J}_f(x) dx.$$

Unser Ziel ist es wie in Lemma 8.3, die rechte Seite als Divergenz zu schreiben. Als ersten Schritt schreiben wir ω als Divergenz: Setze für $r > 0$

$$\varphi(r) := \frac{1}{r^n} \int_0^r w(t) t^{n-1} dt.$$

Für $r > \varepsilon$ gilt $\varphi(r) = \frac{1}{r^n} \cdot \frac{1}{|S^{n-1}|}$. Wir setzen

$$\Phi(y) := \varphi(|y|) \cdot y, \quad \Phi \in C^2(\mathbb{R}^n, \mathbb{R}^n).$$

Dann gilt mit $r = |y|$

$$\begin{aligned} \operatorname{div} \Phi &= \sum_k \partial_k \Phi_k = \varphi(|y|) \cdot n + \partial_r \varphi(|y|) \cdot r \\ &= \varphi(|y|) \cdot n - n \cdot \frac{1}{r} \varphi(|y|) \cdot r + \frac{1}{r^n} w(r) r^{n-1} \cdot r \\ &= w(r). \end{aligned}$$

Wie in Lemma 8.3 gilt nun

$$\omega(f(x) - y_0) \mathcal{J}_f(x) = \operatorname{div} \left[\sum_k \Phi_k(f(x) - y_0) \cdot A_{jk} \right]_j.$$

Mit der Anwendung des Gaußschen Satzes finden wir

$$\begin{aligned}
d(f, G, y_0) &= \int_G w(|f(x) - y_0|) \mathcal{J}_f(x) dx \\
&= \int_G \operatorname{div} \left[\sum_k \Phi_k(f(x) - y_0) \cdot A_{jk} \right]_j dx \\
&= \int_{\partial G} \nu \cdot \left[\sum_k \Phi_k(f(x) - y_0) \cdot A_{jk} \right]_j d\sigma \\
&= \frac{1}{|S^{n-1}|} \int_{\partial G} \sum_j \nu_j \cdot \sum_k \frac{(f(x) - y_0)_k}{|f(x) - y_0|^n} \cdot A_{jk} d\sigma
\end{aligned}$$

und damit die behauptete Formel. \square

Mit obiger Formel lässt sich der Abbildungsgrad aus den Werten von f und Df auf dem Rand berechnen.

Bemerkung: Der Abbildungsgrad hängt nur von der Werten von f am Rand ab. Sei $g : \bar{G} \rightarrow \mathbb{R}^n$ mit $g = f$ auf ∂G , dann ist $h(x, t) := tf(x) + (1-t)g(x)$ eine erlaubte Homotopie und deswegen $d(g, G, y_0) = d(f, G, y_0)$.

8.3. Der Grad für C^0 -Abbildungen.

SATZ 8.8 (Satz von Rouché). *Seien $f, g \in C^0(\bar{G}, \mathbb{R}^n) \cap C^2(G, \mathbb{R}^n)$ und $y_0 \in \mathbb{R}^n$. Falls*

$$|f(x) - g(x)| < |f(x) - y_0| \quad \forall x \in \partial G,$$

dann gilt

$$d(f, G, y_0) = d(g, G, y_0).$$

BEWEIS. Die Homotopie

$$h(x, t) := (1-t) \cdot f(x) + t \cdot g(x)$$

ist C^2 und zulässig, da für alle $x \in \partial G$

$$|h(x, t) - y_0| \geq |f(x) - y_0| - t \cdot |f(x) - g(x)| > 0.$$

Wende die Homotopieeigenschaft (d4) an, gezeigt für glatte h in Proposition 8.4. \square

Wegen des Satzes von Rouché ist die nachfolgende Definition möglich (unabhängig von der Wahl von g).

DEFINITION 8.9 (Grad für C^0 -Funktionen). *Sei $G \subset \mathbb{R}^n$ beschränkt, offen, $f \in C^0(\bar{G}, \mathbb{R}^n)$ und $y_0 \notin f(\partial G)$. Dann wählen wir eine Approximation $g \in C^0(\bar{G}, \mathbb{R}^n) \cap C^2(G, \mathbb{R}^n)$ von f mit $|f(x) - g(x)| < \frac{1}{4}|f(x) - y_0|$ für alle $x \in \partial G$ und definieren*

$$d(f, G, y_0) := d(g, G, y_0).$$

Für die Wohldefiniertheit des Grades: Seien g_1 und g_2 wie in der Definition. Wegen

$$\begin{aligned}|f(x) - y_0| &\leq |f(x) - g_1(x)| + |g_1(x) - y_0| \\ &\leq \frac{1}{4}|f(x) - y_0| + |g_1(x) - y_0|\end{aligned}$$

gilt dann $|f(x) - y_0| \leq \frac{4}{3}|g_1(x) - y_0|$. Es folgt

$$\begin{aligned}|g_1(x) - g_2(x)| &\leq |g_1(x) - f(x)| + |f(x) - g_2(x)| \\ &\leq \frac{1}{2}|f(x) - y_0| < |g_1(x) - y_0|.\end{aligned}$$

SATZ 8.10. *Mit dieser Definition erfüllt d die Axiome (d1)-(d5).*

BEWEIS. (d1), (d3) und (d5) folgen unmittelbar aus der Definition. Zu (d4): Approximiere eine (nur) stetige Homotopie h durch eine C^2 Abbildungen h_1 . Es gilt $d(h(., t), G, y_0(t)) = d(h_1(., t), G, y_0(t))$, und damit konstant. Zu (d2): Falls $d(f, G, y_0) \neq 0$, so haben beliebige Approximationen f_n von f Lösungen $f_n(x_n) = y_0$. Jeder Limes x von (Teilfolgen von) x_n ist eine Lösung zu $f(x) = y_0$. \square

Ein topologisches Resultat: Satz von Hopf

Zu einer Abbildung $\varphi : S^n \rightarrow S^n$ können wir Fortsetzungen $\bar{\varphi} : B_1 \rightarrow B_1$ bilden. Der Abbildungsgrad von $\bar{\varphi}$ hängt nur von den Randwerten ab und man kann definieren

$$\deg(\varphi) := d(\bar{\varphi}, B_1, 0).$$

Wir wissen, dass zwei zueinander homotope Abbildungen $\varphi_1, \varphi_2 : S^n \rightarrow S^n$ denselben Grad (\deg) haben. Es gilt auch die Umkehrung. Ohne Beweis geben wir hier dieses wichtige Resultat von E. Hopf an.

SATZ 8.11 (Satz von Hopf). *Für zwei Abbildungen $\varphi_1, \varphi_2 : S^n \rightarrow S^n$ gilt:*

$$\varphi_1 \sim \varphi_2 \iff \deg(\varphi_1) = \deg(\varphi_2).$$

9. Anwendungen des Abbildungsgrades

9.1. Wichtiges Hilfsmittel: Fortsetzungssätze.

SATZ 9.1 (Fortsetzungssatz von Tietze). *X metrischer Raum, $A \subset X$ abgeschlossen, Y Banachraum. $g : A \rightarrow Y$ stetig. Dann existiert eine stetige Fortsetzung*

$$\bar{g} : X \rightarrow \text{conv}(g(A)), \quad \bar{g}|_A = g.$$

Bemerkungen: (i) Es reicht aus, wenn Y ein topologischer Raum ist.
(ii) Falls g kompakt ist, so kann auch \bar{g} kompakt konstruiert werden.

BEWEIS. Zu jedem $x \in X \setminus A$ wähle $\delta(x) := \frac{1}{2}\text{dist}(x, A) > 0$. Aus der offenen Überdeckung $B_{\delta(x)}(x)$ wähle eine lokal endliche Überdeckung $\{B_j\}_j$ aus. Eine zugehörige Zerlegung der 1 ist

$$\varphi_j(x) := \frac{\text{dist}(x, X \setminus B_j)}{\sum_k \text{dist}(x, X \setminus B_k)}.$$

Zu jedem B_j gibt es $a_j \in A$ mit $\text{dist}(a_j, B_j) < 2\text{dist}(A, B_j)$. Setze

$$\bar{g}(x) := \begin{cases} g(x) & x \in A, \\ \sum_j \varphi_j(x)g(a_j) & x \notin A. \end{cases}$$

$\bar{g} : X \rightarrow \text{conv}(g(A))$ folgt sofort, ebenso Stetigkeit von \bar{g} auf $X \setminus \partial A$.

Zeige Stetigkeit im Punkt $x^* \in \partial A$. Für $x \notin A$ gilt

$$\|\bar{g}(x) - \bar{g}(x^*)\| \leq \sup\{\|g(a_j) - g(x^*)\| : j \text{ mit } x \in B_j\}.$$

Man rechnet nach: Für $\text{dist}(x^*, x) \leq \delta$ gilt $\text{dist}(A, B_j) \leq \delta$ und damit dann auch

$$\begin{aligned} \text{dist}(x^*, a_j) &\leq \delta + \text{dist}(a_j, x) \\ &\leq \delta + \text{dist}(a_j, B_j) + \delta(x_j) \leq 3\delta + \delta(x_j). \end{aligned}$$

Für $\delta(x_j)$ gilt nach Definition

$$\delta(x_j) \leq \frac{1}{2}\text{dist}(x_j, A) \leq \frac{1}{2}(\text{dist}(B_j, A) + \delta(x_j)),$$

also

$$\delta(x_j) \leq \text{dist}(B_j, A) \leq \delta.$$

Für x nahe x^* sind daher alle a_j auch nahe x^* und daher ist \bar{g} stetig. \square

Das nächste Lemma sagt aus, dass im \mathbb{R}^m , $m > n$, viel Platz für Fortsetzungen ist.

LEMMA 9.2. Sei $W \subset \mathbb{R}^n$ ein Würfel, $K \subset W$ kompakt und $\psi : K \rightarrow \mathbb{R}^m \setminus \{0\}$ stetig mit $m > n$. Dann kann ψ stetig fortgesetzt werden zu

$$\bar{\psi} : W \rightarrow \mathbb{R}^m \setminus \{0\}.$$

BEWEIS. a) Sei ψ_1 eine Fortsetzung von ψ nach Tietze. Wir approximieren mit einer Funktion $\psi_2 \in C^\infty$, $|\psi_2 - \psi| < \varepsilon$ auf K . Das Bild von ψ_2 ist eine Nullmenge im \mathbb{R}^m , daher finden wir $y_0 \in \mathbb{R}^m$ mit $|y_0| < \varepsilon$, so dass $\psi_2 - y_0$ die Null nicht trifft. Ohne Einschränkung können wir annehmen, dass ψ_2 die Null nicht trifft (sonst Übergang zu $\psi_2 - y_0$).

b) Modifizierte ψ_2 zu einer Funktion ψ_3 , die von der Null weit weg bleibt: $\psi_3 = \psi_2$ auf K und

$$|\psi_3(x)| \geq \frac{1}{2} \min\{|\psi_2(y)| : y \in K\}$$

für alle $x \in W$. Erreiche dies zum Beispiel mit Hilfe von

$$\eta(t) := \begin{cases} \frac{2t}{c} & t \leq \frac{c}{2}, \\ 1 & t > \frac{c}{2}. \end{cases}$$

Setze

$$\psi_3(x) := \frac{\psi_2(x)}{\eta(|\psi_2(x)|)}.$$

c) Auf K gilt $\psi_3 - \psi = \psi_2 - \psi < \varepsilon$. Setze $\psi_3 - \psi$ wie in Satz 9.1 auf W fort, auch die Fortsetzung bildet nach $B_\varepsilon(0)$ ab. Ziehe diese Fortsetzung von der Funktion ψ_3 ab. Dies liefert die gewünschte Fortsetzung von ψ . \square

Ungerade Fortsetzungen

G heißt *symmetrisch*, falls $G = -G$, eine Funktion $f : G \rightarrow \mathbb{R}^n$ heißt *ungerade*, falls $f(-x) = -f(x)$ für alle $x \in G$.

Das nachfolgende Lemma überträgt das vorige in die Situation, dass ψ ungerade sein soll.

LEMMA 9.3. Sei $D \subset \mathbb{R}^n$ beschränkt, offen und symmetrisch mit $0 \notin \bar{D}$. Sei $\psi : \partial D \rightarrow \mathbb{R}^m \setminus \{0\}$ stetig und ungerade, $m > n$. Dann kann ψ stetig fortgesetzt werden zu einer ungeraden Funktion

$$\bar{\psi} : \bar{D} \rightarrow \mathbb{R}^m \setminus \{0\}.$$

BEWEIS. Induktion über n . Für $n = 1$ elementar. Das Lemma sei nun für $n - 1$ bewiesen.

Betrachte $D_0 := D \cap \{x : x_n = 0\} \subset \mathbb{R}^{n-1}$. Auf ∂D_0 haben wir ein stetiges ψ gegeben, das wir nach Induktionsvoraussetzung fortsetzen zu einer ungeraden Funktion

$$\psi_0 : \bar{D}_0 \rightarrow \mathbb{R}^m \setminus \{0\}.$$

Seien nun $D_+ := D \cap \{x : x_n > 0\}$. Wir setzen nach Lemma 9.2 die Randwerte ψ auf $\partial D_+ \cap \{x_n > 0\}$ und ψ_0 auf $\partial D_+ \cap \{x_n = 0\} = D_0$ zu einer Funktion

$$\tilde{\psi} : \bar{D}_+ \rightarrow \mathbb{R}^m \setminus \{0\}$$

fort. Dieses wird durch Spiegelung ungerade auf ganz \bar{D} fortgesetzt. \square

9.2. Die Sätze von Igel und Borsuk.

SATZ 9.4 (Satz vom Igel). Die Dimension n sei ungerade und das offene, beschränkte Gebiet $G \subset \mathbb{R}^n$ enthalte die 0. Dann gilt:

Zu jedem stetigen $f : \partial G \rightarrow \mathbb{R}^n \setminus \{0\}$ gibt es ein $x \in \partial G$ und $\lambda \in \mathbb{R}$ mit

$$f(x) = \lambda x.$$

BEWEIS. Sei $\bar{f} : \bar{G} \rightarrow \mathbb{R}^n$ eine stetige Fortsetzung der Randwerte. Da n ungerade ist, gilt $d(-\text{id}, G, 0) = -1$. Falls $d(\bar{f}, G, 0) \neq -1$ kann $h(x, t) := (1 - t) \cdot \bar{f}(x) - t \cdot x$ keine zulässige Homotopie sein. Es gibt also $x \in \partial G$ mit

$$0 = h(x, t) := (1 - t) \cdot f(x) - t \cdot x.$$

Da dies nicht für $t = 1$ gilt, ist die Behauptung bewiesen.

Falls $d(\bar{f}, G, 0) = -1$ betrachten wir die analoge Homotopie zur Identität und kommen auf dasselbe Ergebnis. \square

SATZ 9.5 (Satz von Borsuk). *Sei $G \subset \mathbb{R}^n$ symmetrisch mit $0 \in G$. $f : \bar{G} \rightarrow \mathbb{R}^n$ sei stetig, ungerade und $0 \notin f(\partial G)$. Dann ist $d(f, G, 0)$ ungerade.*

BEWEIS. Wir wählen $\varepsilon > 0$ klein, so dass $\bar{B}_\varepsilon(0) \subset G$. Sei \bar{f} eine Fortsetzung von: (a) f auf ∂G und (b) id auf $\partial B_\varepsilon(0)$. Wir benutzen, dass d nur von den Randwerten abhängt, weiterhin (d5) und (d3):

$$\begin{aligned} d(f, G, 0) &= d(\bar{f}, G, 0) = d(\bar{f}, B_\varepsilon(0), 0) + d(\bar{f}, G \setminus \bar{B}_\varepsilon(0), 0) \\ &= 1 + d(\bar{f}, G \setminus \bar{B}_\varepsilon(0), 0). \end{aligned}$$

Es bleibt zu zeigen, dass letzterer Grad gerade ist. Bemerke: Könnten wir annehmen, dass \bar{f} ungerade ist und C^1 , dann wären wir sofort fertig. Dies liefert einen alternativen Beweis, siehe Gromes [5] bzw. [4].

