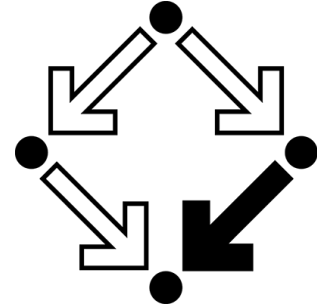




**JOHANNES KEPLER
UNIVERSITY LINZ**



Gewöhnliche Differentialgleichungen und Dynamische Systeme

Josef Schicho

WS 2020/21

Inhaltsverzeichnis

1	Verschiedene Typen von Dynamischen Systemen	2
1.1	Lineare Rekursionen	4
1.2	Kontinuierliche Dynamische Systeme	5
1.3	Zusammenhang zwischen diskreten und kontinuierlichen Systemen	5
2	Skalare autonome Differentialgleichungen	6
3	Stabile und Asymptotisch Stabile Equilibrien	8
4	Bifurkationen	9
5	Lösungsverfahren	10
6	Symbolische Lösungen	12
6.1	Trennung der Variablen	13
6.2	Reduktion der Ordnung	15
6.3	Variation der Konstanten	17
6.4	Unlösbarkeitsresultate	19
7	Lineare Autonome Rekursionen	20

8	Lineare Autonome Differentialgleichungen	25
8.1	Eine Alternative zu VdK	27
8.2	Die Exponentialmatrix	28
9	Der Fundamentalsatz der Algebra	29
10	Existenz und Eindeutigkeit	32
10.1	Vollständigkeit des Funktionenraums	35
10.2	Beweis des Satzes von Picard/Lindelöf	37
10.3	Differenzierbare Vektorfelder	38
10.4	Weitere Folgerungen aus Picard/Lindelöf	39
11	Äquivalenz und lokale Äquivalenz	42
12	Stabile Fixpunkte	48
13	Die Poincaré-Rückkehrabbildung	51
13.1	Der Satz von Poincaré/Bendixson	52
13.2	Chaos	56
14	Ausblick: Differentialgeometrie	57

1 Verschiedene Typen von Dynamischen Systemen

Ein dynamisches System ist ein mathematisches Modell eines Prozesses durch Funktionen von einer Menge T (Zeit) in einen Raum X . Die Menge T ist entweder gleich \mathbb{R} bzw. ein Intervall von \mathbb{R} (kontinuierliches dynamisches System) oder gleich \mathbb{N} (diskretes dynamisches System). Der Raum X ist meistens ein reeller endlich-dimensionaler Vektorraum – bei linearen dynamischen Systemen ist das unabdingbar – oder eine Teilmenge eines Vektorraumes. Das Wesentliche ist, dass die Funktion (manchmal auch eine Menge von Funktionen) nicht durch einen expliziten Ausdruck beschrieben wird, sondern im diskreten Fall durch eine Rekursion und im kontinuierlichen Fall durch eine Differentialgleichung. Wir beginnen mit dem diskreten Fall: $T = \mathbb{N}$, und $X \subset \mathbb{R}$ oder $X \subset \mathbb{R}^n$. Die Funktion $f : T \rightarrow X$ wird beschrieben durch die Rekursionsgleichung der Ordnung $k > 0$, die wie folgt aussieht:

$$\forall t \in T : f(t+k) = F(f(t), f(t+1), \dots, f(t+k-1), t),$$

wobei $F : (X^k \times T) \rightarrow X$ eine gegebene stetige Funktion ist. Sobald die Anfangswerte $f(0), f(1), \dots, f(k-1)$ bekannt sind, ist die Funktion f eindeutig festgelegt und kann mit einem Programm rekursiv an jeder natürlichen Zahl ausgewertet werden.

Sehr oft ist man in der Situation, daß der Wert $F(x_0, \dots, x_{k-1}, t)$ nicht von t abhängt, also dass man eigentlich eine Funktion $F : X^k \rightarrow X$ hat. Wir sprechen in diesem Fall von einer autonomen (zeitunabhängigen) Rekursion. Der nicht-autonome Fall entspricht eigentlich

nicht der ‘‘Philosophie’’ der dynamischen Systeme: wenn wir zeitabhangige Rekursionen erlauben, konnen wir ja auch gleich einen Ausdruck fur die Funktion in die Rekursion stecken:

$$F(x_0, x_1, \dots, x_{k-1}, t) := A(t),$$

und hatzen dann eine explizite Darstellung $f(t) = A(t)$. Genau das, was wir in der Theorie der dynamischen Systeme ja ausschließen wollten. Drum ist der autonome Fall der normale Fall.

Wenn uns jetzt eine nicht-autonome Rekursion uber den Weg lauft und wir trotzdem philosophisch auf der Linie der dynamischen Systeme bleiben wollen, dann ist das moglich: der nicht-autonome Fall lasst sich immer auf den autonomen Fall zururckfuhren. Auerdem konnen Rekursionen beliebiger Ordnung immer auf Rekursionen erster Ordnung zururckgefuhrt werden. Beide Rurckfuhungen sind nicht gratis, sie gehen auf Kosten hoherer Dimension n des Wertebereichs von f .

Wir beschreiben zuerst die Rurckfuhung einer nicht-autonomen Rekursion erster Ordnung fur eine Funktion $f : T \rightarrow X$, gegeben durch die Funktion $F : (X \times T) \rightarrow X$; als Rekursion fur f liest sich das so:

$$\forall t \in T : f(t+1) = F(f(t), t). \quad (1)$$

Wir definieren nun $Y := X \times T$ und die Funktion

$$G : Y \rightarrow Y, (x, t) \mapsto (F(x, t), t+1),$$

die eine Rekursion fur $g : T \rightarrow Y$ definiert:

$$\forall t \in T : g(t+1) = G(g(t)). \quad (2)$$

Wenn nun $f : T \rightarrow X$ die Rekursion 1 erfullt, dann erfullt die Funktion $(x, t) \mapsto (f(t), t)$ die Rekursion 2. Umgekehrt: wenn $g : T \rightarrow Y$ die Rekursion (2) erfullt und einen Startwert $g(0)$ mit t -Komponente gleich 0 hat, dessen t -Komponente gleich 0 ist, dann ist die erste Komponente von g eine Losung von Rekursion (1) und die zweite Komponente ist die Identitat $t \mapsto t$. Bei Rekursionen hoherer Ordnung funktioniert eine ahnliche Rurckfuhung (wobei es beim Ausdenken einer Rekursion fur die Zeitkomponente ein bisschen Platz fur Kreativitat gibt).

Nun beschreiben wir die Rurckfuhung einer autonomen Rekursion von Ordnung $k > 1$ fur eine Funktion $f : T \rightarrow X$, gegeben durch eine Funktion $F : X^k \rightarrow X$ und als Rekursion geschrieben

$$\forall t \in T : f(t+k) = F(f(t), f(t+1), \dots, f(t+k-1)), \quad (3)$$

auf eine Rekursion erster Ordnung. Wir definieren $Y := X^k$ und die Funktion

$$G : Y \rightarrow Y, (x_0, x_1, \dots, x_{k-1}) \mapsto (x_1, x_2, \dots, F(x_0, x_1, \dots, x_{k-1})).$$

Diese definiert wieder eine Rekursion fur $g : T \rightarrow Y$; die Rekursionsgleichung ist die gleiche wie in (2). Wenn $f : T \rightarrow X$ nun die Rekursion 3 erfullt, dann erfullt die Funktion $t \mapsto (f(t), f(t+1), \dots, f(t+k-1))$ die Rekursion (2). Wenn umgekehrt g die Rekursionsgleichung (2) erfullt, dann erfullt jeder ihrer k Komponenten die Rekursion (3).

1.1 Lineare Rekursionen

Es sei nun X ein reeller Vektorraum. Eine nicht-autonome Rekursion gegeben durch eine Funktion $F : (X^k \times T) \rightarrow X$ heißt linear, wenn für jedes $t \in T$ die Funktion $(x_0, \dots, x_{k-1}) \mapsto F(x_0, \dots, x_{k-1}, t)$ linear ist. Die Lösungen bilden einen Vektorraum der Dimension $k = \dim(X)$; die Abbildung, die jedem k -Tupel von Anfangswerten $(x_0, \dots, x_{k-1}) \in X^k$ die eindeutig bestimmte Lösung der Rekursion mit diesen Anfangswerten zuordnet, ist ein Isomorphismus von Vektorräumen.

Die oben beschriebene Rückführung auf eine Rekursion erster Ordnung liefert in der Tat eine lineare Rekursion erster Ordnung. Beschränken wir uns daher (vorläufig) auf lineare Rekursionen erster Ordnung. Es sei $X = \mathbb{R}^n$. Dann kann man die Funktion F auch schreiben als matrixwertige Funktion $A : T \rightarrow \mathbb{R}^{n \times n}$, also $F(x_0, \dots, x_{k-1}, t) := A(t) \cdot (x_0, \dots, x_{k-1})$. Der Vektorraum-Isomorphismus zwischen Anfangswerten und Lösungen läßt sich ebenfalls als matrixwertige Funktion $B \rightarrow \mathbb{R}^{n \times n}$ schreiben. Die Lösung $f : T \rightarrow \mathbb{R}^n$ der Rekursion

$$\forall t \in T : f(t+1) = A(t)f(t)$$

läßt sich dann schreiben als

$$g(t) = B(t)(g(0), \dots, g(k-1)).$$

Die Matrix B erfüllt die Gleichungen

$$\forall t : B(t+1) = A(t)B(t), \quad B(0) = I_n.$$

Die Matrix $B(t)$ kann berechnet werden durch Aufmultiplizieren der Matrizen $A(0), \dots, A(t-1)$ von rechts nach links.

Im autonomen Fall ist das alles viel einfacher: statt einer matrixwertigen Funktion haben wir eine einzige Matrix $A \in \mathbb{R}^{n \times n}$, und $B(t) = A^t$. Diesen Fall werden wir später noch genauer studieren.

Bemerkung 1.1. Wie erwähnt, lassen sich nicht-autonome Rekursionen immer auf autonome Rekursionen zurückführen. Leider wird bei dieser Reduktion die Linearität nicht beibehalten, drum ist die Theorie der nicht-autonomen linearen Rekursionen (bzw. im kontinuierlichen Fall die Theorie der nicht-autonomen linearen Differentialgleichungen) komplizierter als die Theorie der autonomen linearen Rekursionen (bzw. der autonomen linearen Differentialgleichungen).

Wenn nichts dazugesagt wird, verstehen wir unter einer "linearen Rekursion" eine homogene lineare Gleichung. Die inhomogene lineare Rekursion für $f : T \rightarrow \mathbb{R}^n$ sieht so aus:

$$\forall t \in T : f(t+1) = A(t)f(t) + b(t),$$

wobei $A : T \rightarrow \mathbb{R}^{n \times n}$ eine gegebene matrixwertige Funktion und $b : T \rightarrow \mathbb{R}^n$ eine gegebene vektorwertige Funktion ist. Es ist nicht schwierig, eine solche inhomogene Rekursion zurückzuführen auf eine homogene lineare Rekursion für $g : T \rightarrow \mathbb{R}^{n+1}$; die ersten n Komponenten dieser neuen Rekursion bilden f , und die letzte ist konstant 1.

1.2 Kontinuierliche Dynamische Systeme

Hier ist $T = \mathbb{R}$ oder eine zusammenhängende Teilmenge von \mathbb{R} (also ein Intervall); X ist wieder ein endlich-dimensionaler reellen Vektorraum oder eine Teilmenge eines endlich-dimensionalen reellen Vektorraums. Eine Funktion $f : T \rightarrow X$ wird in diesem Fall beschrieben durch eine Differentialgleichung der Ordnung $k > 0$:

$$\forall t \in T : f^{(k)}(t) = F(f(t), f'(t), \dots, f^{(k-1)}(t), t), \quad (4)$$

wobei $F : (X^k \times T) \rightarrow X$ eine gegebene stetige Funktion ist. Im Unterschied zum diskreten Fall ist überhaupt nicht klar, ob die Funktion durch Anfangswerte eindeutig festgelegt ist. Der fundamentale Satz von Picard/Lindelöf, den wir beweisen und ausführlich werden, besagt, daß unter gewissen Voraussetzungen für F die Lösung f der Differentialgleichung eindeutig durch die Anfangswerte $f(0), f'(0), \dots, f^{(k-1)}(0)$ festgelegt ist, und auch daß für gegebene Anfangswerte immer eine Lösung existiert.

In anderer Hinsicht ist die Theorie der kontinuierlichen Systeme dafür wieder einfacher als die diskrete Theorie. Zum Beispiel kann man im kontinuierlichen Fall sehr einfach “die Zeit umdrehen” und ein lineares System aufstellen, das in die Vergangenheit schaut. Man braucht dazu nur in der Differentialgleichung t durch $-t$ ersetzen. Im diskreten Fall kann man die Rekursionen nicht nach rückwärts verfolgen, weil zum Beispiel die Funktion $F : X \rightarrow X$ in der Rekursion $\forall t : f(t+1) = F(f(t))$ im allgemeinen nicht invertierbar ist. Die oben besprochenen Rückführungen für den diskreten Fall lassen sich auch auf den kontinuierlichen Fall übertragen. So kann man eine nicht-autonome Differentialgleichung für $f : T \rightarrow X$

$$\forall t \in T : f'(t) = F(f(t), t)$$

zurückführen auf eine autonome Gleichung für $g : T \rightarrow X \times T$. Und man kann eine Differentialgleichung der Ordnung k für $f : T \rightarrow X$

$$\forall t \in T : f^{(k)}(t) = F(f(t), f'(t), \dots, f^{(k-1)}(t)), \quad (5)$$

auf eine Differentialgleichung erster Ordnung für $g : T \rightarrow X^k$ zurückführen. Im letzteren Fall bleibt Linearität erhalten, im ersteren nicht.

1.3 Zusammenhang zwischen diskreten und kontinuierlichen Systemen

Es gibt mehrere Querverbindungen zwischen dem diskreten und dem kontinuierlichen Fall.

- Falls in einer autonomen Differentialgleichung Existenz und Eindeutigkeit der Lösung des Anfangswertproblems für $f : \mathbb{R} \rightarrow X$

$$\forall t \in \mathbb{R} : f'(t) = F(f(t))$$

erfüllt ist, haben wir für jede Schrittweite $h > 0$ eine Funktion $G : X \rightarrow X$, sodass für alle $t \in \mathbb{R}$ die Gleichung $f(t+h) = G(f(t))$ gilt. Die Funktion G berechnet also

den Wert nach Zeiteinheit h und stellt eine Diskretisierung der Differentialgleichung dar. Die Lösungen der Rekursion für $g : \mathbb{N} \rightarrow X$, $\forall t \in \mathbb{N} : g(t+1) = G(g(t))$ sind Folgen $(f(t_0), f(t_0 + h), f(t_0 + 2h), \dots)$ von Funktionswerten der Lösung der Differentialgleichung.

- Für kleine numerische Werte von h kann ein G wie oben approximativ berechnet werden. Auf diese Weise erhält man numerische Verfahren zum Auswerten von Lösungen von Anfangswertproblemen.
- Wenn die Lösungen einer Differentialgleichung bei 0 analytisch sind – das heißt, dass alle Ableitungen bei 0 existieren und die Taylorreihe in einem geeigneten Konvergenzbereich gegen die Funktion konvergiert –, dann bilden die Ableitungen bei 0 eine Folge. Diese Folge erfüllt eine Rekursion, die sich oft direkt aus der Differentialrechnung berechnen läßt.
- Umgekehrt kann man aus einer Folge, die eine Rekursion löst, eine Potenzreihe bilden und hoffen, daß die Potenzreihe konvergiert. Wenn ja, dann erfüllt sie oft eine Differentialgleichung, die sich direkt aus der Rekursion ableiten läßt.
- Bei der qualitativen Beschreibung von Lösungen von Differentialgleichung für $t \rightarrow \infty$ gibt es eine wichtige Methode, die wir noch diskutieren werden: man bildet eine Durchschnittsmenge $H \subset X$ und untersucht die Folge der Schrittpunkte einer Lösung in der zeitlichen Reihenfolge. Diese Folge erfüllt ebenfalls eine Rekursion, und qualitative Eigenschaften dieser Rekursion hängen eng zusammen mit qualitativen Eigenschaften der zu untersuchenden Differentialgleichung.

2 Skalare autonome Differentialgleichungen

Hier ist $T = \mathbb{R}$. Gegeben ist eine stetige Funktion $F : T \rightarrow X$, und gesucht ist $f : T \rightarrow \mathbb{R}$ differenzierbar, sodaß

$$\forall t \in \mathbb{R} : f'(t) = F(f(t)) \tag{6}$$

gilt. Bevor wir uns konkreten Lösungsverfahren zuwenden, wollen wir versuchen, qualitativ etwas über die Lösung auszusagen ohne die Lösung tatsächlich zu berechnen. Numerische und exakte Methoden zur Lösung werden später noch besprochen.

Das asymptotische Verhalten der Lösung wird durch das Vorzeichen von F bestimmt: wenn für ein $t_0 \in T$ der Wert $F(f(t_0))$ positive bzw. negativ ist, dann ist die Funktion bei t_0 streng monoton steigend bzw. fallend. Wenn die Menge der Nullstellen von F diskret ist, können wir X unterteilen in offene Intervalle, in denen F positiv oder negativ ist, und deren Randpunkte, nämlich die Nullstellen.

Es sei zunächst $x_0 \in X$ eine Nullstelle von F . Dann ist die konstante Funktion $f : t \rightarrow x_0$ eine Lösung der Differentialgleichung. Mit den Begriffen, die im nächsten Abschnitt eingeführt werden, wird der Punkt x_0 als “Gleichgewichtspunkt” oder “Equilibrium” bezeichnet.

Beispiel 2.1. Es sei $F : \mathbb{R} \rightarrow \mathbb{R}$ die “logistische Funktion” $x \mapsto x(1-x)$. Die entsprechende Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$\forall t : f'(t) = f(t)(1 - f(t))$$

wird verwendet zur Beschreibung einer Population. Solange die Bevölkerung $f(t)$ im Verhältnis zu 1 klein ist, ist die Änderungsrate proportional zu x . Wenn $f(t)$ aber in die Nähe des Schwellwerts 1 kommt, wird das Wachstum kleiner und kann für $f(t) > 1$ auch negativ werden.

Die Differentialgleichung besitzt genau zwei Equilibrien, bei denen die Bevölkerung konstant bleibt, nämlich $f_1(t) = 0$ und $f_2(t) = 1$.

Ein wenig schwieriger, aber immer noch leicht zu verstehen, ist die Situation im folgenden Satz:

Satz 2.2. *Es sei $x_0 \in X$ eine Nullstelle von F . Es sei (a, x_0) ein Intervall, in dem F positiv ist ($a = -\infty$ ist erlaubt). Es sei (x_0, b) ein Intervall, in dem F negativ ist. Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Lösung der Differentialgleichung, sodaß $f(0) \in (a, b)$ ist. Dann gilt $\lim_{t \rightarrow +\infty} f(t) = x_0$ (insbesondere existiert der Grenzwert).*

Beweis. Wir definieren die Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto (f(t) - x_0)^2$. Weil f differenzierbar ist – schliesslich ist f ja eine Lösung der Differentialgleichung, ist auch g differenzierbar. Außerdem gilt für alle t

$$g'(t) = 2f'(t)(f(t) - x_0) = 2F(f(t))(f(t) - x_0),$$

und dieser Wert ist in einer Umgebung von $t = 0$ negativ: wenn nämlich $(f(t) - x_0)$ positiv ist, ist $F(f(t))$ negativ, und wenn $(f(t) - x_0)$ negativ ist, ist $F(f(t))$ positiv. Daher ist die Funktion bei 0 monoton fallend, das heißt der Abstand des Funktionswertes zum Equilibrium wird kleiner. Es kann dann auch nicht passieren, daß der Wert von f aus dem Intervall herauskommt, und die Funktion g ist auf ganz \mathbb{R}_+ monoton fallend. Dann ist auch f monoton steigend oder fallend, je nachdem ob $f(0)$ größer oder kleiner als x_0 ist. Daraus folgt auch, daß der Grenzwert $x_1 := \lim_{t \rightarrow +\infty} f(t)$ existiert. Wir müssen noch zeigen, daß $x_1 = x_0$ ist.

Da f bei x_1 eine waagrechte Asymptote hat, gilt

$$0 = \lim_{t \rightarrow +\infty} f'(t) = \lim_{t \rightarrow +\infty} F(f(t)) = F(\lim_{t \rightarrow +\infty} f(t)) = F(x_1).$$

Die einzige Nullstelle von F , die noch näher bei x_0 liegt als $f(0)$, ist aber x_0 selbst. □

Im obigen Beispiel sind die Voraussetzungen für das Equilibrium $x_0 = 1$ erfüllt, daher konvergiert der Funktionswert für alle Startwerte in $(0, \infty)$ gegen 1.

3 Stabile und Asymptotisch Stabile Equilibrien

Im Satz (2.2) wird eine Situation beschrieben, die man gerne verallgemeinern möchte: “stabile Equilibrien” sollten Umgebungen besitzen, sodaß Lösungen mit Startwert in der Umgebung nicht mehr aus dieser Umgebung ausbrechen und schlussendlich gegen das Equilibrien konvergieren. Leider läßt sich so eine Umgebung, aus der es kein Entkommen mehr gibt, nicht immer (und wenn, dann auch nur schwer) nachweisen. Es erweist sich, daß man sich das Leben langfristig leichter macht, wenn man eine zusätzliche “Analysis-Klausel” in die Definition einbaut.

Es sei $n > 0$. Es sei X eine offene Menge in \mathbb{R}^n und $x_0 \in X$ ein Punkt. Es sei $F : X \rightarrow \mathbb{R}^n$ ein stetiges Richtungsfeld, welches ein kontinuierliches dynamisches System durch die Differentialgleichung $\forall t : f'(t) = F(f(t))$ für $f : \mathbb{R} \rightarrow X$ beschreibt.

Definition 3.1. Wenn $F(x_0) = 0$ ist, dann nennen wir x_0 ein Equilibrium für das dynamische System beschrieben durch F .

Wie schon erwähnt, ist für jedes Equilibrium x_0 die konstante Funktion $f(t) = x_0$ eine Lösung der Differentialgleichung mit Anfangswert x_0 .

Definition 3.2. Es sei x_0 ein Equilibrium. Wir nennen es ein stabiles Equilibrium, wenn für jede Umgebung U von x_0 eine Umgebung V von x_0 existiert, sodaß für jede Lösung f der Differentialgleichung mit Startwert $f(0) \in V$ ein $t_1 > 0$ existiert, sodaß $f(t) \in U$ gilt für alle $t > t_1$.

Ein Beispiel, bei dem die Analysis-Klausel schlagend wird, ist das dynamische System auf $X = \mathbb{R}^1$ gegeben durch die Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$, $F(t) = t \sin(1/t)$ (für $t \neq 0$ definiert, in den Wert $t = 0$ stetig durch $F(0) = 0$ fortgesetzt). Die Equilibrien sind 0 und die Zahlen $\frac{1}{k\pi}$, $k \in \mathbb{Z} \setminus \{0\}$. Für ungerade k ist die Voraussetzung von Satz (2.2) erfüllt, und diese Gleichgewichtspunkte sind daher auch stabil. Für gerade k haben wir eine zeitliche Umkehrung der Stabilität: die Lösungen in der Nähe driften vom Gleichgewichtspunkt weg. Diese sind also nicht stabil. Der interessanteste Gleichgewichtspunkte ist der Nullpunkt. In jeder Umgebung dieses Gleichgewichtspunkts befinden sich sowohl stabile als auch instabile Equilibrien vom vorigen Typ. Jede nichtkonstante Lösung wird von einem stabilen Equilibrium eingefangen, und zwar von einem der nicht weit weg vom Startwert ist (zumindest durch kein anderes Equilibrium getrennt). Für jede Umgebung U von 0 können wir daher eine Umgebung von V angeben, sodass U sowohl V als auch alle Equilibrien enthält, die mit die Lösungen mit Startwert in V einfangen. Mit anderen Worten: 0 ist ein stabiles Equilibrium.

Der Begriff Stabilität drückt zwar das “Nicht-Ausbrechen” aus, aber nicht die Limes-Eigenschaft. Wir definieren daher zusätzlich:

Definition 3.3. Es sei x_0 ein stabiles Equilibrium. Wir nennen es asymptotisch stabil, wenn eine Umgebung W von x_0 existiert, sodaß jede Lösung mit Startwert in W gegen x_0 konvergiert.

Wenn die Voraussetzungen von Satz (2.2) erfüllt sind, dann haben wir ein asymptotisch stabiles Equilibrium, wie man sich leicht überlegen kann. Im Beispiel vorhin ist das stabile Equilibrium bei 0 nicht asymptotisch stabil: die Grenzwerte existieren zwar immer, aber sind in der Regel andere Equilibrien (solche mit ungeraden k). Wir haben aber noch ein viel einfacheres Beispiel eines stabilen, aber nicht asymptotisch stabilen Equilibriums, nämlich für das Richtungsfeld $F : X \rightarrow \mathbb{R}^n, x \mapsto 0$. Jeder Punkt ist ein Equilibrium, alle Equilibrien sind stabil, keines davon ist asymptotisch stabil.

In der Mathematik versuchen wir bei den Definitionen, so wenig wir möglich zu verlangen. Bei der Definition der asymptotischen Stabilität könnte man meinen, daß die Voraussetzung der Stabilität vielleicht überflüssig ist: es könnte sein, dass aus der Aussage

es existiert eine Umgebung W von x_0 , sodaß jede Lösung mit Startwert in W gegen x_0 konvergiert

schon die Stabilität von x_0 folgt. Im Fall $X \subset \mathbb{R}^1$ kann man zeigen, dass das tatsächlich der Fall ist: wenn die obige Aussage erfüllt ist, dann können sich die Nullstellen von F nicht bei x_0 häufen, sonst hätte jede mögliche Umgebung W ein weiteres Equilibrium und damit eine Lösung, die nicht gegen x_0 konvergiert. Die Voraussetzung von Satz (2.2) ist daher erfüllt, und damit haben wir Stabilität. Wir werden aber später zwei-dimensionale dynamische Systeme sehen, die ein Equilibrium haben, welches die obige Aussage erfüllt, die aber nicht stabil sind (und daher auch nicht asymptotisch stabil).

Das Phasenportrait ist eine graphische Darstellung eines dynamischen Systems, in dem die einzelnen Lösungen als Kurven in X dargestellt werden, und die Durchlaufrichtung der Kurve mit einem Pfeil bezeichnet wird. Die Kurven bzw. Punkte (im Fall von Equilibrien) nennen wir auch *Bahnen*. Voraussetzung für das Erstellen des Phasenportraits sind die Eindeutigkeit (Bahnen sind disjunkt) und zumindest lokale Existenz (jeder Punkt hat eine Bahn) von Lösungen. Wenn die Voraussetzungen erfüllt sind, bilden die Bahnen eine Partition von X .

Im Fall $X \subset \mathbb{R}^1$ kann man das Phasenportrait zeichnen der Differentialgleichung für $f : T \rightarrow X, \forall t : f'(t) = F(f(t))$ zeichnen, wenn man die Nullstellen von F kennt und das Vorzeichen des Wertes von F in den Intervallen zwischen den Nullstellen. Die Nullstellen entsprechen konstanten Lösungen und werden als Punkte dargestellt. Die Intervalle sind ebenfalls Bilder von Lösungskurven. Die Richtung ist aufwärts in Intervallen, in denen F positiv ist, und abwärts in Intervallen, in denen F negativ ist.

4 Bifurkationen

Angenommen, wir haben ein dynamisches System, das stetig von einem (oder mehreren) Parameter abhängt. Welchen Einfluss haben kleine Änderungen der Parameter auf Stabilitäts-Eigenschaften der Equilibrien? Diese Fragestellung wird im Lauf der Vorlesung noch ausführlich behandelt; an dieser Stelle nur eine Vorschau auf den Begriff der Bifurkation mit einigen Beispielen.



Abbildung 1: Phasenporträt der logistischen Differentialgleichung $f'(t) = f(t)(1 - f(t))$. Der Bildraum wird besteht aus 5 Bahnen. Zwei davon sind Punkte, der Punkt 1 ist asymptotisch stabil.

In den folgenden Beispielen gehen wir von einer stetigen Funktion $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, $(\lambda, x) \mapsto F(\lambda, x)$ aus. Für ein fixes $\lambda \in \mathbb{R}$ sei F_λ die Funktion $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto F(\lambda, x)$. Diese definiert die parameterabhängige Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\forall t : f'(t) = F_\lambda(f(t)).$$

Beispiel 4.1. Wir beginnen mit $F_\lambda(x) = -x + \lambda$. Dieses System hat für jedes $\lambda \in \mathbb{R}$ genau ein Equilibrium, und zwar bei λ . Die Voraussetzungen von Satz 2.2 sind erfüllt, also ist das Gleichgewicht asymptotisch stabil. Das Equilibrium ändert zwar die Position, aber der Typ bleibt unverändert, wenn man den Parameter ändert. Dieses Phänomen werden wir später als *strukturell stabil* bezeichnen.

Beispiel 4.2. (Sattelnoden) Das nächste Beispiel ist $F_\lambda(x) = x^2 - \lambda$. Für $\lambda < 0$ gibt es keine Equilibrium. Für $\lambda > 0$ existieren zwei Equilibrien bei $\pm\sqrt{\lambda}$. Das Equilibrium bei $-\sqrt{\lambda}$ ist asymptotisch stabil, das andere nicht. Im Grenzfall $\lambda = 0$ haben wir ein Equilibrium, welches “halbseitig stabil” ist: Bahnen von unten konvergieren, Bahnen von oben nicht.

Beispiel 4.3. (transkritisch) Nun sei $F_\lambda(x) = x(x - \lambda)$. Hier gibt es für fast alle λ zwei Equilibrien bei 0 und bei λ , außer bei $\lambda = 0$, wo beide zusammenfallen. Beim Übergang von negativen zu positiven Parametern wechseln die Equilibrien den Typ: für negative λ ist das Equilibrium bei λ stabil und das bei 0 nicht. Für positive λ ist es umgekehrt.

Im Bifurkationsdiagramm verwenden wir λ und x als Koordinaten. Die Nullstelle von F in der (λ, x) -Ebene ist im allgemeinen eine Kurve. Wir zeichnen die Teile der Kurve, die stabilen Equilibrien entsprechen, mit dicker Strichstärke, und die anderen Teile strichliert. Die Regel ist: wenn unterhalb eines Kurvenzweiges die Funktion F positiv ist und oberhalb negativ, dann wird dieser Teil dick gezeichnet, sonst strichliert.

5 Lösungsverfahren

Es gibt Anekdoten von Mathematikern, die mit einem rechnerischen Problem konfrontiert werden, eine Nacht lang nachdenken und danach stolz verkünden: “ich habe gezeigt, dass

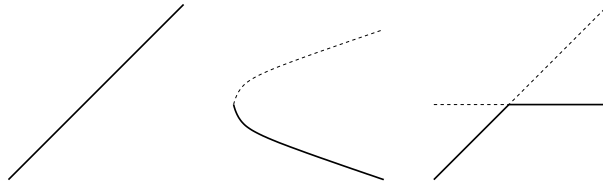


Abbildung 2: Bifurkationsdiagramme für die parameterabhängigen Differentialgleichungen $f'(t) = -f(t) + \lambda$, $f'(t) = f(t)^2 - \lambda$, $f'(t) = f(t)(f(t) - \lambda)$. Stabile Equilibrien sind fett durchgezeichnet, die anderen strichliert.

eine Lösung existiert”. Meistens bleibt der Applaus aus; man erwartet sich oft mehr. Doch was heißt das eigentlich genau, eine Gleichung lösen?

Ein eklatanter Fall ist die allgemeine Gleichung vierten Grades für $x \in \mathbb{C}$:

$$x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0,$$

wobei $a_0, a_1, a_2, a_3 \in \mathbb{C}$ gegeben sind. Nach dem Fundamentalsatz der Algebra existieren 4 Lösungen, die auch zusammenfallen können. Die Formel von Ferrara ist ein Ausdruck für die Lösung, der neben den arithmetischen Operationen $(+, -, \cdot, :)$ auch Wurzeln enthält. Die Wurzel ist im komplexen eine mehrdeutige Operation, und bei der richtigen Wahl der Auswertung bekommt man die 4 Lösungen.

Es besteht also kein Zweifel, dass die Formel von Ferrara die obige Gleichung vollständig löst. Auf der anderen Seite wird die Formel kaum verwendet zum numerischen Berechnen der Lösung, und zwar aus zwei Gründen. erstens ist die Formel kompliziert und die Wurzeln sind im komplexen nicht leicht auszuwerten (zu Ferrara’s Zeiten verwendete man dazu Winkelfunktionen). Zweitens steht ein einfaches numerisches Näherungsverfahren zur Verfügung, das schnell konvergiert, nämlich das Newton-Verfahren. Ist dann nicht die Newton-Methode gemeinsam mit dem Fundamentalsatz der Algebra, der die Existenz und Anzahl der Lösungen garantiert, nicht eine bessere “Lösung” als Ferrara’s Formel?

Im Fall von dynamischen Systemen ist nicht eine Zahl, sondern eine Funktion gesucht, das macht die Frage, was “Lösung” bedeutet, auch nicht gerade einfacher. Wenn wir uns die Option einer numerischen Lösung offenhalten wollen, liegt folgende Definition nahe: *eine Differentialgleichung für f ist gelöst, wenn wir zu einem gegebenen Anfangswert $f(0)$ und zu einer gegebenen Stelle τ die Zahl $f(\tau)$ berechnen können, sei es mit einer Formel oder mit einem numerischen Näherungsverfahren.*

Wenden wir die gleiche Definition auch für die Rekursion für $f : \mathbb{N} \rightarrow X: f(t+1) = F(f(t))$ an, dann ist die Rekursion selbst auch schon eine Lösung, weil sie als rekursiver Algorithmus gelesen werden kann.

Für Differentialgleichungen ist die Sache nicht so einfach. Sei $n \in \mathbb{N}$, $X \subset \mathbb{R}^n$, $F : X \rightarrow \mathbb{R}^n$; und wir wollen die Differentialgleichung

$$\forall t : f'(t) = F(f(t))$$

für $f : \mathbb{R} \rightarrow X$ lösen. Sei $\tau > 0$. Im Eulersche Polygonzugverfahren wählen wir zunächst $N \in \mathbb{N}$ – je größer N , desto höher ist der Rechenaufwand, aber wir erwarten auch eine höhere Genauigkeit – und unterteilen das Intervall $[0, \tau]$ in N Teilintervalle der Länge $h := \frac{\tau}{N}$. Dann definieren wir eine Rekursion für $g_N : \mathbb{N} \rightarrow X$

$$\forall t : g_N(t+1) = g_N(t) + hF(g_N(t))$$

hoffend dass die rechte Seite in X liegt. Wenn nicht, bricht das Verfahren mit einer Fehlermeldung ab (das kann nicht passieren wenn $X = \mathbb{R}^n$ ist). Als Startwert setzen wir $g_N(0) = f(0)$, den gegebenen Startwert von f . Und der Näherungswert von $f(\tau)$ ist $g_N(N)$ – allgemeiner, für alle $i \geq 0$ ist $g_N(i)$ ein Näherungswert für $f(hi)$.

Zwei schwerwiegende Einwände muss man an dieser Stelle machen. Erstens ist noch gar nicht ausgemacht, dass überhaupt ein Lösung f mit dem abgegebenen Startwert existiert. Und zweitens, selbst wenn die Lösung existiert und eindeutig ist, dann ist auch nicht klar, ob $\lim_{N \rightarrow \infty} g_N(N)$ existiert und gleich $f(\tau)$ ist. Beide Fragen werden an späterer Stelle noch behandelt werden.

6 Symbolische Lösungen

Das Beispiel der Gleichung vierten Grades war nicht typisch: mitunter sind Lösungsformeln einfacher und schneller auswertbar als numerische Verfahren (wenn man weiß wie man sie findet). Das Entwickeln von Algorithmen zum Berechnen von Lösungsformeln ist Aufgabe der mathematischen Disziplin “Symbolisches Rechnen”. Es ist natürlich nicht so, dass immer eine Lösungsformel existiert. Ob das so ist, hängt von der Sprache ab, die man vorher definieren muss und die genau festlegt was eine Formel ist.

Ein berühmtes Beispiel aus der Algebra ist die allgemeine Gleichung fünften Grades für $x \in \mathbb{C}$:

$$x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0,$$

wobei $a_0, a_1, a_2, a_3, a_4 \in \mathbb{C}$ gegeben sind. Abel/Ruffini/Galois haben gezeigt, dass es keine Lösungsformel gibt, die aus den arithmetischen Operationen und Wurzeloperationen besteht.

Ein oft verwendeter Begriff einer Formel ist die Menge der elementaren Funktionen: eine Funktion einer Teilmenge von \mathbb{C} nach \mathbb{C} heißt elementar, wenn sie sich ausdrücken lässt durch arithmetische Operationen, Exponentialfunktionen und Logarithmen (Wurzelfunktionen braucht man nicht, weil die sich mit Exponentialfunktionen und Logarithmen ausdrücken lassen). Diese Klasse von Funktionen ist relativ eingeschränkt, und viele Gleichungen sind mit diesem Formelbegriff unlösbar, etwas das Invertieren von Funktionen, das auf eine algebraische Gleichung führt. Auch Integrationsprobleme und damit Differentialgleichungen sind manchmal unlösbar, etwa die Gleichung für f

$$\forall t : f'(t) = e^{t^2}. \tag{7}$$

Die elementare Funktion $t \mapsto e^{t^2}$ hat keine elementare Stammfunktion.

Ein Satz von Lieuville gibt Aufschluss darüber, wie Stammfunktion einer elementaren Funktion ausschaun muss, falls überhaupt eine existiert. Die untenstehende Formulierung ist etwas salopp; die genaue Formulierung ist technisch aufwändig, und wir sparen sie ein, weil wir sowieso in dieser Vorlesung nicht mit diesem Satz arbeiten werden.

Satz 6.1. *Es sei f eine elementare Funktion und F eine elementare Stammfunktion. Dann treten als Argumente von Exponentialfunktionen in F nur solche Ausdrücke auf, die auch schon als Argumente von Exponentialfunktionen in f auftreten; und als Argumente von Logarithmen in F treten entweder Ausdrücke auf, die auch schon als Argumente von f in Logarithmen auftreten, oder in f im Nenner eines Bruches stehen; im zweiten Fall ist F eine lineare Funktion von diesem Logarithmus.*

Der Beweis des Satzes von Lieuville (den wir hier nicht angeben) ist konstruktiv. Er liefert einen Algorithmus, der entscheidet, ob eine gegebene elementare Funktion eine Stammfunktion hat, und berechnet diese im Fall dass sie existiert.

Ein wichtiger Spezialfall ist die Unterklasse der rationalen Funktionen, also Funktionen, die sich nur durch arithmetische Operationen darstellen lassen. Diese haben immer elementare Stammfunktionen, und nach dem Satz von Lieuville treten keine Exponentialfunktionen, nur Logarithmen von Faktoren des Nenners. Mit Hilfe der Methode der Partiabruchzerlegung kann eine elementare Stammfunktion berechnet werden (siehe Skriptum für Analysis). Würde man die Sprache erweitern und zusätzlich zu arithmetischen Operationen, Exponentialfunktion und Logarithmus auch noch unbestimmte Integrale zulassen und Funktionen, die sich in dieser Sprache schreiben lassen, etwa "intelementare Funktionen" nennen, dann hat man natürlich überhaupt kein Problem, die Gleichung (7) zu lösen: $f(x) = \int e^{x^2} dx$. Allgemein hat jede intelementare Funktion g eine intelementare Stammfunktion, nämlich $\int g(x) dx$.

6.1 Trennung der Variablen

Es seien $T \subset \mathbb{R}$ und $X \subset \mathbb{R}$ Intervalle, $a : X \rightarrow \mathbb{R}^*$ und $b : T \rightarrow \mathbb{R}$ stetige Funktionen (zur Erinnerung: wenn K ein Körper ist, dann ist K^* die Menge aller Elemente ungleich 0). Wir lösen die nichtautonome Differentialgleichung für $f : T \rightarrow X$

$$\forall t : f'(t) = \frac{b(t)}{a(f(t))}$$

durch Rückführung auf Integration und Funktionsumkehrung.

Es seien $A := \int a$ und $B := \int b$ Stammfunktionen von a und b . Da die $A'(x) \neq 0$ ist für alle $x \in X$, ist $A : X \rightarrow Y$ invertierbar, wenn Y die Bildmenge von A ist.

Satz 6.2. *In der Notation wie oben ist $A^{-1} \circ B$ eine Lösung. Umgekehrt lässt sich jede Lösung schreiben als $t \mapsto A^{-1}(B(t) + c)$ für eine geeignete Zahl $c \in \mathbb{R}$.*

Beweis. Es sei h eine differenzierbare Funktion von T nach X . Dann ist h genau dann eine Lösung der Differentialgleichung, wenn

$$(A \circ h)' = (A' \circ h) \cdot h' = (a \circ h) \cdot \frac{b}{a \circ h} = b$$

ist, also genau dann wenn $A \circ h$ eine Stammfunktion von b ist (zum Beispiel B). \square

Man beachte, dass die Lösung $g = A^{-1} \circ B$ nicht in ganz T definiert sein muss: damit g beim Wert t definiert ist, muss $B(t)$ in Y liegen (dem Definitionsbereich von A^{-1}). Wenn zusätzlich zur Differentialgleichung eine Anfangsbedingung $f(t_0) = x_0$ gegeben ist, wobei $t_0 \in T$ und $x_0 \in X$ ist, dann ist die Konstante $c = A(x_0) - B(t_0)$ eindeutig festgelegt. Wenn also überhaupt eine Lösung existiert, dann ist diese eindeutig.

Bemerkung 6.1. Es gibt Beispiele, bei denen das Anfangswertproblem keine Lösung hat, die auf ganz T definiert ist (etwa Übung 3, Beispiel 4a). Wenn T und X offen sind, dann gibt es aber immer lokal eine Lösung, d.h. es existiert eine Umgebung von t_0 , auf der die Lösung definiert ist. *Beweis:* Damit die Lösung $t \mapsto A^{-1}(B(t) + A(x_0) - B(t_0))$ für t definiert ist, muss $B(t) + A(x_0) - B(t_0)$ in Y liegen. Für $t = t_0$ ist das der Fall. Nachdem A streng monoton ist, ist $Y = A(X)$ ebenfalls offen. Es sei $V := \{y - A(x_0) + B(t_0) \mid y \in Y\}$; diese Menge ist auch offen. Weil B stetig ist, ist das Urbild $U := B^{-1}(V)$ wieder offen. Weil $B(t_0) \in V$ gilt, ist U eine offene Umgebung von t_0 . Für $t \in U$ ist $B(t) \in V$, daher ist $B(t) + A(x_0) - B(t_0) \in Y$, daher ist die Lösung für t definiert.