Die Randwerte von \bar{f} auf $G \setminus \bar{B}_\varepsilon(0)$ sind ungerade. Diese Information reicht uns: Die Abbildung

$$\bar{f} : \partial(G \setminus B_\varepsilon(0)) \cap \{x_n = 0\} \rightarrow \mathbb{R}^n \setminus \{0\}$$

wird ungerade fortgesetzt zu

$$\bar{\psi} : (\bar{G} \setminus B_\varepsilon(0)) \cap \{x_n = 0\} \rightarrow \mathbb{R}^n \setminus \{0\}$$

nach Lemma 9.3. Damit haben wir ungerade Randwerte g definiert durch:

(a) $g = \text{id}$ auf ∂B_ε , (b) $g = \bar{\psi}$ auf $\bar{G} \cap \{x_n = 0\}$, (c) $g = \bar{f}$ auf $\partial G \cap \{x_n > 0\}$ für das Gebiet $G_+^\varepsilon := (G \setminus \bar{B}_\varepsilon(0)) \cap \{x_n > 0\}$. Da die Randwerte nirgends verschwinden, können wir ausrechnen:

$$\begin{aligned} d(\bar{f}, G \setminus \bar{B}_\varepsilon(0), 0) &= d(g, G_+^\varepsilon, 0) + d(-g \circ (-\text{id}), -G_+^\varepsilon, 0) \\ &= 2d(g, G_+^\varepsilon, 0). \end{aligned}$$

Die letzte Gleichheit sieht man an den Formeln für den Grad. \square

Anwendungen von Borsuk. Der nächste Satz ist der Satz vom Wetter: Es gibt auf der Erde immer zwei einander gegenüberliegende Punkte, an denen Temperatur und Luftdruck identisch sind.

COROLLAR 9.6 (Der Satz vom Wetter). *Sei $G \subset \mathbb{R}^n$ offen, beschr. und symmetrisch und $0 \in G$. Für $m < n$ sei $f : \partial G \rightarrow \mathbb{R}^m$ stetig. Dann existiert ein $x \in \partial G$ mit $f(x) = f(-x)$.*

BEWEIS. Angenommen, $g(x) = f(x) - f(-x)$ habe keine Nullstelle auf ∂G . Wir setzen zunächst die Randwerte fort zu $g : \bar{G} \rightarrow \mathbb{R}^m \subset \mathbb{R}^n$. Der Satz von Borsuk liefert $d(g, G, 0) \neq 0$. Wegen der Homotopieeigenschaft gilt auch noch $d(g, G, \delta e_n) \neq 0$ für kleines $\delta \neq 0$. Das Lösungskriterium impliziert, dass es eine Lösung x von $g(x) = \delta e_n \notin \mathbb{R}^m$ existiert. Ein Widerspruch. \square

Um den nachfolgenden Satz zu genießen, denke man bei A_1, A_2, A_3 an Mengen, die von Brot, Käse und Schinken eingenommen werden.

COROLLAR 9.7 (Das Sandwich Problem). *Seien A_1, A_2, A_3 drei meßbare Teilmengen des \mathbb{R}^3 mit endlichem Volumen. Dann gibt es eine Ebene, die alle drei Mengen in zwei gleiche Volumina teilt.*

BEWEIS. Sei $x \in S^2$. Wir betrachten die Schar $E(t)$ von Ebenen senkrecht zu x , die die Achse $\mathbb{R}x$ im Punkt tx schneiden. Beginnend mit $t = -\infty$ finden wir einen ersten Punkt $t_1(x)$ und einen letzten Punkt $t_2(x)$, an dem $E(t)$ das Volumen von A_3 halbiert. Wir definieren $t_0(x) := \frac{1}{2}(t_1(x) + t_2(x))$. Die Volumina

$$V_k(x) := |\{y \in A_k : y \cdot x < t_0(x)\}|, \quad k = 1, 2$$

hängen stetig von x ab und definieren $V : S^2 \rightarrow \mathbb{R}^2$. Nach Corollar 9.6 gibt es einen Punkt $x_0 \in S^2$ mit $V_1(x_0) = V_1(-x_0)$ und $V_2(x_0) = V_2(-x_0)$. Dieses x_0 liefert die gesuchte Ebene. \square

SATZ 9.8 (Satz der Gebietsinvarianz). *Sei $G \subset \mathbb{R}^n$ offen und $f : G \rightarrow \mathbb{R}^n$ stetig und lokal injektiv. Dann ist f eine offene Abbildung.*

BEWEIS. Wir betrachten einen beliebigen Punkt in G . O.E. sei der Punkt $0 \in G$ und $f(0) = 0$. Wir müssen zeigen, dass für $r > 0$ gilt:

$$\exists \varepsilon > 0 : B_\varepsilon(0) \subset f(B_r(0)).$$

Durch eventuelles Verkleinern von r können wir annehmen, dass $f|_{B_r(0)}$ injektiv ist. Wir betrachten

$$h(x, t) := f\left(\frac{1}{1+t}x\right) - f\left(-\frac{t}{1+t}x\right)$$

für $t \in [0, 1]$ und $x \in B_r(0)$. h ist eine stetige Homotopie von f nach dem ungeraden $h(x, 1) = f(\frac{1}{2}x) - f(-\frac{1}{2}x)$. Die Homotopie ist zulässig, denn $h(x, t) = 0$ impliziert

$$f\left(\frac{1}{1+t}x\right) = f\left(-\frac{t}{1+t}x\right)$$

und wegen der Injektivität von f

$$\frac{1}{1+t}x = -\frac{t}{1+t}x,$$

also $x = 0$. Dies ist kein Randpunkt. Wegen (d4) und dem Satz von Borsuk gilt

$$d(f, G, 0) = d(h(., 1), G, 0) \neq 0.$$

Also ist $f(x) = y$ für eine ganze Umgebung des Nullpunkts (im Bild) lösbar. \square

COROLLAR 9.9 (Ein Atlas mit 3 Karten). *$G \subset \mathbb{R}^n$ offen, beschr. und symmetrisch und $0 \in G$. Sei ∂G überdeckt mit n (das n der Dimension!) abgeschlossenen Mengen A_k . Dann enthält eines der A_k ein Paar von Antipodenpunkten x und $-x$.*

BEWEIS. Angenommen, nicht. Dann ist $\bigcap_k A_k = \{\}$ (denn ein x darin hätte auch $-x$ in einem der A_k und die Antipodenpunkte $x, -x$ wären dann beide in A_k). Definiere

$$d_k(x) := \text{dist}(x, A_k), \text{ und } d(x) := \sum_k d_k(x) > 0,$$

Positivität folgt aus $\bigcap_k A_k = \{\}$. Betrachte jetzt $f : \partial G \rightarrow \mathbb{R}^{n-1} \subset \mathbb{R}^n$,

$$f(x) := \left(\frac{d_1(x)}{d(x)}, \dots, \frac{d_{n-1}(x)}{d(x)} \right).$$

Nach Corollar 9.6 gibt es zu f einen Punkt x mit $f(x) = f(-x)$. x gehört zu einem der A_k .

1. Fall: $x \in A_k$ für ein $k < n$. Dann $d_k(-x) = d_k(x) = 0$. Beide Punkte x und $-x$ sind in A_k .

2. Fall: $x \notin A_k$ für alle $k < n$. Dann gilt $d_k(-x) = d_k(x) > 0$ für alle $k < n$. Also sind beide Punkte x und $-x$ in A_n . \square

9.3. Der Satz von Brouwer.

SATZ 9.10 (Starker Satz von Brouwer). *Sei $K \subset \mathbb{R}^n$ nichtleer, kompakt und konvex oder homöomorph zu einer solchen Menge. $f : K \rightarrow K$ sei stetig. Dann hat f einen Fixpunkt.*

BEWEIS. In Satz 7.5 haben wir dieses Resultat gezeigt für den Fall, dass K eine Kugel ist. Für konvexas $K \subset B_R(0)$ können wir mit Satz 9.1 f fortsetzen zu einer Abbildung

$$\bar{f} : B_R(0) \rightarrow K \subset B_R(0).$$

Satz 7.5 liefert einen Fixpunkt $x \in B_R(0)$ für \bar{f} . Wegen $x = \bar{f}(x) \in K$ ist x auch Fixpunkt für f .

Wenn K nur homöomorph zu einer kompakten und konvexen Menge ist und der Homöomorphismus Φ ist, dann betrachte $\Phi \circ f \circ \Phi^{-1}$. \square

Wir betrachten die gewöhnliche Differenzialgleichung

$$(9.1) \quad u'(t) = f(t, u(t)), \quad u(t=0) = u_0.$$

Wir nehmen an, dass f Lipschitzstetig ist. Dann ist das Problem für jedes u_0 eindeutig lösbar und die Lösung $u(t)$ hängt stetig von den Anfangswerten u_0 ab. Dann gibt es den stetigen Lösungsoperator $T_t : u_0 \mapsto u(t)$. T_t heißt Poincaré-Abbildung oder Zeit- t -Abbildung.

Oft hat man invariante Mengen für den Fluss. Gilt zum Beispiel $\langle x, f(t, x) \rangle < 0$ für alle $x \in \partial B_R(0)$, so können Lösungen u wegen

$$\frac{d}{dt} \|u\|^2 = 2 \langle u, u' \rangle = 2 \langle u, f \rangle < 0$$

auf $\partial B_R(0)$ die Kugel nicht verlassen.

SATZ 9.11 (Periodische Lösungen gew. DGL). *Sei f in der ersten Komponente periodisch mit Periode p und die Menge A invariant. Wenn A homöomorph zu B_1 ist, dann gibt es eine p -periodische Lösung von $u'(t) = f(t, u(t))$.*

BEWEIS. Sei $T_p : A \rightarrow X$ die Poincaré-Abbildung. Nach dem Satz von Brouwer hat T_p einen Fixpunkt. Dieser liefert die gesuchte periodische Lösung. \square

SATZ 9.12 (Fixpunkte von Flüssen). *Sei A homöomorph zu \bar{B}_1 . Weiterhin sei A invariant unter dem Fluss von $u' = f(u)$. Dann hat der Fluss einen Fixpunkt u_0 , d.h. $f(u_0) = 0$.*

BEWEIS. Wähle eine Folge $0 < \tau_k \rightarrow 0$ für $k \rightarrow \infty$. Nach Satz 9.11 gibt es für jedes k eine τ_k -periodische Lösung mit Startwert $x_k = T_{\tau_k}x_k$.

Wir schreiben ein festes t als $t = \gamma_k \tau_k + r_k$ mit $\gamma_k \in \mathbb{N}$ und $r_k < \tau_k$. Ohne Einschränkung können wir $x_k \rightarrow x$ annehmen. Wir wollen $T_t x = x$ nachweisen. Zunächst

$$T_t x = \lim_{k \rightarrow \infty} T_t x_k = \lim_{k \rightarrow \infty} T_{r_k} x_k.$$

Wir rechnen

$$\|T_{r_k} x_k - x\| \leq \|T_{r_k} x_k - x_k\| + \|x_k - x\|.$$

Wegen

$$\|T_{r_k} x_k - x_k\| = \left\| \int_0^{r_k} f(T_\xi x_k) \, d\xi \right\| \leq \int_0^{r_k} M \leq M \cdot r_k \rightarrow 0$$

gehen beide Terme der rechten Seite gegen 0. \square

10. Der Abbildungsgrad im Banachraum

Eine Motivation für den Abbildungsgrad ist der Satz von Schauder aus Theorem 5.7. Wir stellen fest, dass der Beweis für Theorem 9.10 sich auf den unendlichdimensionalen Fall überträgt, sobald man den Abbildungsgrad im Banachraum zur Verfügung hat.

10.1. Konstruktion des Grades. Wir betrachten ab jetzt immer f von der Form $f = \text{id} + g$. Wir führen die Konstruktion nur für beschränkte Mengen G durch.

Eine der wesentlichen Eigenschaften solcher Abbildungen f ist die folgende: Für $A \subset X$ abgeschlossen ist $f(A)$ auch abgeschlossen. In unserer Anwendung bei der Definition des Grades ist also $f(\partial G)$ abgeschlossen und y_0 hat einen positiven Abstand zu ∂G .

PROPOSITION 10.1. *$G \subset X$ beschränkt, $f : \bar{G} \rightarrow X$ von der Form $f = \text{id} + g$ mit $g : \bar{G} \rightarrow X$ kompakt. Dann ist f eine abgeschlossene Abbildung und eigentlich (Urbilder kompakter Mengen sind kompakt).*

BEWEIS. Sei $A \subset \bar{G}$ abgeschlossen. Wir betrachten eine Folge $f(x_n)$ in $f(A)$ mit $f(x_n) \rightarrow y$ für $n \rightarrow \infty$. Dann gilt $x_n - g(x_n) \rightarrow y$. Wegen Kompaktheit von g finden wir eine Teilfolge mit $g(x_n) \rightarrow z$. Dann konvergiert x_n gegen $x = y + z$ und wegen der Stetigkeit gilt $f(x) = y$. Also gilt $y \in f(A)$.

Sei $B \subset Y$ kompakt. Dann ist $M := f^{-1}(B) = \{x : f(x) \in B\}$ abgeschlossen. Wegen $M = \{x : x + g(x) \in B\}$ gilt $M \subset B - \text{cl}(g(\bar{G}))$, also ist M präkompakt und deswegen kompakt. \square

Die nächste Formel besagt, dass, wenn g niederdimensional ist (m), man den Grad von f auch im m -dimensionalen ausrechnen kann.

PROPOSITION 10.2 (Reduktionsformel). *Sei $G \subset \mathbb{R}^n$ offen, beschränkt und $m < n$. Setze $E^m := \mathbb{R}^m \times \{0\}_{\mathbb{R}^{n-m}} \subset \mathbb{R}^n$. Es sei $f = \text{id} + g$ mit*

$$g \in C^0(\bar{G}, E^m).$$

Weiterhin sei $y_0 \in E^m$ mit $y_0 \notin f(\partial G)$. Bezeichne mit π und π^* die Identifikationen $\pi : E^m \rightarrow \mathbb{R}^m$ und $\pi^* : \mathbb{R}^m \rightarrow E^m$. Dann gilt

$$d_{\mathbb{R}^n}(f, G, y_0) = d_{\mathbb{R}^m}(\pi \circ f \circ \pi^*, \pi(G \cap E^m), \pi y_0).$$

BEWEIS. O.E.: $f \in C^1$ und y_0 regulärer Wert. Nach Formel (8.2) reicht es, für $f^* := \pi \circ f \circ \pi^*$ zu zeigen: $\text{sgn}\mathcal{J}_f(x) = \text{sgn}\mathcal{J}_{f^*}(\pi x)$. Begründung: Die Urbilder von y_0 liegen alle in E^m , denn $x + g(x) = y_0$ impliziert $x \in E^m$.

Wir schreiben $g^* = \pi \circ g \circ \pi^*$ für die ersten $m \times m$ Einträge und berechnen

$$\mathcal{J}_f = \det \begin{pmatrix} \text{id}_m + Dg^* & \star \\ 0 & \text{id}_{n-m} \end{pmatrix} = \det(\text{id}_m + Dg^*) = \mathcal{J}_{f^*}.$$

Insbesondere stimmen die Vorzeichen überein. \square

Für $y_0 \notin f(\partial G)$ wissen wir, dass $\varepsilon := \text{dist}(y_0, f(\partial G))$ größer als 0 ist ($f(\partial G)$ ist abgeschlossen, also $X \setminus f(\partial G)$ offen, also für ein $\varepsilon > 0$ $B_\varepsilon(y_0) \cap f(\partial G) = \{\}$). Nach Satz 5.3 können wir die Abbildung f durch endlichdimensionales f_ε approximieren mit $\|f - f_\varepsilon\|_\infty < \frac{\varepsilon}{3}$. Mit einem endlichdimensionalen Unterraum $E_\varepsilon \subset X$ gilt

$$f_\varepsilon|_{E_\varepsilon \cap \bar{G}} : E_\varepsilon \cap \bar{G} \rightarrow E_\varepsilon.$$

Damit können wir den Grad im Banachraum definieren.