Zwei wichtige Spezialfälle sind der autonome Fall und der lineare Fall. Im autonomen Fall ist $X = \mathbb{R}$, $t_0 = 0$, und b die konstante Funktion 1. Im Anfangswertproblem bei der allgemeinen autonomen Differentialgleichung für f

$$\forall t : f'(t) = F(f(t))$$

können wir daher zwei Fälle unterscheiden: entweder ist $f(0)$ ungleich 0, dann können wir eine die Methode TdV anwenden. Oder $f(0) = 0$, dann existiert auf jeden Fall eine konstante Lösung. In einem Übungsbeispiel haben wir gesehen, dass die Lösung im zweiten Fall nicht eindeutig sein muss

Der zweite Spezialfall ist die lineare Gleichung. Hier ist $X = \mathbb{R}$ und die Gleichung für f ist

$$\forall t : f'(t) = b(t)f(t).$$

Wenn der Startwert $f(x_0) \neq 0$ ist, dann verkleinern wir X auf \mathbb{R}^* , setzen $a : X \rightarrow \mathbb{R}$ als $x \rightarrow \frac{1}{x}$ und können die allgemeine Lösung hinschreiben:

$$f(t) = e^{B(t)+c} = c_1 e^{B(t)};$$

wenn man die Stammfunktion B gleich $t \mapsto \int_{t_0}^t b(s)ds$ ist, dann ist c_1 gleich dem Startwert x_0 . Die Lösung ist auf ganz \mathbb{R} definiert und nimmt hat keine Nullstellen.

Wenn der Startwert gleich 0 ist, dann hat man die konstante Lösung $f(t) = 0$. Sie ist auch die einzige, weil alle anderen Lösungen keine Nullstellen haben.

Es sei $F : \mathbb{R} \rightarrow \mathbb{R}$ stetig. Die Euler-homogene Differentialgleichung für $f : \mathbb{R}^* \rightarrow \mathbb{R}$

$$\forall t : f'(t) = F\left(\frac{f(t)}{t}\right)$$

läßt sich durch die Substitution $f(t) = tg(t)$ auf eine Differentialgleichung für g zurückführen, für die an TdV anwenden kann (Übung).

Hier noch ein Trick, der hilft, TdV anzuwenden ohne die Formel auswendig zu lernen. Er besteht aus Umformungen der Differentialgleichung in Pseudo-Geichungen, die zwar mit einem Gleichheitszeichen ausgestattet sind, aber deren beide Seiten nicht immer sinnvolle mathematische Ausdrücke sind.

$$\begin{aligned} f'(t) &= \frac{b(t)}{a(f(t))} \\ \frac{df}{dt} &= \frac{b(t)}{a(f)} \\ a(f)df &= b(t)dt \\ \int a(f)df &= \int b(t)dt \\ A(f) &= B(t) + c \\ f &= A^{-1}(B(t) + c) \end{aligned}$$

6.2 Reduktion der Ordnung

Wir führen nun eine autonome Differentialgleichung zweiter Ordnung auf eine Differentialgleichung erster Ordnung, im allgemeinen nicht autonom. Für allgemeine Differentialgleichungen erster Ordnung gibt es keine symbolische Lösung, aber in manchen Fällen lässt sich dann TdV anwenden.

Es sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ stetigs; wir untersuchen die Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\forall t : f''(t) = F(f(t), f'(t)).$$

Wir nehmen an, dass $f'(t_0) \neq 0$ ist. Dann ist f lokal invertierbar und es existiert lokal eine Funktion $g : U \rightarrow \mathbb{R}$, U eine Umgebung von $f(t_0)$, sodass $f'(t) = g(f(t))$ gilt, wann immer die rechte Seite definiert ist. Wir leiten nun beide Seiten der Gleichung $f'(t) = g(f(t))$ ab und verwenden die Differentialgleichung:

$$F(f(t), g(f(t))) = F(f(t), f'(t)) = f''(t) = g'(f(t))f'(t) = g'(f(t))g(f(t)).$$

In dieser Gleichung kommt t nur als Argument von f vor. Wir setzen $s := f(t)$ und erhalten

$$F(s, g(s)) = g'(s)g(s).$$

Das ist eine Differentialgleichung für g . Wenn F nicht von der zweiten Variable abhängt, dann kann TdV anegewandt werden und man bekommt eine Lösung für g . Um f zu finden, muss man in einem zweiten Schritt nch die Gleichung $f'(t) = g(f(t))$ lösen; hier greift TdV ebenfalls.

Wenn man von einer linearen Differentialgleichung eine Lösung kennt, kann man die Ordnung ebenfalls reduzieren. Es sei $k \in \mathbb{N}$, $T \subset \mathbb{R}$, und $a_0, \dots, a_{k-1} : T \rightarrow \mathbb{R}$ stetig. Für die Differentialgleichung für $f : T \rightarrow \mathbb{R}$

$$\forall t : f^{(k)}(t) + a_{k-1}(t)f^{(k-1)}(t) + \dots + a_1(t)f'(t) + a_0(t)f(t) = 0 \quad (8)$$

brauchen wir eine Lösung $f_0 : T \rightarrow \mathbb{R}$, die auf T keine Nullstellen hat; nennen wir sie $f_1 : T \rightarrow \mathbb{R}$.

Wir ersetzen $f(t)$ durch $f_1(t)g(t)$ und berechnen eine Differentialgleichung für $g : T \rightarrow \mathbb{R}$. Nach Anwenden der Produktregel und Ausmultiplizieren aller Klammern werden alle Terme weggekürzt, in denen g nicht abgeleitet wird. Das Ergebnis ist eine lineare Differentialgleichung der Ordnung $k - 1$ für $g' : T \rightarrow \mathbb{R}$. Wir führen die Rechnung durch für den Fall $k = 2$:

$$\begin{aligned} 0 &= (f_1g)'' + a_1(f_1g_1)' + a_0f_1g_1 = (f_1g' + f_1'g)' + a_1f_1g' + a_1f_1'g_1 + a_0f_1g = \\ &= f_1g'' + 2f_1'g' + f_1''g + a_1f_1g' + a_1f_1'g_1 + a_0f_1g = (f_1'' + a_1f_1' + a_0f_1)g + f_1g'' + 2f_1'g' + a_1f_1g' = \\ &= f_1g'' + 2f_1'g' + a_1f_1g'. \end{aligned}$$

Nachdem f_1 keine Nullstellen hat, darf man mit $\frac{1}{f_1}$ multiplizieren und hat eine lineare Differentialgleichung erster Ordnung, bei der man wie schon erwähnt TdV anwenden kann. Wenn man die allgemeine Lösung g gefunden hat, muss man sie noch integrieren und dann mit f_1 multiplizieren.

Auch wenn man keine Lösung kennt, kann man die Ordnung reduzieren, allerdings erhält man dann eine nichtlineare Differentialgleichung. Wir nehmen an, es existiert eine Lösung f ohne Nullstellen, und erstellen eine Differentialgleichung für $g : T \rightarrow \mathbb{R}$, $g = \frac{f'}{f}$. Wenn wir g gefunden haben, kann f durch TdV gefunden werden.

Zum Auffinden der Gleichung für g schreiben wir die Ableitungen von f in der Form f mal einem Ausdruck in g und seinen Ableitungen:

$$f = f \cdot 1$$

$$f' = f \cdot g$$

$$f'' = f'g + fg' = fg^2 + fg' = f \cdot (g^2 + g')$$

$$f''' = (f \cdot (g^2 + g'))' = f'(g^2 + g') + f(g^2 + g')' = fg(g^3 + g') + f(2gg' + g'') = f \cdot (g^3 + g'' + 3gg')$$

usw.. Dann ersetzt man in Gleichung (8) alle Ableitungen von f durch die rechten Seiten der Gleichungen oben, hebt f heraus und hat die Gleichung für g . Im Fall $k = 2$ lautet das Ergebnis

$$g' + g^2 + a_1g + a_0 = 0$$

(Riccati-Differentialgleichung). Weil die Rückführung auch umgekehrt funktioniert, kann man auch eine gegebene Riccati-Differentialgleichung auf eine lineare Differentialgleichung zweiter Ordnung reduzieren, hoffend dass man eine Lösung errät und damit auf eine lineare Differentialgleichung erster Ordnung reduzieren und TdV anwenden kann.

6.3 Variation der Konstanten

Die Methode der Variation der Konstanten gibts in zwei Variationen, für vektorwertige erster Ordnung und für skalare beliebiger Ordnung. Sie dient zur Lösung einer inhomogenen linearen Differentialgleichung unter der Annahme, dass man die allgemeine Lösung der dazugehörigen homogenen Gleichung schon kennt. Im skalaren Fall für Ordnung 1 kann man bekanntlich die homogene Gleichung mit TdV lösen, also hat man durch Teamwork von TdV und VdK ein allgemeines Lösungsverfahren für inhomogene lineare Differentialgleichung erster Ordnung. Auch für autonome Differentialgleichungen werden wir Lösungsverfahren für den homogenen Fall angeben, und daher sind auch alle inhomogenen Differentialgleichungen lösbar, bei denen der lineare Anteil nicht von t abhängt.

Der vektorwertige Fall ist einfacher zu erklären, darum fangen wir damit an. Es sei $T \subset \mathbb{R}$ ein reelles Intervall, $n \in \mathbb{N}$. Es sei $A : T \rightarrow \mathbb{R}^{n \times n}$ eine matrixwertige stetige Funktion und $b : T \rightarrow \mathbb{R}^n$ eine vektorwertige Funktion. Wir lösen die inhomogene Differentialgleichung für $f : T \rightarrow \mathbb{R}^n$

$$\forall t \in T : f'(t) - A(t)f(t) = b(t)$$

unter der Annahme, dass wir das Anfangswertproblem der homogenen Gleichung für $g : T \rightarrow \mathbb{R}^n$

$$\forall t : g'(t) - A(t)g(t) = 0$$

für jeden Anfangswert bei der Anfangsstelle $t_0 \in \mathbb{R}$ lösen können. Die Annahme ist äquivalent zu der, dass wir eine Matrix B kennen, die die Matrix-Differentialgleichung

$$\forall t : B'(t) = A(t)B(t)$$

und die Anfangsbedingung $B(t_0) = I_n$ erfüllt; die Matrix $B(t)$ ist jene lineare Abbildung, die jedem Anfangswert $v_0 \in \mathbb{R}^n$ den Funktionswert $g(t)$ der Lösung zuordnet.

Alles das setzt natürlich voraus, dass das Anfangswertproblem eine eindeutige Lösung hat; auch das werden wir später für beliebige lineare Differentialgleichungen zeigen. Eine weitere Folgerung aus diesem (später zu zeigenden) Satz ist, dass die Matrix $B(t)$ für alle $t \in T$ invertierbar ist. Die Inverse von $B(t_1)$ ist die Matrix $\bar{B}(t_0)$, wobei $\bar{B}(t)$ der linearen Abbildung entspricht, die jedem Anfangswert v_1 bei t_1 den Wert $f(t)$ der eindeutigen Lösung mit Anfangswert $f(t_1) = v_1$ zuordnet.

Wir ersetzen nun in der inhomogenen Gleichung $f(t)$ durch $B(t)h(t)$ und bekommen eine Differentialgleichung für die vektorwertige Funktion $h : T \rightarrow \mathbb{R}^n$:

$$\begin{aligned} b(t) &= (B(t)h(t))' - A(t)B(t)h(t) = B(t)h'(t) + B'(t)h(t) - A(t)B(t)h(t) = \\ &= B(t)h'(t) + (B'(t) - A(t)B(t))h(t) = B(t)h'(t) \\ (B(t))^{-1}b(t) &= h'(t) \end{aligned}$$

Die Funktion h kann man jetzt durch Integration finden. Um eine Lösung für die Gleichung für f zu finden, muss man noch mit der Matrixfunktion B multiplizieren.

Wir wenden uns dem skalaren Fall beliebiger Ordnung zu. Es sei $T \subset \mathbb{R}$, $k \in \mathbb{N}$, $a_0, \dots, a_{k-1}, b : T \rightarrow \mathbb{R}$ stetige Funktionen. Wir lösen die Differentialgleichung für $f : T \rightarrow \mathbb{R}$

$$f^{(k)} + a_{k-1}f^{(k-1)} + \dots + a_1f' + a_0f = b \quad (9)$$

(eine Gleichung von Funktionen $T \rightarrow \mathbb{R}$). Wir nehmen an, dass wir k linear unabhängige Lösungen g_1, \dots, g_k der homogenen Gleichung für g

$$g^{(k)} + a_{k-1}g^{(k-1)} + \dots + a_1g' + a_0g = 0$$

schon kennen. An dieser Stelle könnten wir die Gleichung (9) schon auf zwei Arten lösen:

- Zurückführen auf eine vektorielle Gleichung erster Ordnung und VdK.
- Zurückführen auf eine homogene lineare Gleichung der Ordnung $k + 1$, dann k mal Reduktion der Ordnung mit den bekannten Lösungen bis auf eine lineare Gleichung erster Ordnung, dann TdV.

Im ersten Fall hat man eine matrixwertige Funktion zu invertieren. Im wesentlichen läuft es darauf hinaus, ein $k \times k$ Gleichungssystem für unbekannte Funktionen zu lösen. Was wir jetzt machen, ist, dieses Gleichungssystem ohne Umwege hinzuschreiben.

Wir suchen Funktionen $h_1, \dots, h_k : T \rightarrow \mathbb{R}$, sodass deren Ableitungen h'_1, \dots, h'_k die folgenden Gleichungen erfüllen:

$$\begin{aligned} g_1h'_1 + \dots + g_kh'_k &= 0 \\ g'_1h'_1 + \dots + g'_kh'_k &= 0 \\ g''_1h'_1 + \dots + g''_kh'_k &= 0 \\ &\dots \\ g_1^{(k-2)}h'_1 + \dots + g_k^{(k-2)}h'_k &= 0 \\ g_1^{(k-1)}h'_1 + \dots + g_k^{(k-1)}h'_k &= b \end{aligned}$$

Die Matrix, die der linken Seite dieses Gleichungssystems entspricht, ist genau die Matrix B (eigentlich: matrixwertige Funktion) in der vektorwertigen Differentialgleichung, auf die man die Gleichung (9) reduzieren kann. Deshalb ist die Matrix auch für alle $t \in T$ invertierbar, und das Gleichungssystem kann nach h'_1, \dots, h'_k aufgelöst werden. Die Funktionen h_1, \dots, h_k können nun durch Integration gefunden werden.

Das Gleichungssystem ist nun gerade so gewählt, dass folgendes gilt. Es sei $f := g_1h_1 + \dots + g_kh_k$. Dann ist

$$\begin{aligned} f &= g_1h_1 + \dots + g_kh_k \\ f' &= g'_1h_1 + \dots + g'_kh_k \\ f'' &= g''_1h_1 + \dots + g''_kh_k \\ &\dots \end{aligned}$$

$$f^{(k-2)} = g_1^{(k-2)} h_1 + \dots + g_k^{(k-2)} h_k$$

$$f^{(k-1)} = g_1^{(k-1)} h_1 + \dots + g_k^{(k-1)} h_k + b$$

Wir multiplizieren die erste Gleichung mit a_0 , die zweite mit a_1 usw. (letzte Zeile mit 1) und summieren alles auf. Das Ergebnis ist – nach einigen Umformungen unter Verwendung der Annahme, dass g_1, \dots, g_k die homogene Gleichung lösen – gleich b . Also erfüllt f die Differentialgleichung.

Bernoulli-Differentialgleichung. Es sei $T \subset \mathbb{R}$, $a, b : T \rightarrow \mathbb{R}$ stetige Funktionen, $\alpha \in \mathbb{N}$. Die Differentialgleichung für $f : T \rightarrow \mathbb{R}^*$

$$\forall t : f'(t) + a(t)f(t) + b(t)f(t)^n$$

ist für $n \neq 0, 1$ zwar nicht linear, aber sie läßt sich durch die Substitution $f(t) = g(t)^{\frac{1}{1-n}}$ auf eine lineare Gleichung für g zurückführen:

$$0 = \left(g^{\frac{1}{1-n}}\right)' + ag^{\frac{1}{1-n}} + bg^{\frac{n}{1-n}} = \frac{1}{1-n}g^{\frac{n}{1-n}}g' + agg^{\frac{n}{1-n}} + bg^{\frac{n}{1-n}}$$

Nach Herausheben von $g^{\frac{n}{1-n}}$ erhält man die inhomogene lineare Differentialgleichung für g

$$\frac{1}{1-n}g' + ag + b = 0.$$

Diese kann man mit VdK und TdV lösen. Wenn man g hat, dann kriegt man f durch die obige Substitution.

Symbolische Methoden im diskreten Fall. Die Frage der Lösbarkeit mit symbolischen Methoden macht auch für Rekursionen Sinn. Wir werden zum Beispiel im nächsten Abschnitt eine symbolische Lösung für lineare autonome Rekursionen angeben. Für nicht-autonome oder nichtlineare Rekursionen gibt es viele Resultate, aber auch offene Fragen. Führende Experten auf diesem Gebiet arbeiten auf der JKU (Kauers, Paule, Schneider, Koutschan).

6.4 Unlösbarkeitsresultate

Für lineare Differentialgleichungen gibt es die *Picard-Vessiot-Theorie*, die ähnlich funktioniert wie die Galois-Theorie in der Algebra. Für eine gegebene Gleichung kann man im Prinzip entscheiden, ob sich die allgemeine Lösung als elementare Funktion schreiben läßt; anders ausgedrückt, ob sich die Lösung zurückführen läßt auf Integration, Logarithmus, und Exponentialfunktion. “Im Prinzip” deshalb, weil der Algorithmus für Ordnung größer als 2 ziemlich kompliziert und rechenaufwändig ist; es gibt auch keine vollständige Implementierung.

Auf jeden Fall weiß man von einigen Differentialgleichungen, dass die allgemeine Lösung nicht elementar ist. Die einfachsten sind die Airy-Differentialgleichung für $a : \mathbb{R} \rightarrow \mathbb{R}$

$$\forall t : a''(t) - ta(t) = 0$$

und die Besselsche Differentialgleichung für $b : (0, \infty) \rightarrow \mathbb{R}$

$$\forall t : t^2 b''(t) + t b'(t) + (t^2 - \lambda^2) b(t) = 0,$$

wobei λ ein reeller Parameter ist, sodass $\lambda - \frac{1}{2}$ nicht ganzzahlig ist. Falls $\lambda - \frac{1}{2}$ ganzzahlig ist, ist die Lösung sogar elementar.

Für allgemeine (nichtlineare) Differentialgleichungen gibt es keine vergleichbare Theorie; wenn es nicht gelingt, die Gleichung zu lösen oder auf eine lineare Gleichung zurückzuführen, dann hat man keine Möglichkeit, festzustellen, ob die Lösung intelementar ist oder nicht.

7 Lineare Autonome Rekursionen

Hier ist $T = \mathbb{N}$, $X = \mathbb{R}^n$ für ein $n \in \mathbb{N}$, $A \in \mathbb{R}^{n \times n}$ eine Matrix, und die Rekursion für $f : T \rightarrow X$ ist $f(t+1) = Af(t)$. Die allgemeine Lösung ist $f(t) = A^t v$ mit einem beliebigen Anfangswert $v = f(0)$. Wir interessieren uns für das Verhalten von f für grosse t .

Was im kontinuierlichen Fall die Equilibrien sind, sind im diskreten Fall die Fixpunkte – im linearen Fall der Nullvektor und die Eigenvektoren zu 1, die zu konstanten Folgen führen. Für andere Eigenwerte bekommen wir Zyklen, d.h., periodische Lösungsfunktionen. Die Definition von Stabilität und asymptotischer Stabilität sind die gleichen wie in Definition 3.2 und in Definition 3.3, man braucht nur das Wort “Equilibrium” durch das Wort “Fixpunkt” ersetzen.

Sei $k \in \mathbb{N}$, $k > 1$. Ein k -Zyklus der Rekursion oben ist eine Lösung f , die für alle $t \in T$ die Gleichung $f(t+k) = f(t)$ erfüllt. Wir können die Begriffe Stabilität und asymptotische Stabilität auch auf k -Zyklen erweitern. Dazu verwenden wir die Tatsache, dass die k -Zyklen den Fixpunkten der Rekursion für $g : T \rightarrow X$, $g(t+1) = A^k g(t)$ entsprechen. Ein (asymptotisch) stabiler k -Zyklus entspricht einem (asymptotisch) stabiler Fixpunkt von A^k .

Es sei $n \in \mathbb{N}$. Ein n -Eck ist gegeben durch eine Folge von n Punkten $(p_1, \dots, p_n) \in (\mathbb{R}^2)^n$. Für jedes n -Eck können wir ein zweites n -Eck durch die Vorschrift “ersetze jeden Punkt durch den Mittelpunkt seiner Nachbarpunkte” definieren. Diese Vorschrift definiert eine Abbildung

$$F : (\mathbb{R}^2)^n \rightarrow (\mathbb{R}^2)^n, (p_1, \dots, p_n) \mapsto \left(\frac{p_n + p_2}{2}, \frac{p_1 + p_3}{2}, \dots, \frac{p_{n-1} + p_1}{2} \right).$$

Die Abbildung ist offensichtlich linear und entspricht einer Blockmatrix aus $n \times n$ Blöcken vom Format 2×2 , und zwar

$$A = \begin{pmatrix} 0 & \frac{1}{2}I_2 & \dots & \frac{1}{2}I_2 \\ \frac{1}{2}I_2 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \frac{1}{2}I_2 & 0 & \dots & 0 \end{pmatrix}$$

Wir iterieren die Abbildung und beobachten, was mit dem n -Eck passiert. Dazu verwenden wir das Programm <http://www.risc.jku.at/jschicho/dg/ngon.py>.

n=3. Eine Iteration liefert ein ähnliches Dreieck mit halber Seitenlänge, das auf den Kopf gestellt ist. Iterieren wir die Abbildung, so werden die Dreiecke immer kleiner und konvergieren gegen den Schwerpunkt.

Die Eigenwerte sind $+1$ mit Vielfachheit 2 und $\frac{-1}{2}$ mit Vielfachheit 4. Der dominierende Eigenwert ist $+1$, drum wird bei fortgesetzter Iteration die Dreiecke auf den Eigenraum zu $+1$ projiziert. Der ist gegeben durch $p_1 = p_2 = p_3$, und die Projektion auf diesen Eigenraum ist das entartete Dreieck, bei dem alle 3 Punkte im Schwerpunkt zusammenfallen. Diese entarteten Dreiecke sind Fixpunkte; sie sind stabil, aber nicht asymptotisch stabil.

Der Eigenraum zu $\frac{-1}{2}$ ist die Menge aller Dreiecke mit Schwerpunkt $(0,0)$. Das ist der Grund für die Beobachtung, dass die Dreiecke bei jedem Schritt auf die Hälfte verkleinert und auf den Kopf gestellt wird.

n=4. Die erste Iteration liefert ein entartetes Viereck mit $p_1 = p_3$ und $p_2 = p_4$. Ab der zweiten passiert nichts mehr, scheinbar ist eine konstante Folge erreicht worden. Im Programm wird allerdings nicht gezeigt, dass die Punkte vertauscht werden, das heißt, in Wahrheit haben wir einen 2-Zyklus $(p_1, p_2, p_1, p_2), (p_2, p_1, p_2, p_1)$. Dieser 2-Zyklus ist stabil, aber nicht asymptotisch stabil.

Die Matrix hat die Eigenwerte 1 (Vielfachheit 2), 0 (Vielfachheit 4), und -1 (Vielfachheit 2). Der Eigenraum zum Eigenwert 0 wird nach einer Iteration gleich wegprojiziert. Die Summe der beiden anderen Teilräume ist genau die Menge aller Vierecke mit $p_1 = p_3$ und $p_2 = p_4$. Zweimalige Anwendung der Abbildung, eingeschränkt auf diese Menge, ist die Identität, darum der 2-Zyklus.

n=5. Die Matrix hat wieder den Eigenwert 1 mit Vielfachheit 2 (das gilt übrigens für alle n). Der Eigenraum ist wie beim Dreieck die Menge aller Fünfecke, bei denen alle fünf Punkte zusammenfallen.

Wir beschränken uns auf den Orthogonalraum, also der Menge aller Fünfecke mit Schwerpunkt $(0,0)$. Es gibt zwei Eigenwerte, nämlich $\lambda_1 := \frac{-1+\sqrt{5}}{4} \approx 0.3$ und $\lambda_2 := \frac{-1-\sqrt{5}}{4} \approx -0.8$, beide mit Vielfachheit 4. Der Eigenraum zu λ_1 ist die Menge aller affinen Bilder eines regelmäßigen Fünfecks. Der Eigenraum zu λ_2 ist das affine Bild einer anderen Figur, die man leicht sehen kann, wenn man die Abbildung oft genug iteriert und ab und zu vergrößert: weil dieser Eigenwert betragsmäßig der grössere ist, setzt sich diese Figur immer mehr durch. Weil $\lambda_2 < 0$ ist, wird die Figur bei jeder Iteration auf den Kopf gestellt.

Es ist klar, dass das asymptotische Verhalten von A^t , $t \in \mathbb{N}$, von den Eigenwerten von A abhängt. Aus der linearen Algebra ist bekannt, dass sich A schreiben lässt als $A = BJB^{-1}$, wobei B invertierbar und A eine Jordan-Normalform ist. Für alle t gilt

$$A^t = (BJB^{-1})^t = BJB^{-1}BJB^{-1}BJ \dots JB^{-1} = BJ^tB^{-1},$$

also genügt es, dass asymptotische Verhalten von Jordan-Normalformen zu untersuchen. Sieht man sich die Formeln für die Potenzen von Jordan-Blöcken an, so sieht man, dass alle Einträge von der Form $\binom{t}{k} \lambda^{t-k}$ sind (sofern sie nicht Null sind), wobei λ der Eigenwert

in der Diagonale des Blocks ist, und k eine natürliche Zahl ist, die kleiner als die Größe des Blocks. Dieser Term konvergiert für $|\lambda| < 1$ gegen 0; für $|\lambda| = 1$ und $k = 0$ ist der Term beschränkt; in allen anderen Fällen ist der Term unbeschränkt.

Für die Potenzen der Matrizen sind daher die folgenden Fälle möglich.

- Alle Eigenwerte haben Betrag kleiner als 1. Dann gilt $\lim_{t \rightarrow \infty} A^t = \mathbf{0}$.
- Alle Eigenwerte haben Betrag kleiner oder gleich 1, und die Jordan-Blöcke, die zu den Eigenwerten mit Betrag 1 gehören, sind 1-Blöcke. Dann ist die Folge der Matrizen $(A^t)_{t \in \mathbb{N}}$ beschränkt.
- In allen anderen Fällen ist die Folge $(A^t)_{t \in \mathbb{N}}$ unbeschränkt.

Satz 7.1. a) Wenn die Folge $(A^t)_{t \in \mathbb{N}}$ gegen 0 konvergiert, dann ist der Nullvektor ein asymptotisch stabiler Fixpunkt.

b) Wenn die Folge $(A^t)_{t \in \mathbb{N}}$ beschränkt ist, dann ist der Nullvektor ein stabiler Fixpunkt.

c) Wenn die Folge $(A^t)_{t \in \mathbb{N}}$ unbeschränkt ist, dann ist der Nullvektor nicht stabil.

Beweis. In allen 3 Fällen verwenden wir die Formel $f(t) = A^t v_0$ für die Lösung f der Rekursion $\forall t : f(t+1) = Af(t)$.

b) Wenn die Folge der Matrixpotenzen beschränkt ist, dann ist auch $K := \{A^t v \mid t \in \mathbb{N}, \|v\| < 1\}$ beschränkt. Es sei R eine obere Schranke für den Maximalbetrag von Elementen in K . Multiplikation mit beliebigen Potenzen von Faktoren kann den Betrag eines Vektors höchstens ver- R -fachen. Es sei U eine offene Umgebung von $\mathbf{0}$. Dann suchen wir und eine Kugel von Radius $\epsilon > 0$, die in U enthalten ist, und wählen V als Einheitskugel von Radius $\frac{\epsilon}{R}$. Jede Folge mit Startwert in U bleibt dann in V .

a) Wenn $\lim_{t \rightarrow \infty} A^t = \mathbf{0}$ gilt, dann ist $\lim_{t \rightarrow \infty} f(t) = 0$ unabhängig vom Startwert v_0 . Das zeigt den Zusatz "asymptotisch".

c) Angenommen, der Nullvektor sei stabil. Wir wählen $U := \{v \mid \|v\| < 1\}$. Wegen der Stabilität existiert eine offene Umgebung V von $\mathbf{0}$, sodass Lösungen mit Startwert in V immer in U bleiben. Es sei ϵ der Radius einer Kugel mit Mittelpunkt $\mathbf{0}$, die in V enthalten ist. Dann kann Multiplikation mit beliebigen Matrizen den Betrag eines Vektors höchstens mit dem Faktor $\frac{1}{\epsilon}$ multiplizieren. Daraus folgt, dass die Folge der Potenzen von A beschränkt ist bezüglich der Norm

$$M \mapsto \|M\|_{op} := \sup_{v, \|v\| < 1} \|Mv\|.$$

Daraus folgt, dass f beschränkt ist durch diese Schranke $\frac{1}{\epsilon}$ mal der Norm von v_0 . □

Wir wenden uns dem skalaren Fall höherer Ordnung zu: gegeben sind $k \in \mathbb{N}$, $a_0, a_1, \dots, a_{k-1} \in \mathbb{R}$. Gesucht ist eine Folge $f : \mathbb{N} \rightarrow \mathbb{R}$, die die Rekursion

$$\forall t : f(t+k) + a_{k-1}f(t+k-1) + \dots + a_1f(t+1) + a_0f(t) = 0 \quad (10)$$

erfüllen. Die Funktion ist durch k Anfangswerte $f(0), \dots, f((k-1))$ eindeutig bestimmt. Ein klassisches Beispiel ist die Fibonacci-Rekursion $k=2$, $a_1 = a_0 = -1$. Die Lösung mit Anfangswerten $f(0) = 0, f(1) = 1$ ist die Fibonacci-Folge $(0, 1, 1, 2, 3, 5, 8, 13, 21, \dots)$.

Um das asymptotische Verhalten zu untersuchen (und nebenbei auch eine geschlossene Lösungsformel für die Rekursion zu finden), reduzieren wir auf den vektoriellen Fall erster Ordnung und erhalten die folgende Rekursion für $g : \mathbb{N} \rightarrow \mathbb{R}^k$

$$\forall t : g(t+1) = Ag(t), A = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & \dots & -a_k \end{pmatrix} \quad (11)$$

Die Matrix A ist die Begleitmatrix des Polynoms $X^k + a_{k-1}X^{k-1} + \dots + a_1X + a_0 \in \mathbb{R}[X]$. Dieses Polynom heißt auch das "charakteristische Polynom der Rekursion (10)".

Wir werden ein wenig später zeigen, dass das Minimalpolynom und das charakteristische Polynom der Begleitmatrix mit dem charakteristischen Polynom der Rekursion übereinstimmen. Aus der Gleichheit von Minimalpolynom und charakteristischem Polynom folgt, dass alle Jordan-Blöcke maximale Grösse haben, oder equivalent dazu dass alle Eigenräume eindimensional sind.

Um die Jordan-Normalform verwenden zu können, lassen wir auch komplexe Lösungen zu, also Funktionen von \mathbb{N} nach \mathbb{C} .

Der Lösungsraum der Rekursion (10) wird von den Einträgen der ersten Zeile in der der matrixwertigen Funktion $t \mapsto A^k$ erzeugt. Diese sind wieder erzeugt von den Einträgen der matrixwertigen Funktion $t \mapsto J^k$, wobei J eine Jordan-Normalform von A ist. Diese Einträge kennen wir bereits: ein Eigenwert $\lambda \neq 0$ von Vielfachheit m führt zu Einträgen $t \mapsto \binom{t}{i} \lambda^{t-i}$, $i = 0, \dots, m-1$. Im Fall $\lambda = 0$ ist λ^{t-i} für $t \leq i$ nicht definiert. Wenn 0 ein m -facher Eigenwert ist, dann nimmt man die Funktionen $\delta_0, \dots, \delta_{m-1}$, wobei $\delta_i(j)$ gleich 1 für $i = j$ und 0 sonst ist. Insgesamt sind das nicht mehr als k Funktionen, also genau die Dimension des erzeugten Vektorraums. Daher sind sie linear unabhängig und bilden eine Basis für den Lösungsraum.

Beispiel 7.2. Das charakteristische Polynom der Fibonacci-Rekursion ist $X^2 - X - 1$, mit den Nullstellen $\frac{1+\sqrt{5}}{2} \approx 1.62$ und $\frac{1-\sqrt{5}}{2} \approx -0.62$. Die allgemeine Lösung der Rekursion ist

$$f(t) = c_1 \left(\frac{1+\sqrt{5}}{2} \right)^t + c_2 \left(\frac{1-\sqrt{5}}{2} \right)^t, \quad c_1, c_2 \in \mathbb{R}.$$

Mit den Anfangswerten $f(0) = 0$ und $f(1) = 1$ bekommt man $c_1 = \frac{\sqrt{5}}{5}$, $c_2 = -\frac{\sqrt{5}}{5}$. Für große t wird der zweite Summand sehr klein. Da die Fibonacci-Zahlen ganzzahlig sind, liegt der erste Summand sehr nahe bei einer ganzen Zahl.

Im obigen Beispiel sind die Nullstellen des charakteristischen Polynoms reell. Wenn die Nullstellen nicht reell sind, dann sind die Basislösungen oben auch nicht reellwertig, sondern

komplexwertig. Es tritt aber zu jeder nichtreellen Nullstelle auch die konjugiert komplexe Zahl als Nullstelle mit der gleichen Vielfachheit auf. Damit hat man zu jeder nichtreellwertigen Basislösung g eine konjugiert komplexe Basislösung \bar{g} . Durch komplexe Linearkombination

$$h_1 = \frac{g + \bar{g}}{2}, h_2 = \frac{g - \bar{g}}{2i}$$

kann man zwei reelle Lösungen finden, die den gleichen Vektorraum wie g, \bar{g} erzeugen. Wenn man das für alle nichtreellen Basislösungen macht, kann man dadurch eine Basis von reellwertigen Lösungen erzeugen.

Beispiel 7.3. Die Rekursion für $f : \mathbb{N} \rightarrow \mathbb{R}$

$$\forall t : g(t+2) + g(t+1) + g(t) = 0$$

hat das charakteristische Polynom $X^2 + X + 1$ mit den Eigenwerten $\frac{-1+i\sqrt{3}}{2} = \cos(2\pi/3) + i \sin(2\pi/3)$ und $\frac{-1-i\sqrt{3}}{2} = \cos(2\pi/3) - i \sin(2\pi/3)$. Die komplexwertigen Basislösungen sind

$$f_1(t) = (\cos(2\pi/3) + i \sin(2\pi/3))^t = \cos(2t\pi/3) + i \sin(2t\pi/3),$$

$$f_2(t) = (\cos(2\pi/3) - i \sin(2\pi/3))^t = \cos(2t\pi/3) - i \sin(2t\pi/3).$$

(Hier haben wir die Formeln von De Moivre für die Potenzen von komplexen Zahlen verwendet.) Als reelle Basisfunktionen findet man daher $f_3(t) = \cos(2t\pi/3)$ und $f_4(t) = \sin(2t\pi/3)$. Die allgemeine reellwertige Lösung ist daher

$$f(t) = c_1 \cos(2t\pi/3) + c_2 \sin(2t\pi/3).$$

Beim asymptotischen Verhalten von Folgen, die durch lineare autonome Rekursionen definiert werden, können wir folgende Möglichkeiten feststellen.

- Wenn alle Eigenwerte Betrag kleiner als 1 haben, dann konvergiert jede Lösungsfolge gegen 0, und 0 ist ein asymptotisch stabiler Fixpunkt.
- Wenn alle Eigenwerte Betrag kleiner gleich 1 haben und wenn alle Eigenwerte mit Betrag 1 einfach sind, dann ist jede Lösung beschränkt. 0 ist ein stabiler Fixpunkt.
- Wenn ein Eigenwert Betrag grösser als 1 hat oder wenn ein Eigenwert mit Betrag 1 und Vielfachheit größer als 1 existiert, dann gibt es unbeschränkte Lösungen. 0 ist instabiler Fixpunkt. Die Lösungen wachsen aber höchstens exponential, d.h., es existieren Konstanten $C, Q \in \mathbb{R}$, sodass $\forall t : f(t) < CQ^t$ gilt.

Zum Beweis reduziert man auf den vektorwertigen Fall erster Ordnung und verwendet Satz (7.1) sowie die Tatsache, dass die Größe der Jordan-Blöcke von Begleitmatrizen gleich der Vielfachheit des zugehörigen Eigenwerts ist.

Erwähnt sei noch eine Deutung des charakteristischen Polynoms, die wir im nächsten Kapitel brauchen werden. Es sei \mathcal{F} die Menge aller Funktionen $\mathbb{N} \rightarrow \mathbb{R}$ (bzw. aller Folgen

in \mathbb{R}), die höchstens exponentiell wachsen. Diese Menge bildet mit gliedweiser Addition und Skalarmultiplikation einen \mathbb{R} -Vektorraum. Die Funktion

$$S : \mathcal{F} \rightarrow \mathcal{F}, f \mapsto (t \mapsto f(t+1))$$

heißt Shift-Operator: sie schmeißt das erste Folgenglied weg, und alle anderen rücken um 1 nach. Sei nun $P \in \mathbb{R}[X]$ ein normiertes Polynom vom Grad k , $P = a_0 + a_1X + \dots + a_{k-1}X^{k-1} + X^k$. Wir definieren den linearen Operator

$$P(S) : \mathcal{F} \rightarrow \mathcal{F}, f \mapsto a_0f + a_1S(f) + \dots + a_{k-1}S^{k-1}(f) + S^k(f).$$

Die linke Seite der Rekursion (10) ist gleich dem t -ten Glied der Folge $P(S)(f)$, und $f : \mathbb{N} \rightarrow \mathbb{R}$ ist genau dann eine Lösung wenn sie im Kern von $P(S)$ ist. Das charakteristische Polynom einer Rekursion liefert also genau jenes Polynom, in das man den Shift-Operator einsetzen muss, um die linke Seite der Rekursionsgleichung zu erhalten.

Satz 7.4. *Das charakteristische Polynom der Rekursion (10) ist gleichzeitig das Minimalpolynom und das charakteristische Polynom der Matrix A in der Rekursion (11).*

Beweis. Es sei $f \in \mathcal{F}$. Wir definieren die Prolongation $\rho(f)$ als das k -Tupel $(f, S(f), S^2(f), \dots, S^{k-1}(f)) \in \mathcal{F}^k$. Die Funktion f erfüllt die Gleichung $A \cdot \rho(f) = S(\rho(f))$ genau dann, wenn f eine Lösung der Rekursion (10) ist (wir haben die Matrix gerade so definiert, sodass diese beiden Rekursionen equivalent sind).

Es sei $M \in \mathbb{R}[X]$ das Minimalpolynom der Matrix A . Für alle Polynome $P \in \mathbb{R}[X]$ und für jede Lösung f gilt die Gleichung $P(A)(\rho(f)) = \rho(P(S)(f))$; das kann man für Monome durch Induktion zeigen, und für beliebige Polynome folgt es dann aus der Linearität der Gleichung. Daher ist für jede Lösung f der Rekursion

$$\mathbf{0} = M(A) \cdot \rho(f) = \rho(M(S)(f)).$$

Wenn wir von dieser vektoriellen Gleichung nur den ersten Eintrag nehmen, erhalten wir $M(S)(f) = 0$. Das ist aber eine Rekursion der Ordnung $k' := \deg(M)$. Da der Lösungsraum der Rekursion (10) Dimension k hat, muss $k = k'$ gelten. Es folgt, dass Minimalpolynom und charakteristisches Polynom der Matrix gleich sind.

Nachdem das charakteristische Polynom Q der Rekursion (10) ebenfalls ein Polynom von Grad k ist und $Q(S)(\rho(f))$ gilt für alle Lösungen f , sind alle drei Polynome gleich. \square

8 Lineare Autonome Differentialgleichungen

Zur Abwechslung beginnen wir mit dem skalaren Fall höherer Ordnung. Gegeben sind $k \in \mathbb{N}$, $a_0, \dots, a_{k-1} \in \mathbb{R}$. Gesucht ist $f : \mathbb{R} \rightarrow \mathbb{R}$, sodass

$$\forall t : f^{(k)}(t) + a_{k-1}f^{(k-1)}(t) + \dots + a_1f'(t) + a_0f(t) = 0 \quad (12)$$

gilt. Beim Anfangswertproblem (AWP) gegeben wir noch zusätzlich numerische Werte von $f(0), f'(0), \dots, f^{(k-1)}(0)$ vor. Das charakteristische Polynom der Differentialgleichung (12) ist das Polynom $X^k + a_{k-1}X^{k-1} + \dots + a_1X + a_0$.

Beispiel 8.1. Es sei $a \geq 0$ und $b > 0$. Die Gleichung für $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\forall t : f''(t) + af'(t) + bf(t) = 0$$

heißt *Schwingungsgleichung*. Sie beschreibt die Schwingung eines Federpendels oder die die Dynamik eines elektrischen Schwingkreises bestehend aus einem Kondensator, einer Spule und einem Widerstand.

Das charakteristische Polynom ist $X^2 + aX + b$.

Ähnlich wie im diskreten Fall definieren wir die Menge \mathcal{D} aller beliebig oft differenzierbaren Funktionen von \mathbb{R} nach \mathbb{R} , der bezüglich punktweiser Addition und Skalarmultiplikation ein reeller Vektorraum ist. Der Operator $D : \mathcal{D} \rightarrow \mathcal{D}$ ist definiert als die Ableitung. Wie im diskreten Fall können wir dann die Differentialgleichung auch schreiben als

$$P(D)(f) = 0,$$

wobei die Null auf der rechten Seite die Nullfunktion ist und P das charakteristische Polynom ist.

Mit der Methode des Potenzreihenansatzes werden wir eine Lösung des AWP konstruieren und zeigen, dass es keine andere Lösung geben kann, die an der Stelle 0 *analytisch* ist. Eine Funktion heißt an einer Stelle analytisch, wenn sie dort beliebig oft differenzierbar ist, wenn der Konvergenzradius der Taylorreihe positiv ist, und wenn die Funktion innerhalb des Konvergenzintervalls mit der Taylorreihe übereinstimmt.

Wenn $f : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion ist, die an der Stelle 0 analytisch ist, wollen wir mit $\text{tay}(f)$ die Folge $(f^r)_{r \in \mathbb{N}}$ der Ableitungen bei 0 bezeichnen. Man kann zeigen, dass $\text{tay}(f)$ in \mathcal{F} ist (Übung). Wenn umgekehrt $c = (c_r)_{r \in \mathbb{N}}$ eine Folge in \mathcal{F} ist, dann ist der Konvergenzradius der Potenzreihe $\sum_{r=0}^{\infty} c_r \frac{t^r}{r!}$ gleich $+\infty$. Wir wollen die Funktion von \mathbb{R} nach \mathbb{R} , die durch diese überall konvergente Potenzreihe definiert ist, mit $\text{pow}(c)$ bezeichnen.