DEFINITION 10.3 (Grad im Banachraum). *Wähle E_ε und f_ε wie oben, ohne Einschränkung mit $y_0 \in E_\varepsilon$. Setze*

$$d(f, G, y_0) := d(f_\varepsilon|_{E_\varepsilon \cap \bar{G}}, E_\varepsilon \cap \bar{G}, y_0).$$

Wir müssen zeigen, dass diese Definition unabhängig ist von der Approximation f_ε . Dazu seien f_1 und f_2 zwei solche Approximationen mit Räumen E_1 und E_2 . Wir berechnen die Grade der rechten Seite in dem größeren Unterraum $F := \text{conv}(E_1, E_2)$. Es gilt

$$d(f_i|_{E_i \cap \bar{G}}, E_i \cap \bar{G}, y_0) = d(f_i|_{F \cap \bar{G}}, F \cap \bar{G}, y_0)$$

wegen des Reduktionssatzes 10.2. Es bleibt zu zeigen

$$d(f_1|_{F \cap \bar{G}}, F \cap \bar{G}, y_0) = d(f_2|_{F \cap \bar{G}}, F \cap \bar{G}, y_0).$$

Dies folgt aber sofort aus dem Satz von Rouché wegen

$$\|f_1 - f_2\| < \frac{2\varepsilon}{3} = \frac{2}{3} \text{dist}(y_0, f(\partial G)) \leq \text{dist}(y_0, f_1(\partial G)).$$

Es bleibt, die Eigenschaften des Grades nachzuweisen.

SATZ 10.4. *Der Abbildungsgrad aus Definition 10.3 erfüllt die Eigenschaften (d1)-(d5).*

BEWEIS. (d1), (d3) und (d5) folgen sofort aus der Definition.

(d2): Sei $d(f, G, y_0) \neq 0$. Dann hat für eine Familie von Abbildungen $f_k \rightarrow f$, $f_k = \text{id} + g_k$, $g_k : \bar{G} \rightarrow E_k$, E_k endlichdimensional, der Punkt y_0 jeweils ein Urbild x_k . Es gilt

$$y_0 = x_k + g_k(x_k) = x_k + g(x_k) + [g_k(x_k) - g(x_k)].$$

Nach Auswahl einer Teilfolge gilt also

$$x_k = y_0 - g(x_k) - [g_k(x_k) - g(x_k)] \rightarrow y_0 - g_0 =: x_0$$

für ein g_0 wegen der Kompaktheit von g . Wegen der Stetigkeit von f ist x_0 dann auch ein Urbild von y_0 ,

$$f(x_0) = \lim f(x_k) = \lim f_k(x_k) = y_0.$$

(d4): Bemerke, dass wir im Banachraum fordern: $h(., t) = \text{id} + g(., t)$, $g(., t)$ kompakt, für alle $t \in [0, 1]$. Dann ist g auch auf $\bar{G} \times [0, 1]$ eine kompakte Abbildung (Übung) und wir können g endlichdimensional approximieren. Wir dürfen diese Approximation zur Berechnung des Grades verwenden und folgern aus dem endlichdimensionalen Resultat, dass der Grad konstant ist. \square

Ohne dies vorzurechnen bemerken wir, dass sich die Sätze von Borsuk und von der Gebietsinvarianz auf die unendlichdimensionale Situation übertragen.

Als abstrakte Anwendung geben wir hier einen eleganten Beweis der Fredholm-Alternative, Corollar 6.5.

BEWEIS VON COROLLAR 6.5. Betrachte $L = \lambda - K$. Wir nehmen an, dass $\ker(L) = \{0\}$ und wollen folgern, dass $R(L) = Y$. Wir betrachten $f = L|_{B_1} : B_1(0) \rightarrow Y$. Die Abbildung f ist linear, also $f(-x) = -f(x)$, daher ist f ungerade. Wegen $\ker(L) = \{0\}$ gilt $0 \notin f(\partial B_1(0))$. Der Satz von Borsuk liefert

$$d(f, B_1(0), 0) \neq 0.$$

Dann ist f invertierbar auf einer Umgebung der Null. Wegen Linearität ist L dann global invertierbar, $R(L) = Y$. Dies war zu zeigen. \square

10.2. Der Index einer Lösung. Wir betrachten für $G \subset X$, X ein Banachraum über \mathbb{R} , Abbildungen $f \in C^1(G, X)$ und deren Nullstellen. Sei x_0 eine solche Nullstelle. Nun definieren wir den Index der Nullstelle x_0 als die ganze Zahl

$$\text{Index}(f, x_0) := d(f, B_\varepsilon(x_0), 0),$$

wobei ε hinreichend klein ist. Wir nehmen dabei an, dass $A = Df(x_0)$ invertierbar ist; in diesem Fall ist insbesondere obiger Grad definiert. Für einen endlichdimensionalen Raum X (also, nach der Wahl einer Basis, $X = \mathbb{R}^n$), erhalten wir wegen der Summenformel

$$\begin{aligned} \text{Index}(f, x_0) &= d(f, B_\varepsilon(x_0), 0) = d(A, B_\varepsilon(0), 0) = \operatorname{sgn} \det \mathcal{J}_A \\ &= \operatorname{sgn} \det A = (-1)^\beta, \end{aligned}$$

wobei β die Anzahl (mit Vielfachheit) der negativen (also insbesondere reellen) Eigenwerte ist. Insbesondere sehen wir, dass die Definition des Index für zwei verschiedene, aber hinreichend kleine Werte von ε dieselbe Zahl liefert.

Erinnerung: Die (algebraische) Vielfachheit n_λ des Eigenwertes $\lambda \in \mathbb{C}$ des Operators T ist definiert als

$$n_\lambda := \dim \left(\bigcup_{p=1}^{\infty} \ker(\lambda \operatorname{id} - T)^p \right).$$

Dabei ändert sich ab einem p_0 der Kern auf der rechten Seite nicht mehr, es wird also in Wirklichkeit eine endliche Vereinigung gebildet. Zur letzten Zeile in obiger Rechnung bemerken wir, dass tatsächlich Eigenwerte in $\mathbb{C} \setminus \mathbb{R}$ nicht beitragen. Für reelle Matrizen A ist für einen nichtreellen Eigenwert λ auch die konjugiert komplexe Zahl $\bar{\lambda}$ ein Eigenwert, und $\bar{\lambda}$ hat dieselbe Vielfachheit. In der Determinante taucht in diesem Fall also der Term $\lambda^{n_\lambda} \cdot \bar{\lambda}^{n_\lambda}$ auf, welcher positiv ist.

Für einen allgemeinen Banachraum X nehmen wir an, dass

- (1) x_0 regulär ist, also $A := Df(x_0) \in \mathcal{L}(X, X)$ invertierbar
- (2) A Fredholmoperator der Form $A = \operatorname{id} + K$ mit $K \in \mathcal{L}(X)$ kompakt.

SATZ 10.5 (Index). Sei X reeller Banachraum, $G \subset X$, $f \in C^1(G, X)$, und x_0 eine Nullstelle von f . Es gelte (1) und (2). Dann ist der Index von x_0

$$\text{Index}(f, x_0) := d(f, B_\varepsilon(x_0), 0) = (-1)^\beta,$$

wobei

$$\beta = \sum_{\lambda \in \mathbb{R}, \lambda < -1} n_\lambda(K).$$

Bemerkung: Das Theorem verwendet den Spektralsatz für kompakte Operatoren, siehe z.B. [1], 9.6. Dieser liefert für den kompakten Operator K die Endlichkeit $n_\lambda(K) < \infty$ und die Endlichkeit der Anzahl der Eigenwerte in der Menge $\mathbb{C} \setminus B_1(0)$. Daher verwendet die Formel den

Operator K ; die Eigenwerte $\lambda < -1$ von K können aber interpretiert werden als die negativen Eigenwerte von $A = \text{id} + K$.

BEWEIS. *1. Schritt. Übergang zur Ableitung.* Wir betrachten die Homotopie $h(x, t) = (1 - t)f(x) + tAx$. Diese ist zulässig, falls keine Nullstellen auf dem Rand der Kugel $B_\varepsilon(0)$ auftreten. Wir setzen $g(x) := f(x) - Ax$ und rechnen

$$0 = h(x, t) = (1 - t)f(x) + tAx = Ax + (1 - t)g(x),$$

also, wegen Invertierbarkeit von A , $x = (1 - t)A^{-1}g(x)$. Für ε klein genug gilt wegen der Differenzierbarkeit von f dass $\|A^{-1}g(x)\| \leq \varepsilon/2$ für alle $x \in B_\varepsilon(0)$. Insbesondere ist dann h zulässig auf $B_\varepsilon(0)$ und es gilt

$$d(f, B_\varepsilon(0), 0) = d(A, B_\varepsilon(0), 0)$$

2. Schritt. Endlichdimensionale Approximation. Der Spektralsatz [1] 9.6 liefert eine Zerlegung von X in $X = X_1 \oplus X_2$ mit zugehörigen Projektionen π_1, π_2 , wobei X_1 durch die verallgemeinerten Eigenräume von K mit Betrag größer als $1/2$ aufgespannt wird, und X_2 invariant ist. Wir betrachten nun die Homotopie $\text{id} + K_t$ mit $K_t := K \circ \pi_1 + (1 - t)K \circ \pi_2$ zwischen K und dem endlichdimensionalen Operator $K \circ \pi_1$. Wir behaupten, dass die Homotopie zulässig ist. Tatsächlich folgt für Nullstellen x mit $x = x_1 + x_2$, $x_i = \pi_i x$,

$$0 = (\text{id} + K_t)(x) \Rightarrow x_1 + Kx_1 = 0, x_2 + (1 - t)Kx_2 = 0.$$

Wegen Invertierbarkeit von $A = \text{id} + K$ verschwindet dann x_1 , da keine Eigenvektoren x_2 mit Eigenwerten mit Betrag größer 1 existieren, auch x_2 . Wir können also den Grad mit $K \circ \pi_1$ berechnen,

$$d(f, B_\varepsilon(0), 0) = d(\text{id} + K \circ \pi_1, B_\varepsilon(0), 0) = (-1)^\beta.$$

Dabei haben wir das endlichdimensionale Ergebnis verwendet, also

$$\beta = \sum_{\lambda \in \mathbb{R}, \lambda < 0} n_\lambda((\text{id} + K)|_{X_1}) = \sum_{\lambda \in \mathbb{R}, \lambda < -1} n_\lambda(K|_{X_1}) = \sum_{\lambda \in \mathbb{R}, \lambda < -1} n_\lambda(K).$$

□

Teil 3

Verzweigungstheorie

11. Lokale Verzweigung

Wir wollen die Lösungsmenge der Gleichung

$$f(x, \lambda) = 0$$

studieren. Dabei betrachten wir λ als einen Parameter, der in einem physikalischen System zum Beispiel die Temperatur, der Druck oder eine Viskosität sein könnte. Er ist uns im Prinzip vorgegeben. Wir können ihn allerdings auch als Variable ansehen und uns fragen: Wenn wir die Eigenschaften des Systems für den Parameterwert $\lambda = \lambda_0 - \varepsilon$ kennen, was wissen wir dann über die Eigenschaften des Systems für den Parameterwert $\lambda = \lambda_0 + \varepsilon$?

Statt abstrakt von 'Eigenschaften' zu sprechen, wollen wir nun die Anzahl der Nullstellen untersuchen, fragen also nach der Anzahl der Lösungen der Gleichung $f(., \lambda) = 0$ für verschiedene Werte von λ .

Schon mit $\Lambda = \mathbb{R}$ und $X = \mathbb{R}$ sieht man: Alles ist möglich!

- (1) $f(x, \lambda) := x - \lambda$ (Nullstelle lässt sich verfolgen)
- (2) $f(x, \lambda) := (x^2 - \lambda)$ (zwei Nullstellen tauchen auf)
- (3) $f(x, \lambda) := x \cdot (x - \lambda)$ (Variante von 1 mit Verzweigung)
- (4) $f(x, \lambda) := x \cdot (x^2 - \lambda)$ (Variante von 2, zusätzliche Nullstellen)
- (5) $f(x, \lambda) := x \cdot (x^2 - \lambda)(x^4 - \lambda)$ (Pitchfork)

Wir beobachten: Außer in der ersten Gleichung geschieht eine *Verzweigung*. Im Verzweigungspunkt $(x, \lambda) = (0, 0)$ gilt: Neue Lösungen tauchen auf oder Lösungen verschwinden oder Lösungsäste treffen sich.

In der ersten Gleichung ist man in der einfachen Situation, dass man lokal nach x auflösen kann:

$\exists g : \Lambda \rightarrow X$ stetig:

Für jedes λ hat $f(x, \lambda) = 0$ die eindeutige Lösung $x = g(\lambda)$.

Die physikalische Interpretation ist, dass eine Änderung des Parameters (z.B. Temperatur) zwar den Zustand des System ändert, aber diese Änderung ist stetig, vorhersehbar und umkehrbar. Die Verzweigungstheorie beschäftigt sich mit Situationen, in denen dies nicht gilt.

SATZ 11.1 (Notwendige Bedingung für Verzweigung). *Seien X, Λ, Y Banachräume, $f : X \times \Lambda \rightarrow Y$ sei C^1 , $f(x_0, \lambda_0) = 0$. Falls*

$$D_x f(x_0, \lambda_0) : X \rightarrow Y \quad \text{ein Isomorphismus,}$$

so kann man lokal die Gleichung $f(x, \lambda) = 0$ eindeutig durch $x = g(\lambda)$ lösen. (x_0, λ_0) ist dann kein Verzweigungspunkt.

BEWEIS. Dies ist wörtlich eine der Aussagen des Satzes über implizite Funktionen. \square

Wir werden im folgenden davon ausgehen, dass wir eine 'triviale Lösung' kennen, $f(x_0, \lambda_0) = 0$. Dann gilt:

Bemerkung 1: Die Probleme beim Auflösen durch $x = x(\lambda)$ treten auf, obwohl f glatt ist. Das Problem ist nicht durch eine Drehung des (x, λ) -Koordinatensystems behebbar.

Bemerkung 2: Überall bis auf einzelne Punkte kann man lokal durch $x = x(\lambda)$ auflösen (zumindest nach einer Drehung im (x, λ) -Koordinatensystem). Im Fall $\Lambda = \mathbb{R}$ und $X = \mathbb{R}$ und $f \in C^1$ mit $f(x_0, \lambda_0) = 0$ kann man immer lokal auflösen, falls

$$D_x f(x_0, \lambda_0) \neq 0.$$

Falls dies nicht gilt, so entscheiden zweite (höhere) Ableitungen über das lokale Aussehen der Lösungsmenge.

11.1. Ein topologisches Kriterium für eine Verzweigung.
Wir betrachten den Fall, dass ein Ast trivialer Lösungen gegeben ist, also den Fall $f(0, \lambda) = 0$ für alle λ .

SATZ 11.2. *Sei X endlichdimensionaler Banachraum und $\Lambda = \mathbb{R}$, $f : X \times \Lambda \rightarrow X$ stetig differenzierbar mit $f(0, \lambda) = 0$ für alle λ , und $D_x f(0, \lambda)$ invertierbar für $\lambda \neq \lambda_0$. Falls*

$$\begin{aligned} \text{Index}(f(., \lambda), 0) &= \sigma \in \{\pm 1\} & \forall \lambda > \lambda_0, \\ \text{Index}(f(., \lambda), 0) &= -\sigma & \forall \lambda < \lambda_0, \end{aligned}$$

dann ist λ_0 Verzweigungspunkt in dem Sinne, dass für ein $\varepsilon_0 > 0$ gilt

$$(11.1) \quad \forall \varepsilon \in (0, \varepsilon_0) \exists x(\varepsilon), \lambda(\varepsilon) : f(x(\varepsilon), \lambda(\varepsilon)) = 0, \|x(\varepsilon)\| = \varepsilon.$$

Dabei kann $\lambda(\varepsilon)$ so gewählt werden, dass $\lambda(\varepsilon) \rightarrow \lambda_0$ für $\varepsilon \rightarrow 0$.

BEWEIS. Wir wählen eine Folge $\mu_k \searrow 0$ und betrachten die Parameterwerte $\lambda_0 \pm \mu_k$. Für jedes k gibt es Radien $\varepsilon_k > 0$, so dass die Abbildungsgrade $d(f(., \lambda_0 \pm \mu_k), B_\varepsilon(0), 0)$ unabhängig von ε_k mit dem Index übereinstimmen. Insbesondere gilt für alle $\varepsilon < \varepsilon_k$

$$d(f(., \lambda_0 + \mu_k), B_\varepsilon(0), 0) \neq d(f(., \lambda_0 - \mu_k), B_\varepsilon(0), 0).$$

Wir können ε_k monoton fallen wählen. Für $\varepsilon \in (\varepsilon_{k+1}, \varepsilon_k]$ wählen wir $\lambda_- = \lambda_0 - \mu_k$ und $\lambda_+ = \lambda_0 + \mu_k$. Für die Homotopie $f(x, \lambda)$ mit $\lambda \in [\lambda_-, \lambda_+]$ ist der Abbildungsgrad in den Endpunkten verschieden, die Homotopie kann also nicht zulässig sein. Wir folgern

$$\exists x \in \partial B_\varepsilon(0), \lambda \in [\lambda_-, \lambda_+] : f(x, \lambda) = 0.$$

Dies liefert die Behauptung. \square

11.2. Der Satz von Krasnoselskii. Ziel: Finde eine hinreichende Bedingung für eine Verzweigung.

Sei X ein Banachraum und $f : X \times \mathbb{R} \rightarrow X$ von der Form

$$f(x, \lambda) = x - (\mu_0 + \lambda)Tx + g(x, \lambda).$$

Wir machen folgende Annahmen:

- (1) $\mu_0 \neq 0$,
- (2) $T : X \rightarrow X$ ist linear und kompakt,

- (3) g ist eine nichtlineare, kompakte Abbildung $X \times \mathbb{R} \rightarrow X$,
- (4) $g(0, \lambda) = 0$ für alle λ und $g(x, \lambda) = o(\|x\|)$ gleichmäßig in $\lambda \in (-\varepsilon, \varepsilon)$.

Bemerkung: (1)-(4) sind z.B. erfüllt, falls $f(x, \lambda) = \text{id} - (\mu_0 + \lambda)\tilde{T}(x)$, mit \tilde{T} nichtlinear, kompakt und C^1 .

Falls $f \in C^1$, so ist nach Satz 11.1 die *notwendige* Bedingung für eine Verzweigung

$$D_x f(0, 0) = \text{id} - \mu_0 T \quad \text{nicht invertierbar.}$$

äquivalent dazu: $1/\mu_0$ ist ein Spektralwert von T . Wegen Kompaktheit von T äquivalent: $1/\mu_0$ ist ein Eigenwert von T . Wir haben allerdings nicht vorausgesetzt, dass $g \in C^1$.

DEFINITION 11.3. Falls ein trivialer Ast gegeben ist, d.h. $f(0, \lambda) = 0 \forall \lambda$, dann sagen wir:

$(0, \lambda_0)$ ist Verzweigungspunkt, falls in jeder Umgebung von $(0, \lambda_0)$ nichttriviale Lösungen von $f = 0$ liegen.

Behauptung: In obigem Fall ist $\text{id} - \mu_0 T$ nicht invertierbar notwendige Voraussetzung für eine Verzweigung.

Beweis: Wir nehmen an, dass $\text{id} - \mu_0 T$ invertierbar ist. Eine nichttriviale Lösung (x, λ) erfüllt

$$0 = f(x, \lambda) = x - (\mu_0 + \lambda)Tx + g(x, \lambda),$$

also

$$x = (\text{id} - \mu_0 T)^{-1}[\lambda Tx - g(x, \lambda)].$$

In einer Umgebung von $(0, 0)$ erreichen wir, dass die rechte Seite beschränkt ist durch $\frac{1}{2}\|x\|$. Ein Widerspruch.

SATZ 11.4 (Krasnoselskii). Es gelte (1)-(4) und $1/\mu_0$ sei ein Eigenwert von T mit ungerader (algebraischer) Vielfachheit. Dann ist $(0, 0)$ ein Verzweigungspunkt für f .

BEWEIS. Annahme: $(0, 0)$ ist kein Verzweigungspunkt. Dann sind in $B_\varepsilon(0) \times (-\varepsilon, \varepsilon)$ keine nichttrivialen Lösungen von $f = 0$. Insbesondere ist der Index

$$d(f(., \lambda), B_\varepsilon, 0) \in \mathbb{Z}$$

definiert. Wegen der Homotopieeigenschaft ist der Index unabhängig von $\lambda \in (-\varepsilon, \varepsilon)$. Dies wird zu einem Widerspruch führen.

Um im Folgenden den Index zu berechnen, betrachten wir die Homotopie

$$x - (\mu_0 + \lambda)Tx + \frac{1}{t}g(tx, \lambda),$$

die f nach $\text{id} - (\mu_0 + \lambda)T$ deformiert, und zulässig ist, da es sonst nichttriviale Lösungen in $B_\varepsilon(0)$ gäbe. Nach dem Satz von Leray-Schauder, Satz 10.5, gilt für $\lambda \in (-\varepsilon, \varepsilon)$

$$d = (-1)^{\beta(\lambda)}, \quad \beta(\lambda) = \sum_{\sigma > 1} n_\sigma((\mu_0 + \lambda)T).$$

Es gilt

$$\begin{aligned} & (\mu_0 + \lambda)Tx = \sigma x, \quad \sigma > 1 \\ \iff & Tx = \frac{\sigma}{\mu_0 + \lambda}x, \quad \sigma > 1 \\ \iff & Tx = \bar{\sigma}x, \quad \bar{\sigma} > \frac{1}{\mu_0 + \lambda}. \end{aligned}$$

Daher gilt

$$\beta(\lambda) = \sum_{\bar{\sigma} > (\mu_0 + \lambda)^{-1}} n_{\bar{\sigma}}(T).$$

Für $\varepsilon > 0$ klein genug und $-\varepsilon < \lambda_- < 0 < \lambda_+ < \varepsilon$ gilt dann

$$\beta(\lambda_+) - \beta(\lambda_-) = n_0(T) \in 2\mathbb{Z} + 1,$$

also

$$d = (-1)^{\beta(\lambda_-)} = -(-1)^{\beta(\lambda_+)} = -d,$$

ein Widerspruch wegen $d \in \{\pm 1\}$. \square

Wir wollen mit einem Beispiel zeigen, dass im Falle gerader Vielfachheit tatsächlich keine Verzweigung stattfinden muss.

Beispiel. Zu $x = (x_1, x_2) \in X := \mathbb{R}^2$ betrachten wir

$$f(x, \lambda) := -\lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} -x_2^3 \\ x_1^3 \end{pmatrix} = 0.$$

Dann gelten (1)-(4) mit $T = \text{id}$ und $\mu_0 = 1$. Dann ist μ_0 ein Eigenwert von T mit Vielfachheit 2.

Angenommen, wir hätten eine Lösung (x, λ) . Dann gilt

$$0 = f(x, \lambda) \cdot \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = x_2^4 + x_1^4.$$

Also $x = (x_1, x_2) = (0, 0)$. Es gibt also nur die trivialen Lösungen, $\mu_0 = 1$ ist kein Verzweigungspunkt.

11.3. Lokale Verzweigung im einfachen Eigenwert. In diesem Abschnitt werden wir eine weitere hinreichende Bedingung für eine Verzweigung kennenlernen, die mit dem Satz über implizite Funktionen arbeitet. Da $D_x f = 0$ im Verzweigungspunkt, werden wir etwas über zweite Ableitungen voraussetzen. Als Erinnerung an Bezeichnungen: Für $f : X \times Y \rightarrow Z$ gilt $D_x f(x, y) : X \rightarrow Z$, wir schreiben $D_x f(x, y) \langle \bar{x} \rangle = D_x f(x, y) \bar{x}$. Die zweite Ableitung ist eine Abbildung (z.B.) $D_y D_x f(x, y) : Y \rightarrow \mathcal{L}(X, Z)$. Wir identifizieren mit $D_y D_x f(x, y) : X \times Y \rightarrow Z$ und schreiben $D_y D_x f(x, y) \langle \bar{x}, \bar{y} \rangle$.

SATZ 11.5 (Verzweigung im einfachen Eigenwert). *Sei $\Lambda = \mathbb{R}$, X, Y Banachräume, $f : X \times \Lambda \rightarrow Y$ sei C^2 . Wir setzen voraus:*

(1) trivialer Ast:

$$f(0, \lambda) = 0 \quad \forall \lambda \in \Lambda.$$

(2) Fredholm Eigenschaft:

$$L(\lambda) := D_x f(0, \lambda) : X \rightarrow Y$$

ist ein Fredholm Operator mit Index 0.

(3) notwendige Bed. und einfacher Eigenwert:

$$X_0 := \ker L(0) = \mathbb{R}x_0.$$

(4) Transversalitat:

$$D_\lambda L(0)x_0 \notin R(L(0)).$$

Dann kann man lokal mit zwei Zweigen auflosen: Fur Umgebungen $(0, 0, 0) \in U \times V \times \Sigma \subset X_0 \times X_1 \times \Lambda$, $X = X_0 \oplus X_1$ gibt es eindeutige $\tilde{v} : U \rightarrow V$, $\tilde{v}(0) = 0$, $\tilde{\lambda} : U \rightarrow \Sigma$, $\tilde{\lambda}(0) = 0$, so dass

$$f(x_0 + x_1, \lambda) = 0 \iff \begin{cases} x = x_0 + x_1 = 0 & \text{oder} \\ x_1 = \tilde{v}(x_0), \lambda = \tilde{\lambda}(x_0) \end{cases}$$

Falls $f \in C^k$ gilt $\tilde{v}, \tilde{\lambda} \in C^{k-1}$.

Das Bild ist wie in den ersten beiden mathematischen Beispielen; die Losungsmenge ist die Vereinigung zweier aste, des trivialen Astes mit einem Ast, der uber x_0 parametrisiert wird (und nicht uber λ).

Zum Namen '... im einfachen Eigenwert': Betrachte $f(x, \lambda) = \lambda x - Tx$ fur einen nichtlinearen kompakten Operator $T : X \rightarrow X$ mit $T(0) = 0$ (wir setzen $Y = X$). Sei $\tilde{T} := D_x T(0)$. Wir wollen die vier Voraussetzungen fur dieses f uberprufen.

- (1) ist immer erfullt.
- (2) fordert, dass $D_x f(0, 0) = \lambda \text{id} - \tilde{T}$ ein Fredholmoperator mit Index 0 ist. Dies ist erfullt nach Lemma 5.4 und Satz 6.2.
- (3) fordert, dass $X_0 = \ker(\lambda \text{id} - \tilde{T}) = \mathbb{R}x_0$. Dies ist gerade die Bedingung, dass λ ein Eigenwert ist mit geometrischer Vielfachheit 1.
- (4) Wegen $D_\lambda L(0) = \text{id}$ fordert man hier, dass $x_0 \notin R(\lambda \text{id} - \tilde{T})$. Dies bedeutet, dass λ ein algebraisch einfacher Eigenwert sein soll.

BEWEIS. Der Beweis erfolgt in 2 Schritten; der erste Schritt reduziert das Problem auf eine Dimension, der Kern des Beweises geschieht im zweiten Schritt.

1. Schritt. Wir fuhren eine Ljapunov-Schmidt Reduktion bezuglich $L(0)$ durch. Wir zerlegen $Y = Y_0 \oplus Y_1$ mit $Y_1 := R(L(0))$ und verwenden die Projektion $Q : Y \rightarrow Y$ auf Y_0 mit Kern Y_1 . Dann gilt lokal fur eine Auflofungsfunktion $v : X_0 \times \Lambda \rightarrow X_1$

$$f(x, \lambda) = 0 \iff Qf(x_0 + v(x_0, \lambda), \lambda) = 0.$$

Wir setzen jetzt

$$\Phi(r, \lambda) := Qf(rx_0 + v(rx_0, \lambda), \lambda),$$

$\Phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit der offensichtlichen Identifikation von $\mathbb{R}x_0$ mit \mathbb{R} . Damit ist das Problem auf ein eindimensionales reduziert. Es bleibt zu zeigen, dass Φ wieder die Voraussetzungen (1)-(4) erfullt.

Ad (1). $Qf(0 + v(0, \lambda), \lambda) = 0$, da $v(0, \lambda) = 0$ ($x_1 = v$ ist die eindeutige Lösungsfunktion der Qf -Gleichung und $x_1 = 0$ ist sicherlich eine Lösung der Qf -Gleichung, falls $x_0 = 0$).

Ad (2). $D_r\Phi(\lambda) : \mathbb{R} \rightarrow \mathbb{R}$ ist eine lineare Abbildung zwischen endlichdimensionalen Räumen und daher immer Fredholm mit Index 0.

Ad (3). Dies ist die Bemerkung nach Satz 4.9. $D_r\Phi(0, 0) = QD_x f(0, 0) \langle x_0 + D_{x_0}v \cdot x_0 \rangle = 0$, da $D_x f(0, 0)$ nach $R(L(0))$ abbildet.

Ad (4). Wir berechnen

$$\begin{aligned} & D_\lambda D_r\Phi(0, 0) \\ &= QD_\lambda [D_x f(rx_0 + v(rx_0, \lambda), \lambda) \langle x_0 + D_{x_0}v(rx_0, \lambda) \cdot x_0 \rangle] (0, 0) \\ &= QD_x^2 f(0, 0) \langle x_0 + D_{x_0}v(0, 0) \cdot x_0, D_\lambda v(0, 0) \rangle \\ &\quad + QD_\lambda D_x f(0, 0) \langle x_0 + D_{x_0}v(0, 0) \cdot x_0, 1 \rangle \\ &\quad + QD_x f(0, 0) \langle D_\lambda D_{x_0}v(0, 0) \langle x_0, 1 \rangle \rangle. \end{aligned}$$

Der dritte Term verschwindet, weil $1 - Q$ auf ein Element in $R(L(0))$ angewendet wird. Der erste Term verschwindet auch: Wegen $v(0, \lambda) = 0$ (s.o.) gilt $D_\lambda v(0, 0) = 0$. Wir müssen zeigen, dass der zweite Term nicht verschwindet. In Satz 4.9 haben wir $D_{x_0}v(0, 0) = 0$ gezeigt. Daher gilt

$$D_\lambda D_r\Phi(0, 0) = QD_\lambda D_x f(0, 0) \langle x_0, 1 \rangle,$$

dies ist nach Voraussetzung (4) ungleich Null.

Wenn man $\Phi = 0$ durch $\lambda = \tilde{\lambda}(x_0)$ lösen kann, dann folgt der Satz mit $x_1 = \tilde{v}(x_0) := v(x_0, \tilde{\lambda}(x_0))$.

2. Schritt. Wir zeigen nun den Satz für $X = Z = \mathbb{R}$. Wir wollen die nichttrivialen Lösungen von $f(x, \lambda) = 0$ finden. Idee: Die trivialen Lösungen werden 'herausdividiert'. Wir setzen

$$\psi(x, \lambda) := \begin{cases} \frac{1}{x} f(x, \lambda) & \text{für } x \neq 0, \\ D_x f(0, \lambda) & \text{für } x = 0. \end{cases}$$

Lokal gilt für Punkte (x, λ) mit $x \neq 0$:

$$f(x, \lambda) = 0 \iff \psi(x, \lambda) = 0,$$

also eine Äquivalenz der beiden Gleichungen. Wir werden die ψ -Gleichung lösen.