Nach dem Satz von Taylor sind die Funktionen $\text{tay} : \mathcal{D} \rightarrow \mathcal{F}$ und $\text{pow} : \mathcal{F} \rightarrow \mathcal{D}$ invers zueinander, wenn wir sie geeignet einschränken: $\tau \circ \text{pow} = \text{id}_{\mathcal{D}}$ ohne Einschränkung, und wenn $f \in \mathcal{D}$ analytisch bei 0 ist, gilt $\text{pow}(\text{tay}(f)) = f$.

Satz 8.2. *Es sei $k \in \mathbb{N}$ und $P \in \mathbb{R}[X]$ ein normiertes Polynom vom Grad k . Dann ist der Kern von $P(D)$ – also die Menge der Lösungen der Differentialgleichung, deren charakteristisches Polynom gleich P ist – gleich dem pow -Bild des Kerns von $P(S)$ – also der Menge der Lösungen der Rekursion, deren charakteristisches Polynom gleich P ist.*

Beweis. Weil man Potenzreihen innerhalb des Konvergenzbereichs gliedweise differenzieren kann, ist $\text{pow} \circ S = D \circ \text{pow}$. Durch Induktion nach r zeigt man

$$\text{pow} \circ S^r = \text{pow} \circ S^{r-1} \circ S = D^{r-1} \circ \text{pow} \circ S = D^{r-1} \circ D \circ \text{pow} = D^r \circ \text{pow},$$

und wegen der Linearität von pow gilt daher $\text{pow} \circ P(S) = P(D) \circ \text{pow}$. Sei $f \in \mathcal{D}$ analytisch. Dann ist f genau dann in $\ker(P(D))$, wenn $\text{tay}(f)$ im Kern von $\text{pow} \circ P(S) = P(D) \circ \text{pow}$ ist; da der kern von pow Null ist, genau dann wenn $\text{tay}(f)$ in $\ker(P(S))$ ist. □

Wir berechnen die Potenzreihen von den Basislösungen:

$$\lambda \neq 0 : \text{pow} \left(\binom{m}{i} \lambda^{m-i} t^m \right) (t) = \sum_{m=0}^{\infty} \frac{\binom{m}{i} \lambda^{m-i} t^m}{m} = \frac{t^i e^{\lambda t}}{i}$$

$$\lambda = 0 : \text{pow} (\delta_i) (t) = t^i$$

Den konstanten Faktor $\frac{1}{i}$ werden wir natürlich weglassen. Dann kriegen wir als Basislösungen für eine Nullstelle mit Vielfachheit r die Basislösungen

$$t \mapsto e^{\lambda t}, t \mapsto t e^{\lambda t}, \dots, t \mapsto t^{r-1} e^{\lambda t}$$

unabhängig davon, ob λ Null ist oder nicht. Wenn λ nicht reell, also $\lambda = a + bi$ mit $b \neq 0$, dann ersetzen wir die komplexen Basislösungen

$$t \mapsto t^r e^{(a+bi)t}, t \mapsto t^r e^{(a-bi)t}$$

durch die reellen Lösungen

$$t \mapsto t^r \frac{e^{(a+bi)t} + e^{(a-bi)t}}{2} = t^r e^{at} \cos(bt), t \mapsto t^r \frac{e^{(a+bi)t} - e^{(a-bi)t}}{2i} = t^r e^{at} \sin(bt).$$

Mit diesen Basislösungen können wir das Stabilitätsverhalten des Equilibriums beim Nullvektor analysieren.

Satz 8.3. *Wenn die Nullstellen des charakteristischen Polynoms Realteil kleiner als 0 haben, dann ist 0 ein asymptotisch stabiles Equilibrium.*

Wenn die Nullstellen des charakteristischen Polynoms Realteil kleiner gleich 0 haben und die Nullstellen mit Realteil Null einfach sind, dann ist 0 ein stabiles Equilibrium.

In allen anderen Fällen ist 0 ein instabiles Equilibrium.

Proof. Beweis Für die Basislösungen gilt: die Funktion $t \mapsto t^r e^{\lambda t}$ ist für $|\lambda| > 1$ und für $(|\lambda| = 1 \text{ und } r \geq 1)$ unbeschränkt, für $|\lambda| = 1$ und $r = 0$ beschränkt, und für $|\lambda| < 1$ konvergent mit Grenzwert 0. Daraus folgt die Behauptung ähnlich wie im diskreten Fall. \square

8.1 Eine Alternative zu VdK

Es sei $P \in \mathbb{R}[X]$ ein normiertes Polynom vom Grad k . Bei der linearen inhomogenen Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}$

$$\forall t : P(D)(f)(t) = b(t), \tag{13}$$

– wobei b eine gegebene stetige Funktion von \mathbb{R} nach \mathbb{R} ist –, kennen wir eine Basis des Lösungsraums. Mit TdV könnte man die Gleichung lösen. Das kostet die Lösung eines $k \times k$ Gleichungssystems mit Funktionen-Koeffizienten und k Integrale.

Wenn die rechte Seite b selbst eine Lösung einer linearen autonomen Differentialgleichung ist, zum Beispiel $Q(D)(b) = 0$ für ein Polynom $Q \in \mathbb{R}[X]$ vom Grad $l \in \mathbb{N}$, dann geht es auch schneller. Dazu wenden wir den Operator $Q(D)$ auf beide Seiten von (13) an und erhalten eine homogene autonome Gleichung der Ordnung $k + l$:

$$Q(D)P(D)(f) = (QP)(D) = \mathbf{0}.$$

(Das Einsetzen von D in ein Polynom ist ein Ringhomomorphismus.) Wir schreiben unsere Basis für den Kern von $(QP)(D)$ hin. Die $k + l$ Basisfunktionen sind günstigerweise so gewählt, dass die k Basisfunktionen von $P(D)$ alle enthalten sind. Diese streichen wir aus und setzen eine allgemeine Linearkombination der l übrigen in die Gleichung (13) ein. Durch Koeffizientenvergleich bekommen wir ein Gleichungssystem. Wir lösen es und haben eine spezielle Lösung von (13). Um die allgemeine Lösung zu erhalten, addieren wir noch eine allgemeine Lösung des homogenen Gleichungssystems. Kostenfaktor: Ableitungen, Koeffizientenvergleich, und Lösung eines $l \times l$ -Gleichungssystems mit konstanten Koeffizienten.

Wenn die charakteristischen Polynome von P und Q teilerfremd sind, dann kann man sogar noch die Lösung des linearen Gleichungssystems einsparen und statt dessen einen erweiterten Euklidischen Algorithmus für P und Q durchführen. Dieser liefert Polynome $A, B \in \mathbb{R}[X]$ von Grad $\deg(A) = l$ und $\deg(B) = k$, sodass $AP + BQ = 1$ ist. Die Lösung von (13) ist dann einfach $A(D)(b)$. Hier ist der Beweis:

$$P(D)A(D)(b) = (PA)(D)(b) = (1 - BQ)(D)(b) = b - (BQ)(D)(b) = b - B(D)Q(D)(b) = b.$$

Man beachte, dass der lineare Operator $\mathcal{D} \rightarrow \mathcal{D}$, der dem konstanten Polynom 1 entspricht, der Identitätsoperator ist.

8.2 Die Exponentialmatrix

Es sei $n \in \mathbb{N}$ und $A \in \mathbb{R}^{n \times n}$. Bei der Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}^n$

$$\forall t : f'(t) = Af(t)$$

führt der Potenzreihenansatz zur Folge der Ableitungen $f^{(k)}(0) = A^k f(0)$ und damit zur analytischen Lösung

$$f(t) = \sum_{r=0}^{\infty} \frac{A^r t^r}{r!} f(0).$$

Die matrixwertige Funktion $t \mapsto \sum_{r=0}^{\infty} \frac{A^r t^r}{r!}$, die zu jedem Anfangswert die Auswertung der Lösung an der Stelle t liefert, heißt *Exponentialfunktion* der Matrix A ; der Funktionswert wird notiert mit e^{At} oder $\exp(At)$. Die folgenden Gesetze gelten für alle $A, B \in \mathbb{R}^{n \times n}$ und $s, t \in \mathbb{R}$:

1. $\frac{de^{At}}{dt} = Ae^{At} = e^{At}A.$

2. $e^{As}e^{At} = e^{A(s+t)}$.
3. Wenn $AB = BA$, dann ist $e^{(A+B)t} = e^{At}e^{Bt} = e^{Bt}e^{At}$.
4. Wenn B invertierbar ist, dann ist $e^{BAB^{-1}t} = Be^{At}B^{-1}$.

Die letzte Identität können wir verwenden, um die Exponentialmatrix mit der Darstellung in Jordan-Normalform zu berechnen. Für Jordan Blockmatrizen $J_r = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in \mathbb{R}^{r \times r}$ haben wir die Exponentialmatrix

$$e^{J_r t} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2!}e^{\lambda t} & \cdots & \frac{t^{r-1}}{(r-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & \cdots & \frac{t^{r-2}}{(r-2)!}e^{\lambda t} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & e^{\lambda t} \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{r-1}}{(r-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{r-2}}{(r-2)!} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Die Exponentialfunktion von Jordan-Normalformen hat man damit eine Formel (blockweise Anwendung der Formel oben). Für beliebige Matrizen $A = BJB^{-1}$ erhalten wir $e^{At} = Be^{Jt}B^{-1}$.

Für das asymptotische Verhalten spielt die Exponentialfunktion eine ähnliche Rolle wie die Folge der Matrixpotenzen im diskreten Fall. Die entscheidene Eigenschaft für die Eigenwerte ist aber nicht, ob der Betrag kleiner als 1 ist, sondern ob der Realteil kleiner als Null ist – genau dann konvergiert die Exponentialfunktion $t \mapsto e^{\lambda t}$ gegen Null.

Satz 8.4. *Wenn alle Eigenwerte von λ negativen Realteil haben, dann ist der Nullvektor ein asymptotisch stabiles Equilibrium.*

Wenn alle Eigenwerte von λ nichtpositiven Realteil haben und wenn die Jordan-Blöcke zu Eigenwerten mit Realteil Null alle 1×1 -Blöcke sind, dann ist der Nullvektor ein stabiles Equilibrium.

In allen anderen Fällen ist der Nullvektor ein instabiles Equilibrium.

Der Beweis ist ähnlich wie im diskreten Fall.

9 Der Fundamentalsatz der Algebra

Wir werden später beliebige (nichtlineare) Differentialgleichungen für $f : \mathbb{R} \rightarrow X \subset \mathbb{R}^n$

$$\forall t : f'(t) = F(f(t))$$

qualitativ beschreiben, wobei $F : X \rightarrow \mathbb{R}^n$ eine differenzierbare Funktion ist. Die Equilibrien sind die Nullstellen von F . In der Nähe der Equilibrien kann man F durch seine

Ableitung approximieren. Ersetzt man F in der Differentialgleichung durch diese Approximation, dann erhält man die linearisierte Differentialgleichung für $g : \mathbb{R} \rightarrow \mathbb{R}^n$

$$\forall t : g'(t) = Ag(t), \quad A = \frac{\partial F}{\partial x}(\mathbf{0}).$$

Man hätte gerne, dass die Lösungen der Linearisierungen nahe bei Lösungen der Ausgangsgleichung sind und sich asymptotisch gleich verhalten. Das ist nicht immer der Fall, und wir werden später genaue Bedingungen dafür angeben.

Als Vorbereitung für diese spätere Untersuchung ist es notwendig, zu bestimmen, wie sich das Stabilitätsverhalten ändert, wenn die Matrix A geringfügig gestört wird. Das Stabilitätsverhalten wird von den Nullstellen des charakteristischen Polynomes bestimmt (und von der Größe der Jordan-Blöcke). Wir würden deshalb gerne ein Resultat verwenden, das besagt, dass die Nullstellen eines Polynoms stetig von den Koeffizienten abhängen. Das kann natürlich nur für den komplexen Fall stimmen: im reellen Fall macht selbst die Anzahl der Nullstellen Sprünge. Aber auch im komplexen Fall gibt es Schwierigkeiten: es lässt sich keine stetige Funktion finden, die jedem Polynom stetig einen Vektor von Nullstellen zuordnet; nicht einmal für Grad 2, wo es sogar eine Lösungsformel gibt. Das Problem ist die Mehrdeutigkeit der Wurzeloperation im komplexen.

Zunächst soll sichergestellt werden, dass die Anzahl der Nullstellen, mit Vielfachheit gezählt, gleich dem Grad des Polynoms ist. Das ist der Fundamentalsatz der ALgebra. Wir benötigen ein Lemma, das die Nullstellen eines Polynoms beschränkt.

Lemma 9.1. *Es sei $p := X^d + b_1X^{d-1} + \dots + b_d$ ein Polynom von Grad d , $b_1, \dots, b_d \in \mathbb{C}$. Es sei $\epsilon > 0$. Es sei $M \in (0, \infty)$ eine reelle Zahl, sodass $|b_i| < \frac{M^i}{d+1}$ für $i = 1, \dots, d$ und $\epsilon < \frac{M^d}{d+1}$ gilt. Dann gilt für jedes z mit $|z| > M$ die Ungleichung $|p(z)| > \epsilon$. Insbesondere hat p keine Nullstelle mit Betrag größer als d .*

Proof. Es sei $|z| > M$. Dann gelten für $i = 1, \dots, d$ die Ungleichung $|b_i z^{d-i}| < \frac{|z^d|}{d+1}$ und obendrein $\epsilon < \frac{|z^d|}{d+1}$. Aus der Dreiecksungleichung folgt

$$|p(z)| = |z^d + \sum_{i=1}^d b_i z^{d-i}| > |z^d| - \sum_{i=1}^d |b_i z^{d-i}| > |z^d| - d \frac{|z^d|}{d+1} = \frac{|z^d|}{d+1} > \epsilon.$$

□

Satz 9.2 (Fundamentalsatz der Algebra). *Jedes normierte Polynom $p \in \mathbb{C}[X]$ von Grad d lässt sich schreiben als Produkt von d Linearfaktoren*

$$p = (X - a_1)(X - a_2) \cdots (X - a_d)$$

mit $a_1, \dots, a_d \in \mathbb{C}$.

Beweis. Es reicht zu zeigen, dass p eine Nullstelle hat. Wenn a eine Nullstelle ist, dann ist $(z - a)$ ein Faktor, und der Quotient ist von kleinerem Grad. Mit Induktion kann man die Behauptung zeigen.

Wir nehmen indirekt an, dass p keine Nullstellen hat. Die Menge $\{z \mid p(z) \leq p(0)\}$ ist abgeschlossen wegen Lemma 9.1 beschränkt, also kompakt. Also nimmt die Funktion $z \mapsto |p(z)|$ auf dieser Menge ein Minimum an, und dieses Minimum ist auch das globale Minimum der Funktion $z \mapsto |p(z)|$ auf ganz \mathbb{C} . Es sei z_0 eine Minimalstelle. Die Taylorentwicklung von p um z_0 sei

$$p(z_0 + w) = b_0 + b_1w + b_2w^2 + \dots + b_dw^d.$$

Es gilt $b_0 \neq 0$, sonst wäre z_0 eine Nullstelle, und $b_d = 1$. Es sei k der kleinste positive Index mit $b_k \neq 0$. Wir wählen nun $s > 0$, sodass die Ungleichungen $|b_i|s^i < \frac{|b_k|s^k}{d+1}$ für $i = k+1, \dots, d$ und $|b_k|s^k < |b_0|$ erfüllt sind. Als nächstes wählen wir $w \in \mathbb{C}$ mit $|w| = s$, sodass $\frac{b_k w^k}{b_0}$ reell und negativ ist. Dann ist

$$\begin{aligned} |p(z_0 + w)| &\leq |b_0 + b_k w^k| + \sum_{i=k+1}^d |b_i w^i| = |b_0| - |b_k|s^k + \sum_{i=k+1}^d |b_i|s^i \leq \\ &\leq |b_0| - |b_k|s^k + (d-k) \frac{|b_k|s^k}{d+1} = |b_0| - \frac{(k+1)|b_k|s^k}{d+1} < |b_0| = |p(z_0)|, \end{aligned}$$

ein Widerspruch. □

Wir brauchen noch eine Art Stetigkeits-Aussage für die Abhängigkeit der Nullstellen vom Polynom. Dazu definieren wir für fixed $d \in \mathbb{N}$ eine stetige Funktion z_d von \mathbb{C}^d in die Menge P_d aller normierten Polynome vom Grad d wie folgt.

$$(a_1, \dots, a_d) \mapsto (X - a_1) \cdots (X - a_d).$$

Das Polynom $z_d(a_1, \dots, a_d)$ ist also das Polynom, das als a_1, \dots, a_d als Nullstellen hat, mit der richtigen Vielfachheit.

Lemma 9.3. *Die Abbildung z_d ist offen: das Bild jeder offenen Menge ist offen.*

Beweis. Es sei $U \subset \mathbb{C}^d$ offen. Es sei V die Menge aller Permutationen von Elementen in U . Es sei $p : \mathbb{N} \rightarrow (P_d \setminus z_d(V))$ eine Folge im Komplement von $z_d(V)$, die in P_d konvergent ist. Nach dem Fundamentalsatz der Algebra ist z_d surjektiv. Es sei $n : \mathbb{N} \rightarrow \mathbb{C}^d$ eine Folge von Urbildern: für jedes $i \in \mathbb{N}$ wählen wir $n(i) \in z_d^{-1}(p(i))$. Da p eine konvergente Folge ist, ist p beschränkt. Aus Lemma 9.1 folgt, daß n beschränkt ist. Daher besitzt n eine konvergente Teilfolge $n' : \mathbb{N} \rightarrow \mathbb{C}^d$. Es sei $p' : \mathbb{N} \rightarrow P_d$ die entsprechende Teilfolge von p (d.h. $p'(i) = z_d(n'(i))$ für alle i). Die Menge V ist offen als Vereinigung (endlich vieler) offener Mengen. Das Komplement von V ist daher abgeschlossen. Also ist $\lim_{i \rightarrow \infty} n'(i)$ im Komplement von V . Es folgt

$$\lim_{i \rightarrow \infty} p(i) = \lim_{i \rightarrow \infty} p'(i) = \lim_{i \rightarrow \infty} z_d(n'(i)) = z_d(\lim_{i \rightarrow \infty} n'(i)) \in z_d(\mathbb{C}^d \setminus V) = P_d \setminus z_d(V).$$

Also ist $P_d \setminus z_d(V)$ abgeschlossen und daher ist $z_d(U) = z_d(V)$ offen. □

Lemma 9.3 ist eine mathematisch präzise Formulierung dessen, was wir uns vorher als “eine Art Stetigkeits-Aussage” gewünscht haben. Es erlaubt uns, zu sagen, dass die Menge aller Polynome mit Nullstellen in einer festgelegten offenen Teilmenge von \mathbb{C} – zum Beispiel der Menge der Zahlen mit negativem Realteil – eine offene Menge ist. Wenn man ein Polynom p in dieser Menge hat, dann gibt es ein $\epsilon > 0$, sodass jedes Polynom in einer ϵ -Umgebung von p_0 wieder Nullstellen in der festgelegten offenen Teilmenge hat.

Wir wollen das Lemma nun anwenden auf charakteristische Polynome von Matrizen.

Satz 9.4. *Es sei $n \in \mathbb{N}$ und $a, b \in \mathbb{N}$ mit $a + b = n$. Die Menge $\mathcal{H}_{a,b} \subset \mathbb{R}^{n \times n}$ aller Matrizen, die a Eigenwerte mit negativem Realteil und b Eigenwerte mit positivem Realteil haben, wobei die Eigenwerte mit (algebraischer) Vielfachheit gezählt werden, ist eine offene Teilmenge von $\mathbb{R}^{n \times n}$.*

Beweis. Es sei A die Menge aller komplexen Zahlen mit negativem Realteil und B die Menge aller reellen Zahlen mit positivem Realteil. Die Menge $A^a \times B^b$ ist offen in \mathbb{C}^n . Nach Lemma 9.3 ist $z_n(A^a \times B^b)$ offen in P_d . Die Teilmenge $P_{a,b}$ der reellen Polynome in $z_n(A^a \times B^b)$ ist dann auch offen in \mathbb{R}^n . Die Abbildung, die jeder Matrix ihr charakteristisches Polynom zuordnet, ist stetig. Und $\mathcal{H}_{a,b}$ ist das Urbild von $P_{a,b}$ unter dieser Abbildung, also ebenfalls offen. \square

Wenn wir also eine lineare Differentialgleichung mit einer Matrix $A \in \mathcal{H}_{a,b}$ haben, dann können wir die eine ϵ -Umgebung von A finden, die immer noch in $\mathcal{H}_{a,b}$ enthalten ist. Wenn A geringfügig gestört wird, genauer gesagt, die Störung $\|A_1 - A\|$ ist kleiner als ϵ , dann haben A_1 und A die gleiche Anzahl von Eigenwerten mit negativem Realteil, und beide Matrizen haben keine Eigenwerte mit Realteil gleich Null. Es folgt, dass der Nullpunkt genau dann stabil für A ist (und wenn ja, dann immer auch asymptotisch stabil), wenn er stabil für A_1 ist. Das Stabilitätsverhalten bleibt also unter kleinen Störungen der Matrix erhalten.

Die Anzahl der Nullstellen mit positivem bzw. negativem Realteil (also a und b im Satz 9.4) bleibt ebenfalls erhalten unter kleinen Störungen (unter der Voraussetzung $a + b = n$). Im Fall $n = 1$ nennen wir den Nullpunkt eine Senke für $(a, b) = (2, 0)$, eine Quelle für $(a, b) = (0, 2)$, und einen Sattelpunkt für $(a, b) = (1, 1)$. Im höherdimensionalen Fall geben die Werte von a und b nicht eine feinere Unterscheidung von Sattelpunkten.