Es gilt $\psi \in C^1$, da $f(0, \lambda) = 0$ und $f \in C^2$. Weiterhin $\psi(0, 0) = D_x f(0, 0) = 0$ und

$$D_\lambda \psi(0, 0) = D_\lambda D_x f(0, 0) \neq 0.$$

Deswegen können wir $\psi = 0$ lokal mit dem Satz über implizite Funktionen auflösen. Wir finden einen Weg $\lambda = \tilde{\lambda}(x)$, so dass lokal gilt

$$\psi(x, \lambda) = 0 \iff \lambda = \tilde{\lambda}(x).$$

$\tilde{\lambda}$ ist die gesuchte Funktion. \square

Beispiel 1: Eine elementare Funktion. Für $x \in \mathbb{R}^3$ wollen wir die Gleichung

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} x + f(x) = 0$$

untersuchen, wobei $f \in C^2(\mathbb{R}^3, \mathbb{R}^3)$, $f(0) = 0$, $Df(0) = 0$. Der Satz liefert die Existenz eines nichttrivialen Astes, der in $(x, \lambda) = (0, 0)$ aus dem trivialen Ast verzweigt. Zu den Voraussetzungen:

- (1) erfüllt wegen $f(0) = 0$
- (2) erfüllt, da $X = \mathbb{R}^3$ endlichdimensional
- (3) erfüllt, denn

$$\ker D_x f(0, 0) = \ker \begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{R}e_1$$

(4) erfüllt, denn es gilt

$$\begin{aligned} D_\lambda D_x f(0, 0) \langle e_1 \rangle &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot e_1 = e_1 \\ &\notin R \left(\begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \right) = \mathbb{R}e_2 + \mathbb{R}e_3. \end{aligned}$$

Beispiel 2: Transversalität. Wir betrachten den allgemeinen endlichdimensionalen Fall $f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ von der Klasse C^2 mit trivialem Ast ($f(0, \lambda) = 0 \forall \lambda$). Wir nehmen an, dass die Linearisierung $L(\lambda) := D_x f(0, \lambda)$ die Eigenwerte

$$\mu_1(\lambda) \leq \mu_2(\lambda) \leq \dots < \mu_k(\lambda) < \dots \leq \mu_n(\lambda)$$

besitzt, und dass diese differenzierbar von λ abhängen. Weiterhin gelte im kritischen Punkt λ_0 für den k -ten Eigenwert $\mu_k(\lambda_0) = 0$, der Eigenwert sei algebraisch einfach, und Transversalität gelte im dem Sinne, dass

$$\partial_\lambda \mu_k(\lambda_0) \neq 0.$$

Dann ist in $(x, \lambda) = (0, \lambda_0)$ obiger Satz anwendbar, d.h. es gibt einen nichttrivialen Ast. Zu den Voraussetzungen:

- (1) und (2) sind erfüllt
- (3) erfüllt, denn nach Voraussetzung ist $\mu_k(\lambda_0)$ geometrisch einfach. (4) Um etwas über $D_\lambda D_x f(0, \lambda_0)$ aussagen zu können, können wir nur von der Eigenwertgleichung ausgehen:

$$\mu_k(\lambda) x_0(\lambda) = L(\lambda) x_0(\lambda).$$

Wir differenzieren nach λ ,

$$\partial_\lambda \mu_k(\lambda_0) x_0(\lambda_0) + \mu_k(\lambda_0) \partial_\lambda x_0(\lambda_0) = D_\lambda L(\lambda_0) x_0(\lambda_0) + L(\lambda_0) \partial_\lambda x_0(\lambda_0).$$

Der erste Term ist von der Form αx_0 mit $\alpha \neq 0$, der zweite Term verschwindet wegen $\mu_k(\lambda_0) = 0$, der dritte Term wird von uns benötigt, der vierte ist sicherlich im Bild von $L(\lambda_0)$. Wir erhalten (4), nämlich

$$D_\lambda L(\lambda_0)x_0 \notin R(L(0)),$$

falls αx_0 nicht im Bild $R(L(0))$ ist. Aber x_0 ist nicht in $R(L(0))$ wegen der algebraischen Einfachheit des Eigenwertes. Wir sehen, dass algebraische Einfachheit zusammen mit der Transversalität des Eigenwertes die Voraussetzungen liefert.

Beispiel 3: Euler-Stab. Ein mit Kraft λ belasteter Stab der Länge L und mit gleichmäßiger Dicke wird beschrieben durch die Gleichung

$$(11.2) \quad \begin{aligned} x''(s) + \lambda \rho x(s) \sqrt{1 - |x'(s)|^2} &= 0 & \forall s \in (0, L), \\ x(0) &= x(L) = 0. \end{aligned}$$

(sh. Deimling [4]). Dabei ist s die Bogenlänge, x die (Quer-) Auslenkung des Stabes, $\rho > 0$ der Elastizitätskoeffizient, der Einfachheit nehmen wir hier ein konstantes ρ an.

Um obige Gleichungen mathematisch zu fassen, setzen wir

$$X := \{x \in C^2([0, L], \mathbb{R}) : x(0) = x(1) = 0\},$$

$$Y := C^0([0, L], \mathbb{R}),$$

$$\Lambda := \mathbb{R},$$

$$f(x, \lambda) := x'' + \lambda \rho x \sqrt{1 - |x'(s)|^2}.$$

Die Funktion f ist C^2 in einer Umgebung von $0 \in X$, denn dort gilt $\|x'\|_\infty \leq 1/2$. Wir prüfen die Voraussetzungen des Satzes nach:

(1) $f(0, \lambda) = 0$ ist erfüllt für alle $\lambda \in \mathbb{R}$.

(2) Der Operator $L(\lambda)$ ist

$$L(\lambda) = D_x f(0, \lambda) : x \mapsto x'' + \lambda \rho x.$$

Also ist $L(\lambda)$ ein Fredholm Operator vom Index 0 ($X \ni x \mapsto x'' \in Y$ ist Fredholm Operator vom Index 0, $L(\lambda)$ ist eine kompakte Störung davon).

(3) Die Eigenwerte des Laplace Operators sind einfach: Zum Eigenwert $-k^2\pi^2/L^2$ gehört die Eigenfunktion $u_k(s) = \sin(k\pi s/L)$. Die möglichen Verzweigungspunkte sind

$$\lambda_k = k^2 \frac{\pi^2}{L^2 \rho}, \quad k \in \mathbb{N}.$$

(4) Es gilt $D_\lambda L(\lambda) = \rho \text{id}$, also

$$D_\lambda L(\lambda) u_n = \rho u_n \notin R(L(\lambda)) = \ker(L(\lambda))^\perp.$$

Hierbei verwenden wir, dass $L(\lambda)$ selbstadjungiert in L^2 ist.

Wir finden das Verzweigungsbild der Stabgleichung in der Umgebung von $x = 0$: In den Werten λ_k findet eine Verzweigung statt, außer der trivialen Lösungen gibt es lokal um Punkte $(x, \lambda) = (0, \lambda_k)$ nicht-triviale eindimensionale Mannigfaltigkeiten von Lösungen. Die Zweige

sind symmetrisch: Wegen $f(x, \lambda) = f(-x, \lambda)$ ist mit (x, λ) auch $(-x, \lambda)$ eine Lösung. Wegen der lokalen Eindeutigkeit der Zweige gilt

$$\tilde{v}(-x_0) = -\tilde{v}(x_0), \quad \tilde{\lambda}(-x_0) = \tilde{\lambda}(x_0).$$

12. Stabilität der stationären Zweige

Oft kann die Gleichung $f(x) = 0$ als stationäre Gleichung zur folgenden instationären Gleichung angesehen werden

$$\frac{d}{dt}x = f(x). \quad (*)$$

Falls uns eine solche Gleichung gegeben ist, dann sind folgende Bezeichnungen sinnvoll:

- x_0 stationäre Lösung : $\iff f(x_0) = 0$.
- (stat. Lsg.) x_0 linear stabil : $\iff \exists c > 0$: alle Spektralwerte λ von $D_x f(x_0)$ erfüllen $\operatorname{Re}(\lambda) \leq -c$.
- (stat. Lsg.) x_0 linear instabil : \iff es gibt einen EW λ mit $\operatorname{Re}(\lambda) > 0$.

Wir haben diese Begriffe im Prolog untersucht und gesehen, dass linear stabile (instabile) stationäre Lösungen auch im anschaulichen Sinne stabil (instabil) sind:

- x_0 linear stabil $\Rightarrow \exists \varepsilon > 0$: alle Lösungen $x(t)$ von $(*)$ erfüllen $x(t) \rightarrow x_0$ für $t \rightarrow \infty$.
- x_0 linear instabil $\Rightarrow \exists \delta > 0 \forall \varepsilon > 0 \exists$ Lösung $x(t)$ von $(*)$ und t mit $x(t) \notin B_\delta(x_0)$.

Betrachte jetzt die Situation von Satz 11.5. Gegeben ein Verzweigungsdiagramm und eine Stabilitätsaussage für einen der 4 Teiläste wollen wir die lineare Stabilität der anderen 3 Teiläste bestimmen. Nach Voraussetzung hat $D_x f(0, 0)$ einen einfachen Eigenwert 0. Falls alle anderen Eigenwerte negativen Realteil haben, so entscheidet über die Stabilität eines Astes $(x(s), \lambda(s))$ das Vorzeichen von $\tilde{\mu}(s)$ mit

$$D_x f(x(s), \lambda(s)) \tilde{u}(s) = \tilde{\mu}(s) \tilde{u}(s), \quad \tilde{u}(s) \neq 0.$$

Wir müssen zeigen, dass wir einfache Eigenwerte bei einer Störung des Operators wie angedeutet fortsetzen können.

DEFINITION 12.1. Seien $T_0, K : X \rightarrow Y$ linear und stetig. μ_0 heißt K -einfacher Eigenwert von T_0 , falls

- (1) $T_0 - \mu_0 K$ Fredholm-Operator mit Index 0.
- (2) $\ker(T_0 - \mu_0 K) = \mathbb{R}x_0$.
- (3) $Kx_0 \notin R(T_0 - \mu_0 K)$.

Bem: Für $X = Y$, $K = \operatorname{id}$ und T_0 kompakt ist ein id-einfacher Eigenwert einfach ein einfacher Eigenwert.

LEMMA 12.2 (Fortsetzung von einfachen Eigenwerten). *Sei μ_0 ein K -einfacher Eigenwert von T_0 mit Eigenvektor x_0 . Dann gibt es für ein $\delta > 0$ eine Fortsetzung des Eigenwertes,*

$$\mu : \{T \in \mathcal{L}(X, Y) \mid \|T - T_0\| < \delta\} \rightarrow \mathbb{R} \quad \text{stetig}$$

mit $\mu(T_0) = \mu_0$ und $\mu(T)$ ist K -einfacher Eigenwert von T . Diese Fortsetzung μ ist eindeutig. Weiterhin gilt

$$\ker(T - \mu(T)K) = \mathbb{R}x(T), \quad x(T) = x_0 + x_1(T) \in \mathbb{R}x_0 \oplus X_1 = X.$$

μ und x sind analytisch in T .

BEWEIS. o.E. $\mu_0 = 0$. Wir wollen $(T - \mu(T)K)x(T) = 0$ lösen. Dazu definieren wir auf einer Umgebung von $(0, 0, T_0) \in \mathbb{R} \times X_1 \times \mathcal{L}(X, Y)$ die Funktion

$$F(r, x_1, T) := (T - rK)(x_0 + x_1) \in Y.$$

Hier wird T als ein Parameter von F aufgefasst.

Wir berechnen die Ableitung und zerlegen dazu $Y = Y_0 \oplus Y_1 = Y_0 \oplus R(T_0)$, $\dim Y_0 = 1$, mit Projektionen P_1 und $P_0 := \text{id} - P_1$.

$$\begin{aligned} D_{(r, x_1)} F(0, 0, T_0) : (\bar{r}, \bar{x}_1) &\mapsto -\bar{r}Kx_0 + T_0\bar{x}_1 \\ &= (-\bar{r}P_0(Kx_0), P_1(-\bar{r}Kx_0 + T_0\bar{x}_1)). \end{aligned}$$

Da $P_0(Kx_0) \neq 0$ und $P_1 \circ T_0 : X_1 \rightarrow Y_1$ invertierbar, ist auch $D_{(r, x_1)} F(0, 0, T_0)$ invertierbar. Nach dem Satz über implizite Funktionen finden wir $\mu(T)$ und $x(T)$ für alle T nahe T_0 .

Es bleibt zu zeigen: a) Eindeutigkeit. b) Eigenschaften (1)-(3) der Definition.

a) Wir haben mit $x = x_0 + x_1$ einen speziellen Ansatz gewählt. Für einen anderen Ansatz bzw. für x_1 groß könnte es weitere Lösungen geben. Wir nehmen jetzt an, dass $x = \beta x_0 + x_1$ eine normierte Lösung ist, also

$$(T - rK)(\beta x_0 + x_1) = 0.$$

Dies schreiben wir als

$$T_0 x_1 - r\beta K x_0 = (T_0 - T)(\beta x_0 + x_1) + rK x_1.$$

Die Terme links erfüllen $T_0 x_1 \in R(T_0)$ und $r\beta K x_0 \notin R(T_0)$. Daher gilt eine Abschätzung

$$\|x_1\| + |r\beta| \leq C \{\|T_0 - T\|(|\beta| + \|x_1\|) + |r|\|x_1\|\}.$$

Falls $\|T_0 - T\|$ und r klein sind können wir die x_1 -Terme in die linke Seite absorbieren und finden

$$\|x_1\| + |r\beta| \leq C' \|T_0 - T\| |\beta|.$$

Da β beschränkt ist (Normierung) liefert dies die Kleinheit von x_1 , und dies (wieder Normierung) die Kleinheit von $1 - \beta$. Insbesondere war unser Ansatz $\beta = 1$ keine Einschränkung. Für kleines x_1 liefert der Satz über implizite Funktionen die Eindeutigkeit.

b) Die Eindeutigkeit von x_1 liefert insbesondere Eigenschaft (2).

Für (3) müssen wir zeigen, dass

$$(12.1) \quad K(x_0 + x_1(T)) \notin R(T - \mu(T)K).$$

Die Abbildungen

$$\hat{T} : (x_1, r) \mapsto (T - \mu(T)K)x_1 + rK(x_0 + x_1(T)) \in Y$$

sind invertierbar für T nahe T_0 , weil \hat{T}_0 invertierbar ist. Dies liefert (12.1).

Eigenschaft (1) folgt aus der nachfolgenden allgemeinen Feststellung.

Beh.: Sei $T_0 : X \rightarrow Y$ Fredholm Operator. Für $\|T - T_0\|$ hinreichend klein gilt: Auch T ist Fredholmoperator und es gilt

$$\dim \ker T \leq \dim \ker T_0, \quad \text{codim } R(T) \leq \text{codim } R(T_0).$$

Falls die Dimensionen auf der rechten Seite jeweils 1 sind, so ist T ebenfalls Fredholm Operator mit Index 0.

Beweis der Beh.: Für den Fredholm Operator $T_0 : X \rightarrow Y$ zerlegen wir wie in Lemma 6.4 $X = X_0 \oplus X_1$, $Y = Y_0 \oplus Y_1$, $T_0|_{X_1} : X_1 \rightarrow Y_1$ invertierbar, $\tilde{T}_0 : X_1 \times Y_0 \rightarrow Y$, $(x_1, y_0) \mapsto T_0 x_1 + y_0$ ebenfalls invertierbar. Für Operatoren T mit $\|T - T_0\|$ klein ist auch die kleine Störung $\tilde{T} : (x_1, y_0) \mapsto T x_1 + y_0$ invertierbar. Also gilt $\ker T \cap X_1 = \{0\}$ und $R(T) + Y_0 = Y$, insbesondere die Fredholm Eigenschaft und die Aussage über die Dimensionen.

Für den Index müssen wir zeigen $R(T) = Y \Rightarrow \ker(T) = \{0\}$. Im Falle $T(X_1) = Y$ wäre \tilde{T} nicht invertierbar (Lösung nicht eindeutig). Also $T(X_1) \neq Y$ und daher $T x_0 \notin T(X_1)$, da sonst T nicht surjektiv ist. Dies impliziert aber $\ker(T) = \{0\}$. \square

Situation: Wir nehmen an, dass $X \subset Y$ und setzen $K = j$, die Inklusion. Weiterhin sei die Situation von Satz 11.5 gegeben und wir haben zwei Äste $(0, \lambda), \lambda \in (-\varepsilon_1, \varepsilon_1)$ und $(x(s), \lambda(s)), s \in (-\varepsilon_2, \varepsilon_2)$, $x(s) = sx_0 + x_1(s)$. Wir wollen die Eigenwerte der Linearisierung auf den Ästen betrachten. Nach Lemma 12.2 gibt es $\bar{u}, \bar{\mu}, \tilde{u}, \tilde{\mu}$ mit

$$\begin{aligned} D_x f(0, \lambda) \bar{u}(\lambda) &= \bar{\mu}(\lambda) \bar{u}(\lambda), \\ D_x f(x(s), \lambda(s)) \tilde{u}(s) &= \tilde{\mu}(s) \tilde{u}(s). \end{aligned}$$

Unser Ziel ist es, die Vorzeichen von $\bar{\mu}$ und $\tilde{\mu}$ zu bestimmen.

SATZ 12.3 (Crandall-Rabinowitz). *In obiger Situation gilt $\bar{\mu}'(0) \neq 0$, $s\lambda'(s) \rightarrow 0$ und $\tilde{\mu}(s) \rightarrow 0$ für $s \rightarrow 0$ verschwinden mit entgegengesetztem Vorzeichen. Genauer: Für jede Folge $s_n \rightarrow 0$ mit $\tilde{\mu}(s_n) \neq 0$ gilt der Limes*

$$(12.2) \quad \frac{s\lambda'(s)\bar{\mu}'(0)}{\tilde{\mu}(s)} \rightarrow -1 \quad \text{für } s \rightarrow 0.$$

BEWEIS. **a) trivialer Zweig:** Wir benutzen die Charakterisierung der Eigenwertfortsetzung auf dem trivialen Zweig,

$$D_x f(0, \lambda) \bar{u}(\lambda) = \bar{\mu}(\lambda) \bar{u}(\lambda),$$

und differenzieren in $\lambda = 0$ nach λ . Dies liefert

$$D_\lambda D_x f(0, 0) \bar{u}(0) + D_x f(0, 0) \partial_\lambda \bar{u}(0) = \partial_\lambda \bar{\mu}(0) \bar{u}(0) + \bar{\mu}(0) \partial_\lambda \bar{u}(0).$$

Wegen $D_\lambda D_x f(0, 0) \bar{u}(0) = D_\lambda L(0)x_0$ und $D_x f(0, 0) \partial_\lambda \bar{u}(0) \in R(L(0))$ und Voraussetzung 4 in Satz 11.5 ist die linke Seite ungleich 0. Auf der rechten Seite gilt $\bar{\mu}(0) = 0$ und wir schließen daher $\partial_\lambda \bar{\mu}(0) \neq 0$.

Wir werden die obige differenzierte Eigenwertgleichung nochmals verwenden. Dazu wollen wir sie vereinfachen: Mit einem linearen, stetigen Funktional $y^* : Y \rightarrow \mathbb{R}$ mit $\ker y^* = R(L(0))$ und $y^* \neq 0$ gilt

$$(12.3) \quad \langle y^*, D_\lambda D_x f(0, 0) \bar{u}(0) \rangle = \partial_\lambda \bar{\mu}(0) \langle y^*, \bar{u}(0) \rangle.$$

b) nichttrivialer Zweig: Wir haben zwei Gleichungen zur Verfügung. Die eine ist die Charakterisierung der Eigenwertfortsetzung

$$D_x f(x(s), \lambda(s)) \tilde{u}(s) = \tilde{\mu}(s) \tilde{u}(s).$$

Die zweite ist die Gleichung selbst, $f(x(s), \lambda(s)) = 0$, die wir nach s differenzieren:

$$D_x f(x(s), \lambda(s)) \partial_s x(s) + D_\lambda f(x(s), \lambda(s)) \partial_s \lambda(s) = 0.$$

Wir subtrahieren die beiden Gleichungen voneinander und erhalten

$$(12.4) \quad \begin{aligned} D_x f(x(s), \lambda(s)) (\partial_s x(s) - \tilde{u}(s)) + D_\lambda f(x(s), \lambda(s)) \partial_s \lambda(s) \\ + \tilde{\mu}(s) \tilde{u}(s) = 0. \end{aligned}$$

Der Beweis besteht nun im Wesentlichen darin, in Gleichung (12.4) $D_x f$, $D_\lambda f$ und \tilde{u} in einer Taylor-Reihe zu entwickeln.

$$(12.5) \quad \begin{aligned} & D_x f(0, 0) (\partial_s x(s) - \tilde{u}(s)) + o(1) (\partial_s x(s) - \tilde{u}(s)) \\ & + D_x D_\lambda f(0, 0) \langle \partial_s x(0) s \rangle \partial_s \lambda(s) + D_\lambda D_\lambda f(0, 0) \langle \partial_s \lambda(0) s \rangle \partial_s \lambda(s) \\ & + o(s) \partial_s \lambda(s) + \tilde{\mu}(s) \tilde{u}(0) + \tilde{\mu}(s) o(1) = 0. \end{aligned}$$

Wir benutzen nun $x(s) = sx_0 + x_1(s)$ aus Satz 11.5, und $\tilde{u}(s) = x_0 + u_1(s)$ aus Lemma 12.2. Sie liefern $\partial_s x(s) - \tilde{u}(s) \in Y_1 = R(L(0))$. Da $L(0)|_{X_1} : X_1 \rightarrow R(L(0))$ invertierbar ist, folgt aus (12.5)

$$(12.6) \quad \| \partial_s x(s) - \tilde{u}(s) \| \leq C (|s \partial_s \lambda(s)| + |\tilde{\mu}(s)|).$$

Wir projizieren nun (12.5) mit Hilfe von y^* und finden wegen $\langle y^*, D_x f(0, 0) \dots \rangle = 0$ und $D_\lambda D_\lambda f(0, 0) = 0$

$$s \partial_s \lambda(s) \langle y^*, D_x D_\lambda f(0, 0) x_0 \rangle + \tilde{\mu}(s) \langle y^*, x_0 \rangle = o(1) (|s \partial_s \lambda(s)| + |\tilde{\mu}(s)|).$$

Mit Gleichung (12.3) können wir im ersten Term die Klammer ersetzen und erhalten

$$(12.7) \quad s \partial_s \lambda(s) \partial_\lambda \bar{\mu}(0) + \tilde{\mu}(s) = o(1) (|s \partial_s \lambda(s)| + |\tilde{\mu}(s)|).$$

Dies impliziert die Behauptungen. \square

Beispiel 1 (für Sätze 11.5 und 12.3). Sei $G \subset \mathbb{R}^n$ beschränktes Gebiet mit glattem Rand. Wir betrachten die Gleichung

$$\begin{aligned} f(u, \mu) &:= (\Delta - \mu)u - u^2 = 0 && \text{in } G, \\ u &= 0 && \text{auf } \partial G. \end{aligned}$$

Info aus PDE-Theorie: Die Eigenwerte von Δ sind $\dots \leq \mu_2 \leq \mu_1 < \mu_0 < 0$. Dabei ist μ_0 (algebraisch) einfach und hat eine positive Eigenfunktion u_0 .

Aus Satz 11.1 folgt: Eine Verzweigung aus dem trivialen Ast $u = 0$ findet höchstens für $\mu = \mu_i$ statt, denn nur dort ist $D_x f(0, \mu) = \Delta - \mu$ nicht invertierbar.

Aus Satz 11.5 folgt: In μ_0 findet eine Verzweigung statt, denn μ_0 ist algebraisch einfach und

$$D_\lambda D_x f(0, \mu_0)u_0 = -\text{id}(u_0) = -u_0 \notin R(\Delta - \mu_0).$$

Es gibt also einen nichttrivialen Ast $(u(s), \lambda(s))$ mit

$$\mu(s) = \mu_0 + \lambda(s), \quad u(s) = su_0 + u_1(s), \quad f(u(s), \mu(s)) = 0.$$

Bestimmung des Verzweigungsbildes. Wir differenzieren die Gleichung

$$(\Delta - \mu_0)u(s) = \lambda(s)u(s) + u^2(s)$$

nach s und erhalten

$$(\Delta - \mu_0)(u_0 + \partial_s u_1(s)) = \lambda(s)\partial_s u(s) + \partial_s \lambda(s)u(s) + 2u(s)\partial_s u(s).$$

Wir differenzieren nochmal nach s und werten in $s = 0$ aus.

$$(\Delta - \mu_0)\partial_s^2 u_1(0) = 2\partial_s \lambda(0)\partial_s u(0) + 2|\partial_s u(0)|^2.$$

Auf der linken Seite ist $\partial_s^2 u_1(0) \in X_1 \perp \mathbb{R}u_0$ und ebenso die ganze linke Seite. Wir multiplizieren die Gleichung mit u_0 und integrieren über das Gebiet G . Es folgt

$$0 = 2\partial_s \lambda(0)\|u_0\|_{L^2}^2 + \int_G 2|u_0|^3.$$

Insbesondere $\lambda'(0) < 0$; damit ist das Verzweigungsbild bestimmt.

Stabilitätsuntersuchung. Die Eigenwerte von $\Delta - \mu_0 - \lambda$ sind alle negativ für $\lambda > 0$, für $\lambda < 0$ gibt es einen positiven Eigenwert (der kritische Eigenwert ist genau $\bar{\mu}(\lambda) = -\lambda$ zur Eigenfunktion u_0). Also ist der triviale Ast stabil für $\lambda > 0$ und instabil für $\lambda < 0$.

Es gilt $\partial_\lambda \mu(0) = -1 \neq 0$ wie in Satz 12.3 ausgesagt. Wir schließen aus der Formel für die Limiten

$$\operatorname{sgn}(s\lambda'(s)) = -\operatorname{sgn}(s), \quad \operatorname{sgn}(\partial_\lambda \mu(0)) = -1 \Rightarrow \operatorname{sgn}(\tilde{\mu}(s)) = -\operatorname{sgn}(s).$$

Also ist der nichttriviale Ast stabil für $s > 0$ und instabil für $s < 0$.

Beispiel 2 Mit demselben G betrachten wir jetzt die Gleichung

$$\begin{aligned} f(u, \mu) &:= (\Delta - \mu)u - u^3 = 0 && \text{in } G, \\ u &= 0 && \text{auf } \partial G. \end{aligned}$$

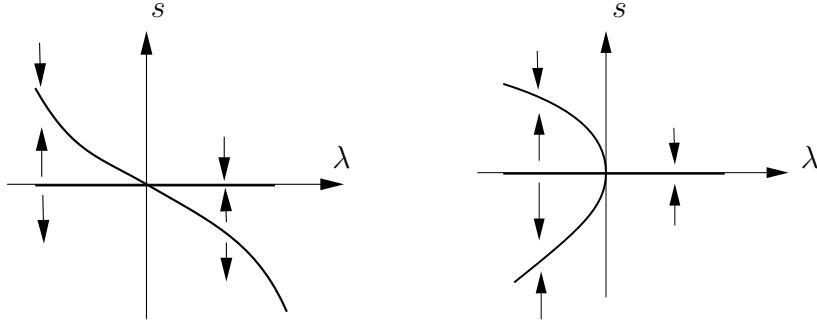


ABBILDUNG 2. Die Verzweigungsbilder mit Stabilitätsresultaten in Beispielen 1 und 2

Die möglichen Verzweigungspunkte (auf der trivialen Achse) sind wieder genau die μ_i , in $(0, \mu_0)$ findet eine Verzweigung statt, der nichttriviale Ast sei gegeben durch $(u(s), \lambda(s))$. Differenzieren der Gleichung liefert

$$(\Delta - \mu_0)(u_0 + \partial_s u_1(s)) = \lambda(s)\partial_s u(s) + \partial_s \lambda(s)u(s) + 3u(s)^2\partial_s u(s).$$

Wir differenzieren nochmal nach s .

$$\begin{aligned} (\Delta - \mu_0)\partial_s^2 u_1(s) &= \lambda(s)\partial_s^2 u(s) + 2\partial_s \lambda(s)\partial_s u(s) \\ &\quad + \partial_s^2 \lambda(s)u(s) + 3u(s)^2\partial_s^2 u(s) + 6u(s)(\partial_s u(s))^2. \end{aligned}$$

Wir multiplizieren die Gleichung mit u_0 im L^2 -Skalarprodukt. übrig bleiben die Terme

$$0 = 2\partial_s \lambda(s)\|u_0\|^2 + \partial_s^2 \lambda(s)s\|u_0\|^2 + O(s^2) + 6 \int_G s|u_0|^2|\partial_s u(s)|^2.$$

Daraus schließen wir in einem ersten Schritt, dass $\partial_s \lambda(s)$ gegen 0 geht für $s \rightarrow 0$, also $\partial_s \lambda(0) = 0$. Die Taylor-Entwicklung des ersten Termes liefert dann

$$2\partial_s^2 \lambda(0)s\|u_0\|^2 + \partial_s^2 \lambda(s)s\|u_0\|^2 + 6 \int_G s|u_0|^2|\partial_s u(s)|^2 = o(s).$$

Wir teilen durch s und bilden dann den Limes $s \rightarrow 0$ und finden

$$\partial_s^2 \lambda(0) = -\frac{2}{\|u_0\|^2} \int_G |u_0|^4.$$

Insbesondere ist $\partial_s^2 \lambda(0)$ negativ und die Verzweigung geht 'nach links', man sagt, die Verzweigung ist subkritisch.

Wir wissen schon, dass der triviale Ast instabil ist für $\lambda < 0$ und stabil für $\lambda > 0$. Mit der Formel von Crandall und Rabinowitz bestätigt man, dass der nichttriviale Ast stabil ist (wie man es sich mit dem 'Pfeilchenbild' auch überlegt).

13. Globale Verzweigung

In diesem Abschnitt wollen wir untersuchen, wie sich die nichttrivialen Äste verhalten können, wenn sie die bisher betrachtete kleine Umgebung des Verzweigungspunktes verlassen.

Die Situation ist wie in Abschnitt 5.4: X ist ein Banachraum, $\hat{\Omega} \subset X \times \mathbb{R}$ eine offene Umgebung von $(0, \lambda_0)$, $T : X \rightarrow X$ linear, stetig und kompakt, $g : \text{cl}(\hat{\Omega}) \rightarrow X$ stetig und kompakt mit $g(x, \lambda) = o(\|x\|)$ für $\|x\| \rightarrow 0$, gleichmäßig in λ . Wir betrachten

$$f(x, \lambda) := x - \lambda T x + g(x, \lambda).$$

$1/\lambda_0$ sei ein Eigenwert von T mit ungerader (algebraischer) Vielfachheit.

Wir betrachten nun die Menge der nichttrivialen Lösungen,

$$M := \left\{ (x, \lambda) \in \hat{\Omega} \mid f(x, \lambda) = 0, x \neq 0 \right\}.$$

Uns interessiert im speziellen der Ast, der in $(0, \lambda_0)$ entstanden ist. Wir setzen

$C :=$ Zusammenhangskomponente von \bar{M} , die $(0, \lambda_0)$ enthält.

Dabei verwenden wir folgende Definitionen: eine Menge heißt zusammenhängend, falls es keine Zerlegung in nichttriviale offene Teilmenge gibt. Eine Zusammenhangskomponente ist eine maximale zusammenhängende Teilmenge.

Unser **Ziel** ist es, zu zeigen:

SATZ 13.1 (Rabinowitz). C läuft entweder gegen $\partial\hat{\Omega}$ oder zurück zum trivialen Ast $\{0\} \times \mathbb{R} \setminus (0, \lambda_0)$.