10 Existenz und Eindeutigkeit

Das Haupt-Thema dieses Kapitels ist der Beweis der Existenz und Eindeutigkeit des Anfangswertproblems. Nach den Überlegungen in Kapitel 5 ist vielleicht klar, dass das Zeigen von Existenz und Eindeutigkeit durchaus ein wesentlicher Bestandteil des “Lösens” eines AWP sein, wenn zum Beispiel ein allgemeines Verfahren existiert, das den Wert einer Lösung berechnet *unter der Voraussetzung, dass eine solche existiert und eindeutig ist*. Hier ist die genaue Formulierung des Satzes in der Form, wie wir ihn beweisen werden.

Satz 10.1 (Picard/Lindelöf). *Es sei T eine zusammenhängende Teilmenge von \mathbb{R} (also ein Intervall, offen oder abgeschlossen, beschränkt oder unbeschränkt). Es sei $k \in \mathbb{N}$. Es sei $F : T \times \mathbb{R}^k \rightarrow \mathbb{R}^k$ eine stetige Funktion, die die Lipschitz-Bedingung erfüllt: es existiert eine Konstante $L \in \mathbb{R}$, sodaß für alle $x \in T, y, z \in \mathbb{R}^k$ gilt $\|F(x, y) - F(x, z)\| \leq L\|y - z\|$. Es sei $t_0 \in T, x_0 \in \mathbb{R}^k$. Dann hat das Anfangswertproblem für $f : T \rightarrow \mathbb{R}^k$*

$$f'(t) = F(t, f(t)), \quad f(t_0) = x_0$$

eine eindeutige Lösung $y : T \rightarrow \mathbb{R}^k$.

Wenn F die Lipschitz-Bedingung erfüllt, sagen wir, dass F Lipschitz-stetig in der zweiten Variable ist.

Bevor wir uns dem Beweis zuwenden, hier ein paar Bemerkungen zur des Reichweite des Satzes.

1. Die Lipschitz-Bedingung ist stärker als die Bedingung der Stetigkeit von F , die wir bis jetzt immer gemacht haben. Beispiele von stetigen Funktionen, die nicht Lipschitz-stetig sind, sind $(x, t) \mapsto x^2$ oder $(x, t) \mapsto \sqrt[3]{x}$, beide definiert auf $\mathbb{R} \times \mathbb{R}$. Wir wissen bereits, das für diese beiden rechten Seiten der Satz falsch wäre: im ersten Fall ist die Lösung nicht auf ganz \mathbb{R} definiert, und im zweiten Fall hat das AWP mit Anfangswert 0 mehr als eine Lösung.
2. Wenn wir T in überlappende Teilintervalle zerlegen können, sodass F in jedem Teilintervall Lipschitz-stetig ist, dann gilt die Schlussfolgerung des Satzes auch. Existenz und Eindeutigkeit auf T ist nämlich equivalent zur Existenz und Eindeutigkeit auf jedem Teilintervall. Man kann die eindeutige Lösung auf T aus den Lösungen auf Teilintervallen zusammenstückeln, indem man die Anfangsstellen in den Schnittpunkten der Intervalle wählt.
3. Im linearen Fall – sowohl homogen als auch nichthomogen – ist Lipschitz-Stetigkeit equivalent zur Stetigkeit, wenn wir T als kompakt annehmen. Diese Annahme der Kompaktheit ist ja kein Problem wegen Punkt (2) oben.
4. Wenn F stetig ist und nach x stetig differenzierbar ist (d.h. wenn die matrixwertige Funktion $(t, x) \mapsto \frac{\partial F(t, x)}{\partial x}$ existiert und stetig ist), dann ist die Lipschitzbedingung ebenfalls auf jedem kompaktem Intervall T erfüllt - hier nimmt man als L eine Schranke für die Operatornorm der Ableitung.
5. Das AWP für Differentialgleichungen höherer Ordnung kann auf das AWP erster Ordnung zurückgeführt werden (hier sind die Anfangswerte von allen Ableitungen bis Ordnung minus eins vorzugeben). Bei dieser Rückführung bleibt die Lipschitz-Bedingung erhalten, also gibt es auch eine Version von Satz 10.1 für Differentialgleichungen höherer Ordnung.

6. Es fällt auf, dass der Satz für den nicht-autonomen Fall formuliert ist, obwohl der ja auf den autonomen Fall zurückgeführt werden. Der Grund ist, dass diese Rückführung leider die Lipschitz-Bedingung zunichte macht, und ohne eine nicht-autonome Version des Satzes 10.1 würden wir für den wichtigen Fall der linearen nicht-autonomen Gleichungen keine Aussage bekommen. – Der lineare nicht-autonome Fall ist, unter uns gesagt, tatsächlich der Hauptgrund für das Beibehalten der Variablen t als Argument von F . In allen anderen Anwendungen werden wir voraussetzen, dass F nicht von t abhängt.
7. Im autonomen Fall ist F unabhängig von der Zeitvariable und kann also Funktion $\mathbb{R}^k \rightarrow \mathbb{R}^k$ aufgefasst werden. In vielen Fällen ist F nur auf einer offenen $U \subset \mathbb{R}^k$ definiert. Es wäre dann unvernünftig, die Existenz einer Lösung beweisen zu wollen; es kann ja jederzeit passieren, dass die Lösung den Bereich U verlässt. Was wir statt dessen machen können, ist, lokale Eindeutigkeit zu zeigen und die eindeutige Lösung bis an den Rand von U verfolgen.

Dem Punkt 7 können wir mit folgender Folgerung aus Satz 10.1 “Ausdruck verleihen”.

Korollar 10.2. *Es sei $U \subset \mathbb{R}^k$ offen, und $F : U \rightarrow \mathbb{R}^k$ eine Lipschitz-stetige Funktion. Dann existiert eine eindeutige Schranke $t_1 \in ((0, \infty) \cup \{+\infty\})$, sodass das AWP für $f : \mathbb{R} \rightarrow U$*

$$\forall t : f'(t) = F(f(t)), \quad f(0) = x_0$$

auf $(0, t_1)$ eine eindeutige Lösung f hat, und, falls $t_1 \neq +\infty$, sich f stetig nach t_1 fortsetzen lässt mit einem Wert auf dem Rand von U .

Beweis. Wir definieren eine Fortsetzung $H : \mathbb{R}^k \rightarrow \mathbb{R}^k$ von F , die wieder Lipschitz-stetig ist. Es ist nicht klar, dass das immer geht, aber nehmen wir mal an es geht. Das AWP für $h : \mathbb{R} \rightarrow \mathbb{R}^k$

$$\forall t : h'(t) = H(h(t)), \quad h(0) = x_0$$

hat nach Satz 10.1 eine eindeutige Lösung. Nun brauchen wir nur noch

$$t_1 := \min\{t \mid h(t) \notin U\}$$

zu definieren, wobei das Minimum der leeren Menge als $+\infty$ definiert ist.

Die Behauptung, dass man Lipschitz-stetige Funktionen Lipschitz-stetig fortsetzen kann, wird gleich noch bewiesen. \square

Lemma 10.3. *Es seien $m, n \in \mathbb{N}$. Es sei $U \subset \mathbb{R}^m$. Dann jede Lipschitz-stetige Funktion $f : U \rightarrow \mathbb{R}^n$ eine Lipschitz-stetige Fortsetzung $h : \mathbb{R}^m \rightarrow \mathbb{R}^n$.*

Beweis. Eine Funktion $x \mapsto (f_1(x), \dots, f_n(x))$ ist genau dann Lipschitz-stetig, wenn alle Komponenten $f_1, \dots, f_n : U \rightarrow \mathbb{R}$ Lipschitz-stetig sind. Deshalb reicht es, die Behauptung für $n = 1$ zu zeigen.

Im Fall $n = 1$ definieren wir

$$h(x) = \sup\{f(y) - L\|x - y\| \mid y \in U\},$$

wobei L eine Lipschitz-Konstante für f ist. \square

10.1 Vollständigkeit des Funktionenraums

Wir werden einen konstruktiven Beweis des Satzes 10.1 zeigen, in dem Sinn, dass eine Folge konstruiert wird, die gegen die Lösung konvergiert. Damit dieser Beweis aber greift, ist die Lösung in einem geeigneten metrischen Raum einzubetten. Wir wünschen uns dabei einen *vollständigen metrischen* Raum in folgendem Sinn.

Definition 10.1. Es sei (X, d) ein metrischer Raum: eine Menge X zusammen mit einer Funktion $d : X^2 \rightarrow \mathbb{R}$, die folgende Eigenschaften erfüllt für alle $x, y, z \in X$:

1. $d(x, y) \geq 0$, und $d(x, y) = 0$ genau dann wenn $x = y$;
2. $d(x, y) = d(y, x)$;
3. $d(x, y) + d(y, z) \geq d(x, z)$.

Eine Folge $(x_i)_{i \in \mathbb{N}}$ heißt *konvergent* mit Grenzwert y , wenn für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ existiert, sodass für alle $i \geq n$ die Ungleichung $d(x_i, y) \leq \epsilon$ erfüllt ist.

Eine Folge $(x_i)_{i \in \mathbb{N}}$ heißt *Cauchy-Folge*, wenn für jedes $\epsilon > 0$ ein $n \in \mathbb{N}$ existiert, sodass für alle $i, j \geq n$ die Ungleichung $d(x_i, x_j) \leq \epsilon$ erfüllt ist.

Der Raum (X, d) heisst *vollständig*, wenn jede Cauchy-Folge konvergent ist.

Der Raum \mathbb{Q} mit der Standard-Metrik $d(x, y) := |x - y|$ ist nicht vollständig. Der Raum \mathbb{R} mit der Standard-Metrik ist vollständig. Eine Teilmenge von \mathbb{R} ist genau dann vollständig, wenn die abgeschlossen ist.

Es sei $T \subset \mathbb{R}$ ein kompaktes Intervall, $n \in \mathbb{N}$. Auf der Menge $X := C^0(T, \mathbb{R}^n)$ aller stetigen Funktionen von T nach \mathbb{R}^n ist

$$(f, g) \mapsto \|f - g\| : \infty := \sup_{t \in T} \|f(t) - g(t)\|$$

überall definiert, weil stetige Funktionen auf kompakten Intervallen beschränkt sind; damit haben wir auf diesem Raum eine Metrik. Um zu zeigen, dass dieser Raum vollständig ist, ist eine weitere Eigenschaft der Funktionen in X nützlich.

Definition 10.2. Es seien (A, d_A) und (B, d_B) metrische Räume. Eine Funktion $f : A \rightarrow B$ heißt *gleichmäßig stetig*, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, sodass gilt:

$$\forall x, y \in A : d(x, y) < \delta \implies d(f(x), f(y)) < \epsilon.$$

Gleichmäßige Stetigkeit impliziert Stetigkeit, ist aber nicht equivalent. In Analysis 1 sind die Begriffe Stetigkeit, gleichmässige Stetigkeit, und Lipschitz-Stetigkeit für $A = B = \mathbb{R}$ verglichen worden (gilt aber für beliebige metrische Räume): Lipschitz-Stetigkeit ist stärker als gleichmässige Stetigkeit, und gleichmässige Stetigkeit ist stärker als Stetigkeit. Ausserdem wird dort der folgende Satz bewiesen (für $A = B = \mathbb{R}$, der gleiche Beweis funktioniert aber für beliebige metrische Räume).

Satz 10.4. *Es sei A kompakt. Dann ist jede stetige Funktion $f : A \rightarrow B$ auch gleichmäßig stetig.*

Nun sind wir soweit, Vollständigkeit von $C^0(T, \mathbb{R}^n)$ beweisen zu können.

Satz 10.5. *Der metrische Raum $C^0(T, \mathbb{R}^n)$ ist bezüglich der oben definierten Supremumsnorm vollständig.*

Proof. Es sei $(f_i)_{i \in \mathbb{N}}$ eine Cauchy-Folge von stetigen Funktionen. Für jedes $\epsilon > 0$ haben wir dann ein n_ϵ , sodass für alle $i, j \geq n$ die Ungleichung $\|f_i - f_j\|_\infty \leq \epsilon$ gilt. Für jedes $t \in T$ ist die Folge $(f_i(t))_i$ eine Cauchy-Folge in \mathbb{R}^n , weil $\|f_i(t) - f_j(t)\| \leq \epsilon$ ist für $i, j \geq n$. Der Raum \mathbb{R}^n ist vollständig, daher ist die Folge $(f_i(t))_i$ konvergent. Wir definieren die Grenzfunktion

$$g : T \rightarrow \mathbb{R}^n, t \mapsto \lim_{i \rightarrow \infty} f_i(t).$$

Für alle $i, j \geq n_\epsilon$ und für alle t ist $\|f_i(t) - f_j(t)\| \leq \epsilon$; wenn wir hier j gegen unendlich gehen lassen, bekommen wir die Ungleichung $\|f_i(t) - g(t)\| \leq \epsilon$. Nachdem das für alle t gilt, folgt

$$\|f_i - g\|_\infty \leq \epsilon$$

für alle $i \geq n_\epsilon$. Die Funktion g wäre also der gesuchte Grenzwert, aber wir müssen noch zeigen, dass g stetig ist.

Es sei $\epsilon > 0$. Die Funktion $f_{n_{\epsilon/3}}$ ist stetig. Der Definitionsbereich ist kompakt, also ist $f_{n_{\epsilon/3}}$ gleichmäßig stetig und es existiert ein δ , sodass

$$\forall s, t \in T : \|s - t\| < \delta \implies \|f_{n_{\epsilon/3}}(s) - f_{n_{\epsilon/3}}(t)\| < \epsilon/3$$

gilt. Für beliebige $s, t \in T$ mit Abstand kleiner als δ folgt dann

$$\begin{aligned} \|g(s) - g(t)\| &\leq \|g(s) - f_{n_{\epsilon/3}}(s)\| + \|f_{n_{\epsilon/3}}(s) - f_{n_{\epsilon/3}}(t)\| + \|f_{n_{\epsilon/3}}(t) - g(t)\| < \\ &< 2\|f_{n_{\epsilon/3}} - g\| + \epsilon/3 \leq 2\epsilon/3 + \epsilon = \epsilon. \end{aligned}$$

□

Bemerkung 10.3. Das erste Lehrbuch der Analysis, wie wir sie heute kennen, ist Cauchy's "Cours d'Analyse" (1821). Es enthält nicht nur die ϵ - δ -Definition der Stetigkeit, sondern auch eine Variante des obigen Satzes: wenn für eine Folge $f_i : T \rightarrow \mathbb{R}^n$ von Funktionen für alle $t \in T$ die Folge der Funktionswerte konvergent ist, dann ist die Grenzfunktion stetig. Der Beweis ist ähnlich dem obigen, er enthält nur einen kleinen Fehler. Und tatsächlich stimmt diese Variante im allgemeinen nicht.

Dieser Widerspruch in der damals neuen Disziplin Analysis führte zu wichtigen Auseinandersetzungen und Entdeckungen; eine ausführliche Diskussion steht im Appendix von J. Worrall und E. Zahar zu dem Buch: Imre Lakatos, Proofs and Refutations, CUP 1976.

10.2 Beweis des Satzes von Picard/Lindelöf

Der folgende Satz liefert ein sehr mächtiges Hilfsmittel, um in einem vollständigen metrischen Raum allerlei Arten von Gleichungen zu lösen.

Satz 10.6 (Banachscher Fixpunktsatz). *Es sei X ein vollständiger metrischer Raum. Es sei $f : X \rightarrow X$ eine Lipschitz-stetige Funktion mit Lipschitz-Konstante $C < 1$. Dann besitzt f einen eindeutigen Fixpunkt. Dieser Fixpunkt ist asymptotisch stabil bezüglich der Rekursion für $x : \mathbb{N} \rightarrow X$ gegeben durch $\forall i : x_{i+1} = f(x_i)$. Jede Folge, die diese Rekursion erfüllt, konvergiert gegen den Fixpunkt.*

Die Konvergenzgeschwindigkeit ist wie folgt beschreibbar: wenn $u \in X$ der Fixpunkt ist, dann gilt die Abschätzung

$$d(x_i, u) \leq \frac{d(x_0, x_1)}{1 - C} C^i.$$

(Man spricht von einer linearen Konvergenz.)

Beweis. Es sei $(x_i)_i$ eine Folge, die die Rekursion erfüllt. Durch Induktion nach i unter Verwendung der Lipschitz-Stetigkeit sieht man, dass die Ungleichung

$$d(x_i, x_{i+1}) \leq C^i d(x_0, x_1)$$

für alle $i \in \mathbb{N}$ erfüllt ist. Um zu zeigen, dass $(x_i)_i$ eine Cauchy-Folge ist, nehmen wir an, dass $N \geq i \geq j$ gilt, und beschränken den Abstand von x_i und x_j :

$$\begin{aligned} d(x_i, x_j) &\leq \sum_{r=i}^{j-1} d(x_r, x_{r+1}) \leq d(x_0, x_1) \sum_{r=i}^{j-1} C^r = \\ &= d(x_0, x_1) C^i \frac{1 - C^{j-i}}{1 - C} < \frac{d(x_0, x_1)}{1 - C} C^i \leq \frac{d(x_0, x_1)}{1 - C} C^N. \end{aligned}$$

Für jedes $\epsilon > 0$ läßt sich ein N finden, sodass die ganz rechte Seite kleiner als ϵ finden; und damit ist gezeigt, dass $(x_i)_i$ eine Cauchy-Folge ist.

Wegen der Vollständigkeit existiert ein Grenzwert; nennen wir ihn u . Da f Lipschitz-stetig, also stetig ist, ist

$$f(u) = f(\lim_{i \rightarrow \infty} x_i) = \lim_{i \rightarrow \infty} f(x_i) = \lim_{i \rightarrow \infty} x_{i+1} = u,$$

also ist u ein Fixpunkt. Wenn umgekehrt v ein Fixpunkt ist, dann ist

$$d(u, v) = d(f(u), f(v)) \leq cd(u, v),$$

und das geht nur wenn $d(u, v) = 0$, also $u = v$ gilt.

Zur Abschätzung der Konvergenzrate verwenden wir noch einmal die Ungleichung $d(x_i, x_j) \leq \frac{d(x_0, x_1)}{1 - C} C^i$, die wir ja schon gezeigt haben, und lassen j gegen unendlich gehen. \square

Nun geht es also darum, das Anfangswertproblem für $f : T \rightarrow \mathbb{R}^n$

$$f'(t) = F(t, f(t)), \quad f(t_0) = x_0$$

equivalent zu einer Fixpunktgleichung umzuformulieren. Wir definieren dazu die Funktion $P : C^0(T, \mathbb{R}^n) \rightarrow C^0(T, \mathbb{R}^n)$ durch

$$P(h)(t) := x_0 + \int_{t_0}^t F(s, h(s)) ds.$$

Eine Funktion von einem Funktionenraum in einen Funktionenraum nennen wir auch Operator.

Lemma 10.7. *Eine Funktion $f \in C^0(T, \mathbb{R}^n)$ ist genau dann ein Fixpunkt von P , wenn sie differenzierbar ist und das AWP löst.*

Proof. Man verwende den Hauptsatz der Differential- und Integralrechnung. □

Lemma 10.8. *Der Operator P ist Lipschitz-stetig mit Lipschitz-Konstante $K := l(T)L$, wobei L die Lipschitz-Konstante von F und $l(T)$ die Länge des Intervalls T ist.*

Proof. Man verwende die Abschätzung des Integrals durch das Produkt einer Schranke für den Integranden mit der Länge des Intervalls. □

Damit ist Satz 10.1 beweisen für den Fall, dass $l(T)L < 1$ ist. Wenn $l(T)L \geq 1$ ist, kann man aber T in Teilintervalle von Länge kleiner als $1/L$ zerlegen. Weil der Satz 10.1 in jedem Teilintervall gilt, gilt er auch auf ganz T (siehe Punkt 2 in Kapitel 10).

10.3 Differenzierbare Vektorfelder

Nachdem nun die Existenz und Eindeutigkeit im wichtigen linearen nichtautonomen Fall geklärt ist, werden wir von nun an immer den autonomen Fall annehmen. Wir betrachten also offene Mengen $X \subset \mathbb{R}^n$ und stetige Funktionen $F : X \rightarrow \mathbb{R}^n$. Eine weitere Annahme, die uns von nun an bis zum Schluss begleiten wird, ist, dass die Funktion F differenzierbar ist. Wir nennen F auch ein *differenzierbares Vektorfeld*. Zu jedem Vektorfeld gehört eine Differentialgleichung für $f : \mathbb{R} \rightarrow U$

$$\forall t : f'(t) = F(f(t)).$$

Die Lösungen nennt man auch *Integralkurven* von F .

Satz 10.9. *Durch jeden Punkt x_0 gibt es eine eindeutige Integralkurve $f : T \subset \mathbb{R}, T \subset \mathbb{R}$ offen, sodass*

- $0 \in T$ und $f(0) = x_0$;
- wenn $\sup(T)$ existiert, dann existiert $\lim_{t \rightarrow \sup(T)} f(t)$ und liegt auf dem Rand von X .

- wenn $\inf(T)$ existiert, dann existiert $\lim_{t \rightarrow \inf(T)} f(t)$ und liegt auf dem Rand von X .

Beweis. Man überdecke X durch offene Mengen mit kompaktem Abschluss in X und wende Korollar 10.2 an. \square

Der Fluss eines Vektorfelds $F : X \rightarrow \mathbb{R}^n$ ist genau dann für ein Paar $(t, x) \in \mathbb{R} \times X$ definiert, wenn $t \in T$ ist, wobei T wie oben der Definitionsbereich der eindeutigen Interalkurve $f : T \rightarrow X$ durch x ist. Der Wert des Flusses ist definiert als $\phi(x, t) := f(t)$. Der Fluss ist auf jeden Fall definiert für $t = 0$ und erfüllt dort die Gleichung $\phi(0, x) = x$. Ausserdem gilt die Funktionalgleichung

$$\varphi(s, \phi(t, x)) = \varphi(s + t, x),$$

wann immer die linke Seite definiert ist. Sollte der Fluss überall definiert sein, dann ist für jedes fixe $t \in \mathbb{R}$ die Abbildung $\varphi_t : X \rightarrow X$, $x \mapsto \varphi(t, x)$ eine bijektive Abbildung. Die Umkehrfunktion ist φ_{-t} .