Oftmals ist $\hat{\Omega} = X \times \mathbb{R}$ der gesamte Raum. Dann bedeutet die erste Möglichkeit, dass C unbeschränkt ist. Wir haben die Aussage bewiesen, falls in folgender Annahme einen Widerspruch finden:

Es gibt eine beschränkte Teilmenge $\Omega \subset \hat{\Omega}$, so dass
 $C \cap \partial\Omega = \{ \}$ und $C \cap \{0\} \times \mathbb{R} = \{(0, \lambda_0)\}$.

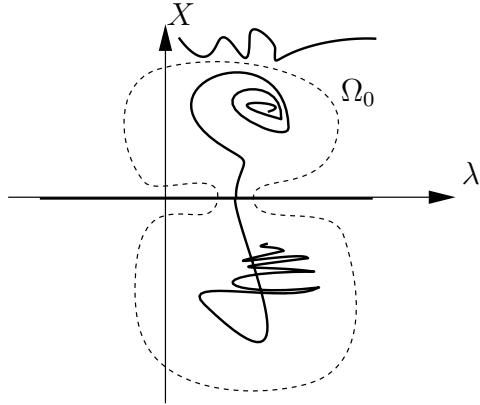
f ist von der Form $\text{id} + kpt$, also nach Proposition 3.6 eigentlich, d.h. Urbilder (in beschränkten Mengen) von kompakten Mengen sind kompakt. Also ist $(f|_{\bar{\Omega}})^{-1}(0)$ kompakt. Es gilt $M = (f|_{\bar{\Omega}})^{-1}(0) \setminus (\{0\} \times \mathbb{R})$ und daher ist auch die abgeschlossene Teilmenge $\bar{M} \cap \bar{\Omega}$ kompakt. Dann ist auch die Zusammenhangskomponente C kompakt.

Annahme: Wir finden eine offene Umgebung Ω_0 von C , die C von $\partial\Omega$ trennt, $C \subset \Omega_0 \subset \bar{\Omega}_0 \subset \Omega$, mit $\bar{M} \cap \partial\Omega_0 = \{ \}$.

Unter Verwendung dieser Annahme finden wir den gesuchten Widerspruch. Zunächst modifizieren wir Ω_0 so, dass gilt

$$\bar{\Omega}_0 \cap (\{0\} \times \mathbb{R}) = [\lambda_0 - \delta, \lambda_0 + \delta],$$

$$B_\delta(0) \times (\lambda_0 - \delta, \lambda_0 + \delta) \subset \Omega_0,$$

ABBILDUNG 3. Die offene Umgebung Ω_0 von C

für ein kleines $\delta > 0$. Dies ist möglich, da jeder Punkt $(0, \lambda) \neq (0, \lambda_0)$ einen positiven Abstand zu C hat (C ist kompakt und kehrt nicht zur Achse zurück); wir können also ein $\bar{B}_\varepsilon(0, \lambda_0)$ aus Ω_0 ausschneiden.

Mit der Notation

$$\Omega_0(\lambda) := \{(x, \lambda') \in \Omega_0 \mid \lambda' = \lambda\}$$

können wir Abbildungsgrade berechnen. Die Zahl

$$d(f(., \lambda), \Omega_0(\lambda), 0)$$

ist unabhängig von $\lambda \in (\lambda_0 - \delta, \lambda_0 + \delta)$. Wohlgemerkt, dies ist nicht nur die Homotopieeigenschaft (d4). Man verwendet zusätzlich, dass die Nullstellenmenge einen positiven Abstand zum Rand hat, um lokal das Gebiet konstant zu wählen. Formal wird dabei das Ausschneideaxiom (d5) verwendet.

Wir wählen nun zwei Zahlen λ_1, λ_2 , beide nahe genug an λ_0 und mit $\lambda_1 < \lambda_0 < \lambda_2$. Wir wollen den obigen Grad an den zwei Stellen λ_i berechnen. Wegen (d5) gilt

$$\begin{aligned} d(f(., \lambda_i), \Omega_0(\lambda_i), 0) \\ = d(f(., \lambda_i), \Omega_0(\lambda_i) \setminus B_\rho(0), 0) + d(f(., \lambda_i), B_\rho(0), 0) \end{aligned}$$

Der erste Term der rechten Seite verschwindet, denn wir können $\tilde{\lambda}_1 = \lambda_1$ soweit nach links verschieben (und $\tilde{\lambda}_2 = \lambda_2$ soweit nach rechts), dass die Mengen $\Omega_0(\tilde{\lambda}_i) \setminus B_\rho(0)$ leer sind. Damit ist der Abbildungsgrad dann 0.

Der zweite Term der rechten Seite ist gerade der Index von $f(., \lambda_i)$ in der 0. Nach Voraussetzung über die ungerade Vielfachheit von $1/\lambda_0$ springt dieser Index. Dies liefert den gesuchten Widerspruch, denn die linke Seite ist gleich für $i = 1, 2$.

Rechtfertigung der Annahme: Die natürliche Wahl für Ω_0 wäre natürlich eine δ -Umgebung $B_\delta(C)$ von C . Die Frage ist allerdings, ob wir durch Wahl eines kleinen δ erreichen, dass $\bar{B}_\delta(C) \cap \bar{M} = C$. Dann könnten wir $\Omega_0 = B_\delta(C)$ setzen, denn dann wären keine Nullstellen

von f auf $\partial B_\delta(C)$. Solch ein kleines $\delta > 0$ könnten wir leicht finden, wenn $(\bar{M} \cap \bar{\Omega}) \setminus C$ kompakt wäre (es reichte: abgeschlossen), denn dann hätte diese Menge von C einen positiven Abstand. Diese Abgeschlossenheit ist allerdings nicht gegeben, und wir müssen ein wenig Arbeit investieren. Vorstellung: Kann sich $\bar{M} \setminus C$ an unendlich vielen Stellen beliebig nah an C annähern?

Wir verwenden das folgende Lemma mit $A := C$, $M := (\bar{M} \cap \bar{\Omega}) \cap \bar{\Omega}$, $B := \bar{M} \setminus B_\delta(C)$. Das Lemma liefert uns dann M_1, M_2 mit positivem Abstand, und wir wählen als unser Ω_0 eine offene Umgebung von M_1 , deren Abschluss M_2 nicht trifft.

LEMMA 13.2. *Sei (M, d) ein kompakter metrischer Raum, $A \subset M$ eine Zusammenhangskomponente und $B \subset M$ abgeschlossen mit $A \cap B = \{\}$. Dann gibt es kompakte $M_1 \supset A$, $M_2 \supset B$, so dass $M = M_1 \cup M_2$ und $M_1 \cap M_2 = \{\}$.*

BEWEIS. Man verwendet den Begriff der ε -Ketten: $a, b \in M$ sind durch eine ε -Kette verbindbar, falls es endlich viele Punkte x_1, \dots, x_n gibt mit $x_1 = a$, $x_n = b$, und $d(x_k, x_{k+1}) < \varepsilon \forall k = 1, \dots, n-1$.

Mit diesem Begriff setzen wir

$$A_\varepsilon := \{x \in M \mid \exists \varepsilon\text{-Kette von } a \in A \text{ nach } x\}.$$

Bild: Vielleicht nähert sich $\bar{M} \setminus C$ an unendlich vielen Stellen C . Dann werden die Teile von $\bar{M} \setminus C$, die näher als ε heranreichen, in die Menge A_ε mitgenommen (obwohl es vielleicht keine Wegverbindung gibt).

Es gilt $A \subset A_\varepsilon$ und A_ε ist offen (eine ε -Umgebung ist auch noch enthalten) und abgeschlossen (da es keine Kette nach $z \in M \setminus A_\varepsilon$ gibt, kann es auch keine zu einem Punkt in $B_\varepsilon(z)$ geben, also $B_\varepsilon(z) \subset M \setminus A_\varepsilon$). Daher sind $M_1 = A_\varepsilon$ und $M_2 = M \setminus A_\varepsilon$ Kandidaten für die gesuchten Mengen. Es bleibt zu zeigen: $B \subset M \setminus A_\varepsilon$ für $\varepsilon > 0$ klein genug.

Wir nehmen das Gegenteil an; zu $\varepsilon = \frac{1}{n}$ gibt es dann ε -Ketten X_n , die einen Punkt b_n aus B mit einem Punkt a_n aus A verbinden. Wegen Kompaktheit von A und B können wir $a_n \rightarrow a_0 \in A$ und $b_n \rightarrow b_0 \in B$ annehmen. Wir betrachten

$$M_0 := \{x \in M \mid \exists x^n \in X_n : x \text{ ist Häufungspunkt von } (x^n)_{n \in \mathbb{N}}\}.$$

Dann ist M_0 abgeschlossen nach Definition, und daher kompakt. Angenommen, M_0 wäre nicht zusammenhängend und könnte also in zwei abgeschlossene Teilmengen C_1 und C_2 zerlegt werden. Dann gilt $\rho := \text{dist}(C_1, C_2) > 0$. Für beliebige Punkte c_1, c_2 und $\varepsilon > 0$ gilt, dass die Punkte durch eine ε -Kette in M verbunden werden können (gehe von einem Punkt nahe c_1 zu a und von einem Punkt nahe c_2 zu a und mache daraus eine Gesamtkette). Wir wählen $\varepsilon = \rho/3$ und finden in den Ketten Punkte x^n , die mindestens Abstand $\rho/3$ zu beiden Mengen C_i haben. Diese Punkte haben einen Häufungspunkt, der in keiner der C_i liegt — ein Widerspruch zu $C_1 \cup C_2 = M_0$. Also ist M_0 zusammenhängend.

Wir schließen $M_0 \subset A$, da A eine Zusammenhangskomponente ist und $a_0 \in M_0$. Es ist aber auch $b_0 \in M_0$ nach Konstruktion und daher $b_0 \in A \cap B$, ein Widerspruch. \square

Ein Beispiel für globale Verzweigung. Wir betrachten auf dem Intervall $(0, \pi)$ für $u \in X := \{v \in C^2((0, \pi), \mathbb{R}) | v(0) = v(\pi) = 0\}$ die Gleichung

$$(13.1) \quad \partial_x^2 u + \lambda u - \Phi(., u(.), \partial_x u(.)) = 0.$$

Dabei sei $\Phi : (0, \pi) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ eine Nichtlinearität mit $|\Phi(x, s, p)| \leq C(|s|^2 + |p|^2) \forall x, s, p$. Der Laplace Operator $\Delta : X \rightarrow C^0((0, \pi), \mathbb{R})$ ist invertierbar, durch Anwendung von $T := \Delta^{-1}$ formen wir die Gleichung um in

$$f(u, \lambda) := u + \lambda T u - N(u) = 0.$$

Diese Gleichung erfüllt alle Voraussetzungen von Satz 11.5 und es gibt Verzweigungen in allen Eigenwerten, also in $\lambda_k = -k^2, k = 1, 2, \dots$ (alle Eigenwerte sind einfach). Seien C_k die zugehörigen Zusammenhangskomponenten der Lösungsmenge. In diesem Beispiel können wir das Verhalten der Lösungsäste global bestimmen, weil man im eindimensionalen Nullstellen zählen kann.

SATZ 13.3. Für jedes $k \in \mathbb{N}$ ist die Nullstellenmenge C_k der Gleichung (13.1) unbeschränkt in $X \times \mathbb{R}$. Sie besteht aus Funktionen u mit genau $k - 1$ einfachen Nullstellen in $(0, \pi)$.

BEWEIS. Sei S^k die Menge von Funktionen $u \in X$ mit genau $k - 1$ einfachen Nullstellen in $(0, \pi)$ und $\partial_x u(0) \neq 0, \partial_x u(\pi) \neq 0$. Dann ist S^k offen in X und $S^k \cap S^m = \{\}$ für $k \neq m$. Für den Lösungszweig wissen wir lokal $u(s) = su_k + o(s)$, wobei $u_k(x) = \sin(kx)$ die Eigenfunktion zu $\lambda_k = -k^2$ ist. Daher erfüllen $u(s)$ aus dem k -ten Zweig $u(s) \in S^k$.

Angenommen, in C_k gäbe es Lösungen $(u, \lambda) \notin S^k$. Dann müßte es auch ein $u \in C_k$ geben, das eine doppelte Nullstelle hat. Da u aber eine gewöhnliche Differentialgleichung zweiter Ordnung erfüllt, wäre dann $u \equiv 0$. Damit wäre C_k wieder zum trivialen Ast zurückgekehrt, was weder in den $\lambda_m, m \neq k$ (Anzahl der Nullstellen), noch in anderen Punkten möglich ist (sind keine Verzweigungspunkte).

Bemerke, dass die Äste 'nach oben' und 'nach unten' im Verzweigungsbild durch das Vorzeichen von $\partial_x u(0)$ global unterschieden werden können. \square

14. Hopf-Verzweigung

In diesem Abschnitt betrachten wir die gewöhnliche Differentialgleichung

$$(14.1) \quad \partial_t x(t) = f(x(t), \lambda),$$

und nehmen immer an, dass $x_0 = 0$ eine stationäre Lösung ist, d.h. $f(0, \lambda) = 0 \forall \lambda$. Wie zuvor interessieren uns Punkte $(0, \lambda_0)$, in denen die Stabilität von x_0 sich ändert.

Wenn man z.B. Gleichungen vom Typus $\partial_t u = \Delta u + f(u)$ untersuchen möchte, dann ist es sinnvoll, anstelle der DGL das zugehörige dynamische System zu betrachten, welches einen Startwert x auf den Wert $x(t)$ wirft, wobei $x(\cdot)$ Lösung der Gleichung zum Parameter λ ist.

$$(14.2) \quad X \ni x \mapsto T(x, t, \lambda) \in X.$$

Wir wollen nun die Grundidee der Hopf-Verzweigung skizzieren.

Lineare Gleichung. Wir betrachten die Gleichung für $x = (x_1, x_2)$,

$$(14.3) \quad \partial_t x = \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} x.$$

Für $\lambda = 0$ hat diese Gleichung sehr spezielle Lösungen, nämlich $x(t) = a \cdot (\sin(t), \cos(t))$, wobei $a \in \mathbb{R}$ beliebig ist. Wir finden also eine Schar von (nichttrivialen) periodischen Lösungen der Gleichung. Wir wollen im Folgenden die Lösungen $a \cdot (\sin(t + \varphi), \cos(t + \varphi))$ nicht als neue Lösungen ansehen, denn es sind die alten, nur phasenverschoben.

Für $\lambda < 0$ spiralen alle Lösungen in die 0 hinein, für $\lambda > 0$ spiralen alle von der 0 weg. Dies ist typisch: Die stationäre Lösung $x_0 = 0$ hat für $\lambda = 0$ ihre Stabilität verloren. Während dieses Prozesses sind keine weiteren stationären Lösungen entstanden, sondern periodische.

Das Ergebnis dieses Abschnittes ist im Wesentlichen: Dieses Bild gilt auch in der nichtlinearen Situation. Die Abbildung zeigt die Menge der (s, λ) , so dass $(0, s)$ Startwert ist von einer periodischen Lösung zu $f(\cdot, \lambda)$.

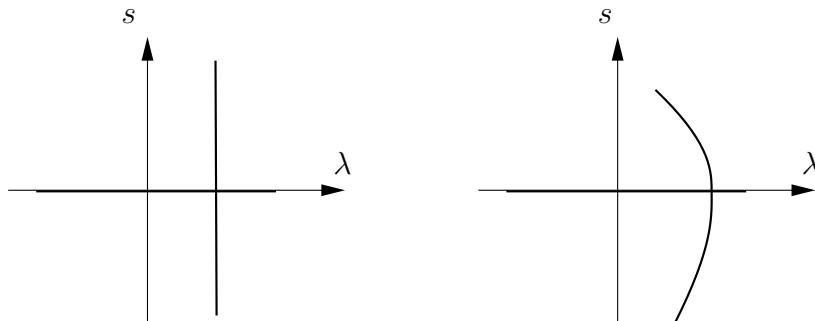


ABBILDUNG 4. Die Menge von Startwerten für periodische Lösungen im linearen und im nichtlinearen Fall

Die Beweisidee ist: Für $\lambda = 0$ wandert ein Punkt $(s, 0)$ der Ebene in der Zeit 2π wieder genau auf den Startpunkt. Eine Nichtlinearität wird die Bahn verschieben, so dass sie nicht genau wieder $(s, 0)$ trifft. Aber: Vergrößern von t schiebt den Zielpunkt nach oben, vergrößern von λ nach rechts. Durch Anpassen dieser zwei Parameter treffen wir genau wieder $(s, 0)$.