10.4 Weitere Folgerungen aus Picard/Lindelöf

Satz 10.1 beinhaltet als Nebenprodukt eine Schranke für den Abstand der Lösungsfunktion f von einer beliebigen vorgegebenen Funktion $g : T \rightarrow \mathbb{R}$. Wegen Satz 10.1 gilt nun

$$\|f - g\|_\infty \leq \frac{\|g - P(g)\|}{1 - Ll(T)}$$

unter der Voraussetzung, dass der Nenner auf der rechten Seite positiv ist. Die rechte Seite können wir abschätzen, ohne irgendwas über f zu wissen.

Es sei zum Beispiel g die stückweise lineare Funktion, die beim Eulerschen Polygonzugverfahren (siehe Kapitel 5) als Näherung für f berechnet wird. Wir nehmen der Einfachheit halber den skalaran Fall $n = 1$ und $T = [0, 1]$ an. Das Intervall T wird in N gleich große Intervalle der Länge $h = 1/N$ unterteilt, und auf jedem Teilintervall $[i/N, (i+1)/N]$ ist $g(t) = y_i + F(y_i)(t - i/N)$, wobei (y_0, \dots, y_N) eine Folge ist, die durch die Rekursion

$$y_{i+1} = y_i + \frac{F(y_i)}{N}, \quad y_0 = x_0$$

berechnet wird. Der Abstand $\|g - P(g)\|$ kann beschränkt werden durch die Dreiecksungleichung

$$\begin{aligned} |P(g)(1) - g(1)| &\leq \sum_{i=0}^{n-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} |P(g)'(s) - g'(s)| = \sum_{i=0}^{n-1} \int_{\frac{i}{N}}^{\frac{i+1}{N}} |F(g(t)) - F(y_i)| dt \leq \\ &\leq \sum_{i=0}^{n-1} L \int_{\frac{i}{N}}^{\frac{i+1}{N}} \|y_i - g(t)\| = \sum_{i=0}^{n-1} L \int_0^{\frac{1}{N}} \|F(y_i)\| s ds \leq N \frac{LM}{2N^2} = \frac{LM}{2N}. \end{aligned}$$

Die Abschätzung von $\|g(t) - P(g)(t)\|$ für $t \in [0, 1]$ ist ebenfalls durch diese Summe beschränkt (nicht alle Summanden werden gebraucht). Wenn N gegen unendlich geht, dann geht also der Fehler gegen Null.

Eine bessere Konvergenz erhält man durch das Zweischrittverfahren

$$z_i = y_i + \frac{F(y_i)h}{2}, y_{i+1} = y_i + hF(z_i).$$

Es sei nun wieder n beliebig, und $F : X \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld mit Lipschitz-Konstante L . Es sei $T \subset \mathbb{R}$ ein kompaktes Intervall mit Länge $l(T) < \frac{1}{L}$, das 0 enthält. Wenn x_0, y_0 zwei mögliche Anfangswerte sind, können wir den Abstand der Integralkurven $f : \mathbb{R} \rightarrow X, t \mapsto (t, x_0)$ und $g : \mathbb{R} \rightarrow X, t \mapsto (t, y_0)$ abschätzen, indem wir den Picard/Lindelöf-Operator $P : C^0(T, \mathbb{R}^n) \rightarrow C^0(T, \mathbb{R}^n)$ auf g anwenden:

$$P(g)(t) = x_0 + \int_0^t F(g(s))ds = x_0 + \int_0^t g'(s)ds = x_0 + g(t) - y_0.$$

Es folgt

$$\|f - g\|_\infty \leq \frac{\|P(g) - g\|}{1 - l(T)L} = \frac{\|x_0 - y_0\|}{1 - l(T)L}.$$

Der Abstand zwischen den Werten zweier Integralkurven vergrößert sich daher nach der Zeit t höchstens um einen Faktor von $\frac{1}{1 - l(t)L}$. Wenn $l(T)L \geq 1$ gilt, dann hilft es auch hier, das Intervall zu unterteilen in N gleich grosse Teilintervalle. Der Abstand der Werte vergrößert sich dann höchstens um den Faktor $\left(\frac{1}{1 - \frac{l(T)L}{N}}\right)^N$. Wir lassen N gegen unendlich gehen:

$$\lim_{N \rightarrow \infty} \left(\frac{1}{1 - \frac{l(T)L}{N}}\right)^N = e^{l(T)L}.$$

Der Abstand vergrößert sich also höchstens um einen Faktor, der exponentiell mit der Zeit wächst.

Wenn man ein differenzierbares Vektorfeld auf einer offenen Teilmenge $X \subset \mathbb{R}^n$ hat, dann kann man X durch beschränkte offene Teilmengen überdecken, deren Abschluss in X liegt (kompakte Ausschöpfung). Auf jeder Teilmenge haben wir einerseits eine Lipschitz-Konstante und andererseits eine Schranke für die Länge der Vektoren, die dann auch eine Schranke der Ableitungen aller Integralkurven ist. Damit kann man folgenden Satz zeigen.

Satz 10.10. *Der Definitionsbereich des Flusses $\phi : D \subset (\mathbb{R} \times X) \rightarrow X$ ist offen, und ϕ ist stetig (sogar Lipschitz-stetig in jedem kompakten Teilintervall).*

Beweis. Es sei $x_0 \in X$. Es sei $V \subset X$ eine beschränkte offene Umgebung von x_0 mit Abschluss in X . Wir schränken das Vektorfeld F zuerst auf V ein. Dort ist die Ableitung von F beschränkt, also ist die Einschränkung Lipschitz. Es sei G eine beschränkte und Lipschitz-stetige Fortsetzung der Einschränkung auf ganz \mathbb{R}^n . Es sei L eine Lipschitz-Konstante von G und M eine Schranke der Länge der Vektoren im Bild von G . Wir

wählen eine offene Umgebung $W \subset V$ von x_0 , deren Abschluss in V enthalten ist. Dann hat jeder Punkt im Abschluss von W einen positiven Abstand zum Rand von V ; es sei R das Minimum aller dieser Abstände (das Minimum wird erreicht, weil der Abschluss von W kompakt ist). Es sei $T := \{t \mid M|t| < R\}$. Dann ist sichergestellt, dass die Intergralkurven durch Anfangswerte in W in der Zeit T die Menge V nicht verlassen.

Es sei $\tilde{\phi} : (\mathbb{R} \times \mathbb{R}^n) \rightarrow \mathbb{R}^n$ der Fluss von G – dieser ist überall definiert, weil G Lipschitzstetig ist. Dann gilt für alle $t_1, t_2 \in \mathbb{R}$ und $x_1, x_2 \in \mathbb{R}^n$ die Abschätzung

$$\|\tilde{\phi}(t_1, x_1) - \tilde{\phi}(t_2, x_2)\| \leq \|\tilde{\phi}(t_1, x_1) - \tilde{\phi}(t_1, x_2)\| + \|\tilde{\phi}(t_1, x_2) - \tilde{\phi}(t_2, x_2)\| \leq e^{L|t_1|} \|x_1 - x_2\| + M|t_1 - t_2|.$$

Wenn (t_2, x_2) gegen (t_1, x_1) geht, geht die rechte Seite gegen Null, also ist $\tilde{\phi}$ stetig. Im Bereich $T \times W$ stimmt der Fluss von F mit dem Fluss von G überein, daher ist auch ϕ stetig im Inneren von $T \times W$. Da x_0 am Anfang beliebig gewählt war, ist ϕ stetig in jedem Punkt $(x, 0)$, $x \in X$; und jeder solche Punkt besitzt eine offene Umgebung, in dem ϕ definiert ist.

Es sei nun (t_0, x_0) ein beliebiger Punkt im Definitionsbereich $D \subset (\mathbb{R} \times X)$. Wir setzen $x_1 := \phi(t_0, x_0)$. Dann ist der Fluss in einer offenen Umgebung von $(0, x_1)$ definiert und stetig. Wegen der Gleichung

$$\phi(t + t_0, x_0) = \phi(t, x_1)$$

ist daher ϕ in einer offenen Umgebung von (t_0, x_0) offen und definiert. \square

Der nächste Satz besagt, dass die Lösung einer Differentialgleichung nicht nur stetig vom Anfangswert abhängt, sondern auch stetig in der Differentialgleichung ist; das soll heißen, stetig in Parametern, die eventuell in der Differentialgleichung vorkommen.

Satz 10.11. *Es sei $n \in \mathbb{N}$. Es sei $X \subset \mathbb{R}^n$. Es sei $P \subset \mathbb{R}$ ein offenes Intervall, Es sei $F : (P \times X) \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Funktion (die wir als parameterabhängiges Vektorfeld verstehen). Für jedes $\lambda \in P$ sei $\varphi_\lambda : D_\lambda \rightarrow X$ mit $D_\lambda \subset (\mathbb{R} \times X)$ der Fluss von $F_\lambda : x \rightarrow F(\lambda, x)$. Es sei $\bar{\varphi} : (\lambda, t, x) \mapsto \varphi_\lambda(t, x)$ die parameterabhängige Flussfunktion, die genau dann definiert ist, wenn $(t, x) \in D_\lambda$ ist. Dann ist der Definitionsbereich von $\bar{\varphi}$ offen und $\bar{\varphi}$ ist stetig.*

Beweis. Wir definieren einfach ein Vektorfeld G auf $P \times X$ durch $(\lambda, X) \mapsto (0, F_\lambda(x))$. Der Fluss von G ist die Funktion $(t, \lambda, x) \mapsto (\lambda, \bar{\varphi}(\lambda, t, x))$. Das Richtungsfeld G ist differenzierbar, daher ist sein Fluss stetig und der Definitionsbereich offen. \square

Die letzte Anwendung ist die Differenzierbarkeit des Flusses. Es ist klar, dass die partielle Ableitung nach der Zeit existiert und stetig ist – es ist ja gerade die rechte Seite der Differentialgleichung – aber die partielle Ableitung nach dem Anfangswert ist noch nicht klar.

Wir beweisen den Satz für $n = 1$. Der Beweis lässt sich auf den vektorwertigen Fall verallgemeinern, aber diese Verallgemeinerungen ist technisch kompliziert; es sei auf <https://www.math.uni-hamburg.de/home/lauterbach/scripts/dsode12/ch3.pdf>, Satz 3.2.14, Seite 89ff, verwiesen.

Satz 10.12. *Es sei $X \subset \mathbb{R}$, $F : X \rightarrow \mathbb{R}$ differenzierbar, $\phi : D \subset (\mathbb{R} \times X) \rightarrow X$ der Fluss von F . Dann ist ϕ stetig differenzierbar.*

Proof. Es sei $(x, t) \in D$. Für kleine $\lambda \neq 0$ ist $(x + \lambda, t) \in D$. Wir setzen

$$h_\lambda(t) := \frac{\varphi(t, x + \lambda) - \varphi(t, x)}{\lambda}$$

und zeigem, dass der Grenzwert für λ gegen Null existiert. Wir berechnen die Ableitung:

$$\begin{aligned} h'_\lambda(t) &= \frac{\frac{\partial \varphi}{\partial t}(t, x + \lambda) - \frac{\partial \varphi}{\partial t}(t, x)}{\lambda} = \frac{F(\varphi(t, x + \lambda)) - F(\varphi(t, x))}{\lambda} = \\ &= \frac{F(\varphi(t, x + \lambda)) - F(\varphi(t, x))}{\varphi(t, x + \lambda) - \varphi(t, x)} \frac{\varphi(t, x + \lambda) - \varphi(t, x)}{\lambda} = \frac{F(\varphi(t, x + \lambda)) - F(\varphi(t, x))}{\varphi(t, x + \lambda) - \varphi(t, x)} h_\lambda(t) \end{aligned}$$

(man beachte, dass der Nenner des linken Faktors ungleich 0 ist, weil ϕ_t injective ist). Dies ist eine parameter-abhängige Differentialgleichung für h_λ . Der Anfangswert ist

$$h_\lambda(0) = \frac{x + \lambda - x}{\lambda} = 1.$$

Wegen Satz 10.11 ist die Lösung stetig in λ . Der Grenzwert $h_0 = \lim_{t \rightarrow 0} h_\lambda$ existiert also, weil er die Lösung des Anfangswertsproblems für $\lambda = 0$ ist:

$$h'_0(t) = F'(\phi(t, x))h_0(t), \quad h_0(0) = 1.$$

In dieser Differentialgleichung kann man x leicht als Parameter hineinbringen, daher ist die Ableitung des Flusses stetig in x . Stetig in t ist sie sowieso, sogar differenzierbar nach t , wie aus der Differentialgleichung hervorgeht. \square

11 Äquivalenz und lokale Äquivalenz

Die Differentialgleichung für $f : \mathbb{R} \rightarrow \mathbb{R}^2$ gegeben durch das Vektorfeld

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (x_1, x_2) \mapsto (x_2 + 2x_1x_2 + 2x_2^3, -x_1 - x_2^2)$$

, also

$$(f'_1, f'_2) = (f_2 + 2f_1f_2 + 2f_2^3, -f_1 - f_2^2),$$

sieht zwar kompliziert aus, ist aber eine verkappte lineare Differentialgleichung mit konstanten Koeffizienten. Mit

$$(x_1, x_2) = (y_1 - y_2^2, y_2), \quad (y_1, y_2) = (x_1 + x_2^2, x_2)$$

haben wir eine nichtlineare umkehrbare Transformation der Ebene auf sich selbst. Wenn $g : \mathbb{R} \rightarrow \mathbb{R}^2$ eine Lösung in den transformierten Koordinaten ist, dann ist

$$(g'_1, g'_2) = ((f_1 + f_2^2)', f'_2) = (f'_1 + 2f_2f'_2, f'_2) = (f_2 + 2f_1f_2 + 2f_2^3 + 2f_2(-f_1 - f_2^2), -f_1 - f_2^2) =$$

$$= (f_2, -f_1 - f_2^2) = (g_2, -g_1).$$

Die Lösung ist

$$g(t) = e \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix} g(0) = (\cos(t)g_1(0) + \sin(t)g_2(0), -\sin(t)g_1(0) + \cos(t)g_2(0)).$$

Die Lösung f für die erste Differentialgleichung bekommen man durch Einsetzen:

$$(f_1(t), f_2(t)) = (g_1(t) - g_2(t)^2, g_2(t)).$$

Allgemein kann man ein differenzierbares Vektorfeld $F : X \rightarrow \mathbb{R}^n$ (mit $X \subset \mathbb{R}^n$ offen) transformieren, wenn man eine bijektive differenzierbare Abbildung $U : X \rightarrow Y$ mit differenzierbarer Umkehrung $U^{-1} : Y \rightarrow X$ hat. Die Formel für das transformierte Vektorfeld ist

$$G : Y \rightarrow \mathbb{R}^n, y \mapsto U'(U^{-1}(y)) \cdot F(U^{-1}(y)).$$

Für die rechnerische Transformation ist es wohl besser, die Formel zu vergessen und wie im Beispiel vorhin die transformierte Gleichung mit der Kettenregel auszurechnen. Die Transformation U bildet Integralkurven nach Integralkurven ab. Vektorfelder, die durch eine solche Transformation ineinander übergeführt werden, heißen *transformations-equivalent*. Ein fundamentales Beispiel einer Transformation ist die Transformation durch den Fluss. Dazu brauchen wir ein Richtungsfeld $F : X \rightarrow \mathbb{R}^n$ und einen Punkt $x_0 \in X$, der kein Equilibrium ist. Wir wählen eine glatte Hyperfläche H , das kann etwa eine Hyperebene sein, die x_0 enthält und deren Tangentialebene bei x_0 *nicht* den Vektor $F(x_0)$ enthält. Die Abbildung

$$K : D \subset (\mathbb{R} \times H) \rightarrow X, (t, p) \mapsto \phi(t, p)$$

ist in einer offenen Umgebung von $(0, x)$ definiert und die Ableitung an diesem Punkt ist eine invertierbare lineare Abbildung: Das Bild der ersten Spalte ist der ein-dimensionale Vektorraum, der von $F(x_0)$ erzeugt ist, und das Bild der restlichen Spalten ist die Tangentialebene von H . Nach dem Satz über inverse Funktionen existiert eine Umgebung von $(0, x)$, sodass K invertierbar ist. Die inverse Abbildung transformiert die Integralkurven auf parallele Geraden, und das transformierte Richtungsfeld ist konstant: $x \mapsto (1, 0, \dots, 0)$. Daraus folgt, dass alle Vektorfelder in der gleichen Dimension bei Nicht-Equilibrien lokal zueinander transformations-equivalent sind.

Wir stellen uns die Frage: welche Dinge bleiben durch die Transformationen von Vektorfeldern erhalten bzw. werden wieder in Dinge mit den gleichen Eigenschaften transformiert? Wir kennen schon ein Beispiel: Integralkurven. Nachdem die Lösung einer Differentialgleichung eine parametrisierte Integralkurve ist, können wir auch von diesen parametrisierten Integralkurven ebenfalls sagen, dass diese Eigenschaft von parametrisierten Kurven (nämlich eine richtig parametrisierte Integralkurve zu sein) erhalten bleibt. Mit der Verwendung der Flüsse φ_F und φ_G lässt sich dieser Sachverhalt durch die Formel

$$U(\varphi_F(t, x)) = \varphi_G(t, U(x))$$

ausdrücken.

Equilibrien werden in Equilibrien transformiert. Die Jacobi-Matrix des transformierten Vektorfelds beim Bild-Equilibrium ist zur Jacobi-Matrix des Urbild-Vektorfelds beim Urbild-Equilibrium ähnlich.

Hier ist eine weitere Operation auf Vektorfeldern, die richtig transformiert wird:

Definition 11.1. Es sei $X \subset \mathbb{R}^n$ offen, $F : U \rightarrow \mathbb{R}^n$ ein Vektorfeld, $f : X \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Dann definieren wir die Funktion $\partial_F(f) : X \rightarrow \mathbb{R}$, genannt die vektorielle Ableitung, als

$$\partial_F(f)(x) := \langle \text{grad}(f)(x) \mid F(x) \rangle;$$

dabei ist $\text{grad}(f)$ der Gradient der Funktion f .

Transformationen von Vektorfeldern transformieren die vektorielle Ableitung richtig, nämlich

$$\partial_F(f)(x) = \partial_G(f \circ U^{-1})(U(x))$$

für alle $x \in X$.

Die vektorielle Ableitung beschreibt die Änderungsrate der Funktion entlang einer Integralkurve: für alle (x, t) im Definitionsbereich des Flusses gilt

$$\frac{df(\varphi(t, x))}{dt} = \partial_F(f)(\varphi(x, t)), \quad (14)$$

wie man aus der Kettenregel herleiten kann. Insbesondere ist jede Funktion, deren vektorielle Ableitung gleich Null ist, auf jeder Integralkurve konstant.

Beispiel 11.1. Es sei $X = (0, \infty)^2 \subset \mathbb{R}^2$ und $F : X \rightarrow \mathbb{R}^2$ das Vektorfeld $(x, y) \mapsto (x(1 - y), y(x - 1))$. Das einzige Equilibrium ist $(x_0, y_0) = (1, 1)$. Die vektorielle Ableitung der Funktion

$$g(x, y) := x - \ln(x) + y - \ln(y)$$

ist gleich Null auf ganz U , daher sind die die Integralkurven Niveaulinien des Graphen von g . Diese Niveaulinien sind geschlossene Kurven; es folgt, dass es rund um das einzige Equilibrium periodische Integralkurven gibt.

Man kann auch zeigen, dass die Funktion ein absolutes Minimum hat und dass der Wert unbeschränkt wächst, wenn x oder y gegen Null oder gegen Unendlich gehen. Daraus folgt, dass jede Integralkurve in U geschlossen ist, ausser der konstanten Lösung im Equilibrium.

Die Transformationsequivalenz für manche Anwendungen zu stark; wir werden noch einen weiteren Äquivalenzbegriff einführen, der etwas weniger verlangt. Zwei stetig differenzierbare Richtungsfelder $F : U \rightarrow \mathbb{R}^n$ und $G : V \rightarrow \mathbb{R}^n$ (die Dimension der beiden Vektorräume muss gleich sein) heißen *topologisch äquivalent*, wenn es einen Homöomorphismus $\tau : U \rightarrow V$ gibt, der Integralkurven erhält und der für jede Integralkurve die Richtung beibehält. Transformationsequivalente Richtungsfelder sind topologisch äquivalent, aber die Umkehrung gilt im Allgemeinen nicht.

Beispiel 11.2. Es sei $\lambda_1, \lambda_2 > 0$ verschiedene reelle Zahlen. Die Integralkurven des Vektorfelds

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (\lambda_1 x_1, \lambda_2 x_2)$$

sind von der Form $(c_1 e^{\lambda_1 t}, c_2 e^{\lambda_2 t})$. Die Transformation

$$(x_1, x_2) \mapsto (y_1, y_2) = (\text{signum}(x_1)|x_1|^{1/\lambda_1}, \text{signum}(x_2)|x_2|^{1/\lambda_2})$$

führt die Lösungskurven in offene Halbgeraden durch den Nullpunkt über. Das Vektorfeld ist daher topologisch äquivalent zu

$$G : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (x_1, x_2).$$

Die Transformation ist ein Homöomorphismus, aber nicht differenzierbar im Nullpunkt. Es kann auch gar keine Transformation geben, die F in G transformiert, weil die Jacobi-Matrizen beim Nullpunkt nicht ähnlich sind.