Nichtlineare Gleichung. Wir wollen auch Gleichungen in Banachräumen zulassen und betrachten daher zunächst die allgemeine Situation: X Banachraum, $T(x, t, \lambda)$ ein dynamisches System auf X (wir verzichten hier auf die Definition von 'dynamischem System', wesentlich ist, dass $T(x, t, \lambda)$ für alle $t \in \mathbb{R}_+$ definiert ist, und $T(., t + s, \lambda) = T(., t, \lambda) \circ T(., s, \lambda)$). Die Abbildung $T(., t)$ heißt Poincaré-Abbildung, falls eine zugrundeliegende Evolutionsgleichung gegeben ist.

LEMMA 14.1. *Wir betrachten eine Abbildung $T : X \times \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ der Klasse C^2 seine Linearisierung*

$$L(t, \lambda) : X \rightarrow X, x \mapsto D_x T(0, t, \lambda) \langle x \rangle.$$

Wir setzen folgendes voraus.

- a) *id* – $L(t_0, \lambda_0)$ ist Fredholm Operator vom Index 0 und hat einen 2-dimensionalen Kern $X_0 = \text{span}(u, v)$.

Insbesondere finden wir eine Projektion P auf ein Co-Bild Y_0 von $Y_1 = R(id - L(t_0, \lambda_0))$ mit $\ker P = Y_1$.

- b) *Folgende Abbildung ist invertierbar:*

$$D_{(t, \lambda)} PL(t_0, \lambda_0) \langle u \rangle : \mathbb{R} \times \mathbb{R} \rightarrow Y_0$$

Dann gibt es eine Schar von Fixpunkten (das dynamische System hat eine Schar periodischer Lösungen): Für kleines s finden wir Startwerte $x(s) = su + x_1(s)$, Perioden $t(s) = t_0 + o(1)$ und Parameter $\lambda(s) = \lambda_0 + o(1)$, so dass

$$T(x(s), t(s), \lambda(s)) = x(s).$$

BEWEIS. Wir machen zunächst eine Ljapunov-Schmidt Reduktion mit Parameterraum $\Lambda = \mathbb{R}^2 \ni (t, \lambda)$. Die Gleichung $x - T(x, t, \lambda) = 0$ wird lokal genau dann von x gelöst, wenn

$$x = x_0 + x_1(x_0, t, \lambda),$$

$$\Phi(x_0, t, \lambda) := Px_0 - PT(x_0 + x_1(x_0, t, \lambda), t, \lambda) = 0.$$

Hierbei gilt $x_1(0, t, \lambda) = 0$ und $D_{x_0} x_1(0, t_0, \lambda_0) = 0$. Wir machen den speziellen Ansatz $x_0 = su$ und schreiben nun die Verzweigungsgleichung äquivalent als

$$(14.4) \quad \Psi(s, t, \lambda) := \begin{cases} \frac{1}{s} \Phi(su, t, \lambda) & \text{für } s \neq 0, \\ D_x \Phi(0, t, \lambda) \langle u \rangle & \text{für } s = 0. \end{cases}$$

Die nichttrivialen Nullstellen von Φ sind genau die nichttrivialen Nullstellen von Ψ ; die Abbildung Ψ ist von der Klasse C^1 . Wir haben eine spezielle Nullstelle von Ψ ,

$$\Psi(0, t_0, \lambda_0) = Pu - PL(t_0, \lambda_0) \langle u \rangle = 0.$$

Wir wollen nun auch für kleine $s \neq 0$ lösen, wir wollen also nach (t, λ) auflösen. Dies geht mit dem Satz über implizite Funktionen, falls

$$D_{(t, \lambda)} \Psi(0, t_0, \lambda_0) = D_{(t, \lambda)} (Pu - PL(t, \lambda) \langle u \rangle)|_{(t_0, \lambda_0)} = D_{(t, \lambda)} PL(t_0, \lambda_0) \langle u \rangle$$

invertierbar. Dies war vorausgesetzt. \square

SATZ 14.2 (Hopf-Verzweigung). Sei $f \in C^2(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$ mit $f(0, \lambda) = 0 \forall \lambda \in (-\delta, \delta)$. Wir nehmen an:

- (1) Imagiäre Eigenwerte: Die Linearisierung $A(\lambda) := D_x f(0, \lambda)$ hat für $\lambda = \lambda_0$ die konjugiert komplexen einfachen Eigenwerte $\pm i\beta_0$, $\beta_0 \in \mathbb{R}_+$.
- (2) Resonanz: Alle Vielfachen $ki\beta_0$ sind keine Eigenwerte von $A(\lambda_0)$.
- (3) Transversalität: Die C^1 -Fortsetzung $\alpha(\lambda) + i\beta(\lambda)$ des Eigenwertes $i\beta_0$ erfüllt $\partial_\lambda \alpha(\lambda_0) \neq 0$.

Dann hat die Gleichung $\partial_t x = f(x, \lambda)$ eine Schar von nichttrivialen periodischen Lösungen der Form $x(t, s)$ zu Parametern $\lambda(s) = \lambda_0 + o(1)$ mit Periode $t(s) = 2\pi/\beta_0 + o(1)$ und Startwerten $x(0, s) = su + o(s)$ für kleine s .

BEWEIS. Die Aussage ist genau die aus Lemma 14.1 für $t_0 := 2\pi/\beta_0$ und T das dynamische System zu f . Wir müssen die Voraussetzungen a) und b) aus dem Lemma nachprüfen. Ausgangspunkt ist folgende Behauptung:

Falls $T(., t, \lambda)$ die Poincaré Abbildung zu $x \mapsto f(x, \lambda)$ ist, dann ist $L(t, \lambda) \langle . \rangle$ die Poincaré Abbildung zu $x \mapsto A \cdot x$.

Der Beweis ist elementar: man betrachtet $\partial_t T(x_0, t) = f(T(x_0, t))$ und differenziert nach x in eine Richtung \bar{x} :

$$\begin{aligned} \partial_t L(t, \lambda) \langle \bar{x} \rangle &= D_x \partial_t T(0, t, \lambda) \langle \bar{x} \rangle \\ &= D_x f(T(0, t)) D_x T(0, t, \lambda) \langle \bar{x} \rangle = AL(t, \lambda) \langle \bar{x} \rangle. \end{aligned}$$

Wir betrachten nun die Komplexifizierung \mathbb{C}^n des Raumes \mathbb{R}^n und zerlegen \mathbb{C}^n in die verallgemeinerten Eigenräume von A . Zu einem Eigenvektor v mit Eigenwert μ gehört die Evolution

$$v(t) := ve^{\mu t} \quad \text{löst} \quad \partial_t v(t) = Av(t).$$

Also wird der Kern von $\text{id} - L(t_0, \lambda_0)$ aufgespannt von den Eigenvektoren von A , für die $e^{t_0 \mu} = 1$, also für $t_0 = 2\pi/\beta_0$ von den Eigenvektoren mit Eigenwerten $ik\beta_0$. Unter Annahmen 1. und 2. ist der Kern 2-dimensional.

Wir wählen nun zwei Eigenvektoren u, v , so dass $Au = i\beta_0 u$, $Av = -i\beta_0 v$. Als entsprechende reelle Basisvektoren können $w_1 := \text{Re } u$ und $w_2 := \text{Im } u$ gewählt werden, für diese gilt $Aw_1 = \text{Re}(Au) = \text{Re}(i\beta_0 u) = -\beta_0 \text{Im } u = w_2$ und ebenso $Aw_2 = \beta_0 w_1$. Als Projektion P wählen wir die Projektion auf $\text{span}(w_1, w_2) \subset X$.

Wir setzen u und v fort zu Eigenvektoren $u(\lambda), v(\lambda)$ mit $Pu(\lambda) = u, Pv(\lambda) = v$ zu Eigenwerten $\alpha(\lambda) \pm i\beta(\lambda)$. Für diese wollen wir $D_{(t, \lambda)} PL(t_0, \lambda_0) : \mathbb{R} \times \mathbb{R} \rightarrow X_0$ aus dem Lemma berechnen. Es gilt

$$L(t, \lambda) : u(\lambda) \mapsto e^{(\alpha(\lambda) \pm i\beta(\lambda))t} u(\lambda),$$

also ebenso auch für die Projektion PL . Wir differenzieren nach t und λ und erhalten

$$\begin{aligned}\partial_t PL(t_0, \lambda_0) \langle u \rangle &= i\beta_0 u, \\ \partial_\lambda PL(t_0, \lambda_0) \langle u \rangle &= (\alpha'(\lambda_0)t_0 + i\beta'(\lambda_0))u.\end{aligned}$$

In der Schreibweise mit reellen Basisvektoren w_1 und w_2 lautet dieses Ergebnis

$$\begin{aligned}\partial_t PL(t_0, \lambda_0) \langle w_1 \rangle &= -\beta_0 w_2, \\ \partial_\lambda PL(t_0, \lambda_0) \langle w_1 \rangle &= \alpha'(\lambda_0)t_0 w_1 - \beta'(\lambda_0)w_2.\end{aligned}$$

Wegen der Transversalität (3) $\alpha'(\lambda_0) \neq 0$ ist die Abbildung $D_{(t,\lambda)}PL(t_0, \lambda_0) \langle w_1 \rangle : \mathbb{R}^2 \rightarrow X_0 = \text{span}(w_1, w_2)$ invertierbar. Dies war zu zeigen. \square

Anmerkung zum Beweis: Unser Beweis rechtfertigt die Vorstellung, dass wir im \mathbb{R}^2 im Punkt $(s, 0)$ starten, und nach Zeit t wieder in $(s, 0)$ sein wollen. Durch Wahl der beiden Parameter t und λ können wir dies erreichen.

Ein Beispiel. Betrachte den Oszillator

$$y'' + (y')^3 - \lambda y' + y = 0.$$

Wir schreiben dies als System, $u = (y, y')$,

$$u' = \begin{pmatrix} 0 & 1 \\ -1 & \lambda \end{pmatrix} u - \begin{pmatrix} 0 \\ u_2^3 \end{pmatrix}.$$

Die Eigenwerte sind $\rho \in \mathbb{C}$ mit

$$-\rho(-\rho + \lambda) + 1 = 0,$$

also

$$\rho_{1,2} = \frac{1}{2}(\lambda \pm \sqrt{\lambda^2 - 4}).$$

Für $\lambda = \lambda_0 = 0$ finden wir konjugiert komplexe Eigenwerte mit

$$\partial_\lambda \text{Re } \rho_{1,2} = \frac{1}{2} \neq 0.$$

Es geschieht also eine Hopf-Verzweigung in $\lambda = 0$; es gibt nichttriviale periodische Lösungen kleiner Amplitude in der Nähe von λ_0 .

Ein alternativer Beweis der Hopf-Verzweigung. Eher im Geiste dieser Vorlesung wäre ein anderer Beweis für die Existenz nichttrivialer periodischer Lösungen. Wir könnten einen Raum von Funktionen $u : t \mapsto u(t)$ definieren und in diesem Raum die Gleichung

$$F(u, \lambda) := \partial_t u(\cdot) - f(u(\cdot), \lambda) = 0$$

betrachten. Gesucht sind periodische Lösungen, wir sollten also einen Funktionenraum periodischer Funktionen wählen. Ein Problem ist,

dass die Periodenlänge T auch unbekannt ist. Daher skalieren wir die Gleichung um:

$$U(\tau) := u(T \cdot \tau), \quad \partial_\tau U(\tau) = T \cdot f(U(\tau), \lambda).$$

Wir suchen T -periodische Lösungen u , also 1-periodische Lösungen U . Also:

Aufgabe: Finde λ nahe λ_0 und T nahe T_0 und nichttriviale

$$\begin{aligned} U \in X &:= C_{per}^1([0, 1], \mathbb{R}^n) \\ &= \{u \in C^1([0, 1], \mathbb{R}^n) | u(1) = u(0), \partial_t u(1) = \partial_t u(0)\}, \end{aligned}$$

welche für

$$\begin{aligned} F : X \times \mathbb{R} \times \mathbb{R} &\rightarrow Y := C_{per}^0([0, 1], \mathbb{R}^n), \\ F(U, \lambda, T) &:= \partial_t U - T \cdot f(U, \lambda) \end{aligned}$$

die Gleichung $F(U, \lambda, T) = 0$ lösen.

Wir können nun die Hopf-Verzweigung wie die vorherigen Verzweigungen (also als stationäre Verzweigung) auffassen kann. Wieder bestimmen wir den Kern der Linearisierung. Diesmal ist dieser Kern zweidimensional:

$$\begin{aligned} U_1 : t &\mapsto w_1 \cdot \sin(2\pi t) + w_2 \cdot \cos(2\pi t), \\ U_2 : t &\mapsto w_1 \cdot \cos(2\pi t) + w_2 \cdot \sin(2\pi t), \\ \ker D_U F(0, \lambda_0, T_0) &= \{U_1, U_2\}, \end{aligned}$$

wobei w_1 und w_2 Vektoren in X sind, bezüglich denen die Matrix $A = D_x f(0, \lambda_0)$ die Darstellung

$$\begin{pmatrix} 0 & -\beta_0 \\ \beta_0 & 0 \end{pmatrix}$$

hat.

Der Kern der Linearisierung von F ist also 2-dimensional, dafür müssen wir auch nach 2 Parametern T und λ auflösen. Dies können wir nicht mit unseren früheren Sätzen tun, da dort ungerade Dimensionen betrachtet wurden. Der Zweidimensionalität liegt eine besondere Symmetrie von F zugrunde: F bildet Zeit- s -translatierte Funktionen wieder auf ebenso Zeit- s -translatierte rechte Seiten ab. Daher besteht auch der Kern aus einer Funktion und ihren Zeit-Translationen, auch die nicht-trivialen Lösungen sind ein Symmetrieorbit einer einzigen Funktion.

Man kann diese Ideen zu einem Beweis machen, indem man die Symmetrie 'herausdividiert'. Daher kann man Hopf-Verzweigung auch verstehen als eine statioäre Verzweigung mit einer Symmetrie. Der Vorteil des von uns durchgeführten Beweises ist, dass er problemlos auch auf z.B. die Wärmeleitungsgleichung anwendbar ist, denn das zugehörige dynamische System hat sehr schöne Eigenschaften (z.B. Kompaktheit).

Literaturverzeichnis

- [1] H.-W. Alt. *Lineare Funktionalanalysis*. Springer, 1985.
- [2] S.N. Chow and J.K. Hale. *Methods of Bifurcation Theory*. Grundlehren der mathematischen Wissenschaften. Springer, 1982.
- [3] M.G. Crandall and P.H. Rabinowitz. The Hopf bifurcation theorem in infinite dimensions. *Arch. Rat. Mech. Anal.*, 67:53–72, 1978.
- [4] K. Deimling. *Nonlinear Functional Analysis*. Springer, 1985.
- [5] W. Gromes. Ein einfacher Beweis des Satzes von Borsuk. *Math. Z.*, 178:399–400, 1981.
- [6] W. Jäger. Nichtlineare Funktionalanalysis. Mitschrift der Vorlesung des WS 1990/1991.
- [7] L. Nirenberg. Topics in nonlinear functional analysis. Courant Institute of Mathematical Sciences, 1974.
- [8] P. Rabinowitz. A global theorem for nonlinear eigenvalue problems and applications. In *Contributions to Nonlinear Functional Analysis*.
- [9] M. Ruzicka. *Nichtlineare Funktionalanalysis*. Springer, 2003.
- [10] M.A. Krasnosel'skii. *Topological Methods in the Theory of Nonlinear Integral Equations*. Macmillan, 1964.