Es sei $n \in \mathbb{N}$, $U \subset \mathbb{R}^n$ offen, $F : X \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Richtungsfeld, $x_0 \in X$ ein Equilibrium. Wir nenne das Equilibrium *hyperbolisch*, wenn alle Eigenwerte (reell oder komplex) der Jacobi-Matrix $F'(x_0)$ Realteil ungleich 0 haben.

Die *Linearisierung* des Vektorfeldes F bei x_0 ist definiert als das Vektorfeld

$$L : \mathbb{R}^n \rightarrow \mathbb{R}^n, x \mapsto F'(x_0)x.$$

Satz 11.3 (Hartmann/Grobmann). *Es sei x_0 ein hyperbolisches Equilibrium des Vektorfeldes $F : X \rightarrow \mathbb{R}^n$ und $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ seine Linearisierung bei x_0 . Dann existiert eine offene Umgebung U von x_0 und eine offene Umgebung V von $\mathbf{0} \in \mathbb{R}^n$, sodass $F|_U$ und $L|_V$ topologisch äquivalent sind.*

Wir beweisen den Satz von Hartmann/Grobmann nur für den (besonders wichtigen) Fall, dass alle Eigenwerte negativen Realteil haben. In diesem Fall ist der Nullpunkt für L ein asymptotisch stabiles Equilibrium, und aus dem Satz folgt, dass auch x_0 ein asymptotisch stabiles Equilibrium ist. Dazu definieren wir eine differenzierbare Funktion $g : X \rightarrow \mathbb{R}$ und eine offene Umgebung U von x_0 mit den folgenden Eigenschaften:

- $g(x_0) = 0$ und $g(x) > 0$ für $x \in U \setminus \{x_0\}$.
- $\partial_F(g)(x) < 0$ für $x \in U \setminus \{x_0\}$.

(Eine Funktion mit solchen Eigenschaften wird *Ljapunov-Funktion* genannt.) Wir wählen eine kompakte Umgebung K von x_0 , die in U enthalten ist. Der Rand ∂K ist ebenfalls kompakt. Wir setzen $c := \min g(x) \mid x \in \partial K$ und $W := \{x \mid g(x) < c\}$. Weil $c > 0$ sein muss, ist W eine offene Umgebung von x_0 , die in U enthalten ist. Es sei $x \in W$. Wir behaupten, dass der Fluss φ von F für (t, x) für alle $t > 0$ definiert ist, dass das Bild in W liegt, und dass $\lim_{t \rightarrow \infty} \varphi(t, x) = x_0$ gilt. Um das zu zeigen, nehmen wir indirekt an, dass ein $t_0 > 0$ existiert mit $\phi(t_0, x) \in \partial W$. Die Menge alle t_0 , für die das gilt, ist

abgeschlossen und hat daher ein kleinstes Element; wir nehmen daher an, dass t_0 schon das kleinste Element ist. Die Funktion $h : [0, t_0] \rightarrow \mathbb{R}$, $t \mapsto g(f(t))$ erfüllt dann $h(0) < c$, $h(t_0) = c$, und $h'(t) = \partial_F(g)(f(t)) < 0$, ein Widerspruch. Damit sind die ersten beiden Teile der Behauptung gezeigt. Die Folge $(f(i))_{i \in \mathbb{N}}$ hat Werte in der kompakten Menge K und hat daher eine konvergente Teilfolge, die durch eine streng monoton steigende Funktion $s : \mathbb{N} \rightarrow \mathbb{N}$ beschrieben werden kann: $s(i)$ ist der Index des i -ten Glieds der Teilfolge. Nun ist h für alle positiven Stellen definiert, streng monoton fallend, und beschränkt. Daher ist

$$0 = \lim_{i \rightarrow \infty} h'(i) = \lim_{i \rightarrow \infty} h'(s(i)) = \lim_{i \rightarrow \infty} \partial_F(g)(f(s(i))) = \partial_F(g)(\lim_{i \rightarrow \infty} f(s(i))),$$

daraus folgt $\lim_{i \rightarrow \infty} f(s(i)) = x_0$, daraus folgt $\lim_{i \rightarrow \infty} h(s(i)) = 0$, daraus folgt $\lim_{t \rightarrow \infty} h(t) = 0$, daraus folgt $\lim_{t \rightarrow \infty} f(t) = x_0$.

Bemerkung 11.2. Die Existenz einer Funktion g mit den genannten Eigenschaften impliziert asymptotische Stabilität (Übung). Sie erlaubt es darüber hinaus, eine konkrete Umgebung anzugeben (nämlich W), die der Fluss nicht mehr verlassen kann. Mehr noch, man kann sogar die Menge weiterverkleinern, $W' = \{x \mid g(x) < c'\}$ für $c' \in (0, c)$, und bekommt immer kleinere Umgebungen von x_0 , die der Fluss nicht mehr verlassen kann.

Wir sollten jetzt aber auch zeigen, dass eine solche Ljapunov-Funktion g und Umgebung U wie oben existiert. Wir nehmen an, dass $x_0 = 0$ ist. Zuerst definieren wir

$$g(x) := \int_0^\infty \|e^{tL}x\|^2 dt,$$

wobei wir hier die lineare Abbildung L als Matrix geschrieben haben – mit anderen Worten, $L = F'(0)$. Um zu zeigen, dass das uneigentliche Integral existiert, verwenden die Jordan-Normalform von L und die Annahme, dass die Eigenwerte negativen Realteil haben. Der Integrand ist nie negativ und nur für $x = 0$ gleich Null, also gilt das gleiche auch für g . Der Gradient von g ist

$$\text{grad}(g)(x) = \int_0^\infty \frac{\partial}{\partial x} \langle e^{tL}x, e^{tL}x \rangle dt = 2 \int_0^\infty (e^{tL})^\dagger e^{tL} x dt,$$

also eine lineare Funktion von x . Die vektorielle Ableitung nach L ist

$$2 \int_0^\infty \langle (e^{tL})^\dagger e^{tL} x, Lx \rangle dt = 2 \int_0^\infty \langle e^{tL}x, Le^{tL}x \rangle dt = \int_0^\infty \frac{\partial \|e^{tL}x\|^2}{\partial t} dt = -\|e^{0L}x\|^2 = -\|x\|^2.$$

Nachdem F bei 0 differenzierbar ist, ist $F(x) = Lx + R(x)$ mit $\lim_{x \rightarrow 0} \frac{\|R(x)\|}{\|x\|} = 0$. Es sei C eine Konstante, sodass $\|\text{grad}(g)(x)\| \leq C\|x\|$ für alle x ist. Wir wählen die Umgebung U so, dass $\|R(x)\| < \frac{1}{2C}\|x\|$ ist für alle $x \in U$. Der Gradient von g linear in x , also ist

$$\partial_F(g)(x) = \delta_L(g)(x) + \langle \text{grad}(g)(x), R(x) \rangle \leq -\|x\|^2 + C\|x\| \cdot \|R(x)\| \leq -\frac{1}{2}\|x\|^2.$$

Damit ist gezeigt, dass g die geforderderten Eigenschaften hat.

Tip: Die Konstruktion für g funktioniert zwar immer, ist aber aufwändig. Wenn man ein konkretes Beispiel hat und eine Ljapunov-Funktion angeben will, ist es eine gute Strategie, ob nicht vielleicht die Funktion $g : x \rightarrow \|x - x_0\|^2$ schon alle Bedingungen erfüllt. Es reicht, zu zeigen, dass die vektorielle Ableitung in der Nähe von x_0 negativ ist. Mit etwas Glück passt dann schon alles. Wenn nicht, muss man eben doch das uneigentliche Integral berechnen.

Satz 11.4 (Hartmann/Grobmann, Spezialfall). *Es sei $X \subset \mathbb{R}^n$, $F : X \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld, $x_0 \in X$ ein hyperbolisches Equilibrium, sodass die Jacobi-Matrix $F'(x_0)$ nur Eigenwerte mit negativem Realteil haben. Dann existiert eine Umgebung W von x_0 und eine Umgebung V von 0 in \mathbb{R}^n , sodass $F|_W$ und das Vektorfeld $x \mapsto -x$ eingeschränkt auf V topologisch äquivalent sind.*

Beweis. Ohne Beschränkung der Allgemeinheit können wir $x_0 = 0$ annehmen. Es seien g und W wie oben. Insbesondere ist g eine positiv definite quadratische Form und $W = \{x \mapsto g(x) < c\}$ für eine positive Zahl c . Wir definieren die Abbildung

$$\tau : \overline{W} \rightarrow \overline{W}, \quad x \mapsto \phi_F \left(\frac{1}{2} \ln \left(\frac{c}{g(x)} \right), \frac{x}{\sqrt{g(x)}} \right);$$

für $x = 0$ ist die rechte Seite nicht definiert, hier setzen wir $\tau(0) = 0$. τ ist stetig, injektiv, und gleich der Identität am Rand. Jede stetige und injektive Abbildung des Balls auf sich selbst, die am Rand die Identität ist, ist ein Homöomorphismus. Sie bildet die Halbgeraden durch 0 auf Integralkurven von F ab, also sind F und $x \mapsto -x$ topologisch äquivalent. \square

Wenn x_0 ein Equilibrium von $F : X \rightarrow \mathbb{R}^n$ und y_0 ein Equilibrium von $G : Y \rightarrow \mathbb{R}^n$ ist, und es existieren Umgebungen U von x_0 und V von y_0 , dann sagen wir: F ist bei x_0 zu G bei y_0 lokal topologisch äquivalent. In dieser Sprechweise gilt also: jedes Vektorfeld ist bei jedem hyperbolischen Equilibrium lokal topologisch äquivalent zu seiner Linearisierung.

Bemerkung 11.3. Man kann sogar zeigen, dass die Einschränkung von τ auf $W \setminus \{0\}$ ein Diffeomorphismus ist, der das Vektorfeld $x \mapsto -x$ in das Vektorfeld F transformiert. Mit anderen Worten, ohne den Punkt hätten wir sogar Transformations-Äquivalenz. Im Equilibrium kann es keine Transformations-Äquivalenz geben, ausser wenn die Jacobi-Matrix $F'(x_0)$ gleich $-I_n$ ist. (Das ist die einzige Matrix, die zu $-I_n$ ähnlich ist.)

Hyperbolische Equilibrien, sodass die Eigenwerte der Jacobi-Matrix negativen Realteil haben, bezeichnen wir als *Senken*. Hyperbolische Equilibrien, sodass die Eigenwerte der Jacobi-Matrix positiven Realteil haben, bezeichnen wir als *Quellen*. Alle anderen hyperbolischen Equilibrien bezeichnen wir als *Sattelpunkte*.

Lemma 11.5. *Es seien $A, B \in \mathcal{H}_{a,b} \subset \mathbb{R}^{n \times n}$ (siehe Kapitel 9). Dann sind die Richtungsfelder $x \mapsto Ax$ und $y \mapsto By$ bei 0 lokal topologisch äquivalent.*

Proof. Für den Fall $a = n, b = 0$ greift Satz 11.4. Im Fall $a = 0, b = 0$ können wir durch Zeitumkehrung $t \mapsto -t$ auf den ersten Fall zurückführen. Im allgemeinen Fall zerlegt man erst in zwei Teilräume von Dimension a und b , wendet in jedem Teilraum einen der beiden obigen Fälle an und hat lokale Homöomorphismen auf den Teilräumen, und die baut man dann zu einem zusammen. \square

Die Kombination von zwei Tatsachen:

1. hyperbolische Senken sind lokal equivalent;
2. die Menge $\mathcal{H}_{a,b}$ aller Matrizen, deren Eigenwerte negativen Realteil haben, ist offen (siehe Kapitel 9);

hat zur Folge, dass Störungen von hyperbolischen Senken wieder hyperbolische Senken sind. Analoges gilt für alle hyperbolischen Equilibrien. Hier ist die genaue Aussage des Satzes.

Satz 11.6. *Es sei $n \in \mathbb{N}$, $X \subset \mathbb{R}^n$ offen, $P \subset \mathbb{R}$ ein offenes Intervall, das 0 enthält, und $F : P \times X \rightarrow \mathbb{R}^n$ ein parameterabhängiges differenzierbares Vektorfeld. Für jedes $\lambda \in P$ setzen wir $F_\lambda : x \mapsto F(\lambda, x)$. Wenn $x_0 \in X$ ein hyperbolisches Equilibrium ist, dann existiert $\epsilon > 0$, sodass für jedes $\lambda \in P$ mit $|\lambda| < \epsilon$ ein Equilibrium x_λ von F_λ existiert, das lokal topologisch zu F_0 bei x_0 equivalent ist.*

Beweis. Es sei $G : P \times X \rightarrow \mathbb{R}^{n+1}$ die Funktion $(\lambda, x) \mapsto (\lambda, F_\lambda(x))$. Diese Funktion ist differenzierbar und an der Stelle $(0, x_0)$ ist die Ableitung invertierbar: in der ersten Spalte steht der erste Einheitsvektor, und der $n \times n$ -Block rechts unten ist $F'_0(x_0)$. Wir wählen eine Umgebung U von x_0 und $\epsilon > 0$, sodass G eingeschränkt auf $(-\epsilon, \epsilon) \times U$ invertierbar ist (Satz der inversen Funktion). Wir wenden G^{-1} auf die Gerade $(t, 0)$ an und erhalten eine Kurve $(t, \xi(t))$. Für alle $|\lambda| < \epsilon$ gilt ist dann $F_\lambda(\xi(\lambda)) = 0$, das heißt $\xi(\lambda)$ ist Equilibrium von F_λ . Die matrixwertige Funktion $H : \lambda \mapsto F'_\lambda(\xi(\lambda))$ ist stetig. Wenn $H(0) \in \mathcal{H}_{a,b}$ ist, kann man daher ein ϵ' wählen, sodass $H(\lambda) \in \mathcal{H}_{a,b}$ ist für alle $\lambda \in (-\epsilon', \epsilon')$. Diese Equilibrien sind dann alle lokal topologisch equivalent. \square

12 Stabile Fixpunkte

Es sei $X \subset \mathbb{R}^n$ eine offene Menge und $F : X \rightarrow X$ stetig differenzierbar. Die Rolle der Equilibrien übernehmen im diskreten Fall die Fixpunkte von F . Es sei $x_0 \in X$ ein Fixpunkt von F . Analog wie im diskreten Fall zeigt das diskrete dynamische System, das durch F gegeben ist, manchmal ein ähnliches Verhalten wie die Linearisierung bei x_0 , die durch $\mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto F'(x_0)$ gegeben ist.

Wir bezeichnen den Fixpunkt als hyperbolischen Fixpunkt, wenn die Eigenwerte der Jacobi-Matrix $F'(x_0)$ (reell oder komplex) Betrag ungleich 1 haben. Hier ist die diskrete Version von Hartmann/Grobmann (Spezialfall):

Satz 12.1. Wir nehmen an, dass $x_0 \in X$ ein hyperbolischer Fixpunkt ist, sodass die Jacobi-Matrix $F'(x_0)$ nur Eigenwerte mit Betrag kleiner als 1 haben. Dann existiert eine offene Umgebung U von x_0 , eine Konstante $c \in (0, 1)$ und eine differenzierbare Funktion $G : U \rightarrow \mathbb{R}$ mit den folgenden Eigenschaften:

- a) g ist positiv außer bei x_0 und $g(x_0) = 0$.
- b) Für alle $y \in U$ gilt $F(y) \in U$, $g(F(y)) \leq cg(y)$.

Beweis. Der Beweis geht analog wie im kontinuierlichen Fall: zuerst konstruieren wir eine positive definite quadratische Form für die Linearisierung, deren Wert bei Multiplikation des Arguments mit der Matrix abnimmt. Dann wählen wir U so klein, dass die Linearisierung nahe genug bei F ist, sodass die Ungleichung (b) noch richtig ist.

Die quadratische Form g kann man (im Fall $x_0 = 0$) wählen als

$$g(x) := \sum_{i=0}^{\infty} \|F'(x_0)^i(x)\|^2.$$

□

Die Existenz einer Funktion g wie in Satz 12.1 hat zur Folge, dass x_0 ein asymptotisch stabiler Fixpunkt ist.

Es sei $k \in \mathbb{N}$. Eine k -tupel $(x_1, \dots, x_k) \in X$ heißt k -Zykel, wenn $F(x_i) = x_{i+1}$ für $i = 1, \dots, k-1$ und $F(x_k) = x_1$ gilt. Jedes der x_i ist dann ein Fixpunkt der Funktion $F^k : X \rightarrow X$. Die Begriffe hyperbolisch, stabil, und asymptotisch stabil, sowie der obige Satz lassen sich verallgemeinern auf k -Zykel. Man kann zeigen, dass die k Fixpunkte von F^k entweder alle oder keiner stabil bzw. asymptotisch stabil sind, und dass die Eigenwerte der Jacobi-Matrix von F^k übereinstimmen.

Beispiel 12.2. Es sei $\lambda \in [2, 4]$ und $F : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \lambda x(1-x)$. Wir untersuchen die rekursive Folge $(x_k)_k$, $x_{k+1} = F(x_k)$, in Abhängigkeit von λ und vom Startwert x_0 .

Die Funktion F bildet $(-\infty, 0]$ auf sich selbst ab. In diesem Bereich ist F monoton steigend und erfüllt $F(x) < x$ für alle $x \in (-\infty, 0)$. Wenn $x_0 \in (-\infty, 0)$ ist, dann ist daher die Folge $(x_k)_k$ streng monoton fallend. Sie kann nicht konvergent sein, denn in $(-\infty, 0)$ hat F keinen Fixpunkt, daher geht die Folge gegen $-\infty$.

Wenn der Startwert x_0 grösser als 1 ist, dann ist $x_1 \in (-\infty, 0)$ und die Folge geht ebenfalls gegen $-\infty$.

Das Intervall $[0, 1]$ wird auf sich selbst abgebildet, die Folge ist daher beschränkt und hat zumindest einen Häufungspunkt. Die Randpunkte sind einfach: 0 ist ein Fixpunkt und führt zu einer konstanten Folge, und 1 wird auf 0 abgebildet und die Folge ist ab dem zweiten Glied konstant. Falls $\lambda = 2$ ist, gilt $F(x) < x$ für $x \in (0, 1)$, daher konvergiert die Folge gegen den linken Randpunkt 0.

Für $\lambda \in (2, 4]$ existiert ein zweiter Fixpunkt an der Stelle $y_1 := \frac{\lambda-1}{\lambda}$. Die Ableitung F' ist an dieser Stelle gleich $2 - \lambda$, also ist y_1 für $\lambda \neq 3$ ein hyperbolischer Fixpunkt. Für $\lambda \in (2, 3)$ ist er asymptotisch stabil.

Sei nun $\lambda > 3$. Dann hat F^2 vier Fixpunkte. Darunter sind die zwei schon bekannten Fixpunkte von F . Die anderen beiden

$$y_2 = \frac{\lambda^2 + \lambda + \sqrt{\lambda^2 - 2\lambda - 3}}{2\lambda}, y_3 = \frac{\lambda^2 + \lambda - \sqrt{\lambda^2 - 2\lambda - 3}}{2\lambda}$$

bilden einen Zweierzyklus (y_2, y_3) . Das Produkt der beiden Ableitungen ist gleich $-\lambda^2 + 2\lambda + 4$. Falls $\lambda < \sqrt{6} + 1 \approx 3.45$ ist, ist der 2-Zyklus asymptotisch stabil.

Für $\lambda > \sqrt{6} + 1$ ist der 2-Zyklus ebenfalls hyperbolisch, aber instabil. Man kann zeigen, dass F dann einen 4-Zyklus hat, der für kleine λ (aber immer noch größer als $\sqrt{6} + 1$) asymptotisch stabil ist. Wird λ größer, dann wird der 4-Zyklus wieder instabil und es entsteht ein stabiler 8-Zyklus. Dieser wird dann wieder instabil etc.

Die Abstände bis zur nächsten Periodenverdopplung werden immer kürzer. Die Folge der Quotienten der Abstände konvergiert, und der Grenzwert heisst *Feigenbaum-Konstante*.

Bei $\lambda \approx 3.57$ existieren alle Zyklen, deren minimale Periode eine Potenz von 2 ist. Geht man mit λ noch weiter nach oben, so entstehen Zyklen mit minimaler Periode k für jede positive ganze Zahl k .

In welcher Reihenfolge entstehen die minimalen Perioden von Zyklen? Das wird durch folgenden Satz (den wir nicht beweisen werden) festgelegt.

Satz 12.3 (Sharkowsky). *Es existiert eine Totalordnung \prec auf der Menge \mathbb{Z}_+ , sodass für alle Zahlen $m \prec n$ gilt:*

- 1) *Wenn eine stetige Funktion $F : \mathbb{R} \rightarrow \mathbb{R}$ einen Zykel von minimaler Periode n hat, dann hat f auch einen Zykel von minimaler Periode m .*
- 2) *Es existiert eine stetige Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, die einen Zykel von minimaler Periode m , aber keinen Zykel von minimaler Periode n hat.*

Die Sharkowsky-Ordnung hat eine kleinste Zahl, 1, und eine größte Zahl, 3. Jede Funktion, die überhaupt einen Zykel hat, hat auch einen Fixpunkt. Jede Funktion mit einem Zykel mit minimaler Periode 3 hat Zyklen von beliebiger minimaler Periode.

Wir brauchen noch zwei Verallgemeinerungen für Funktionen, die nicht immer iteriert werden können. Es sei $n \in \mathbb{N}$, $U \subset V$ offene Mengen, und $F : U \rightarrow V$ offene Mengen. Wir definieren eine möglicherweise endliche, möglicherweise unendliche Folge $(x_k)_k$ mit Startwert x_0 durch die Rekursion $x_{k+1} = f(x_k)$ für alle $k \in \mathbb{N}$, für die die rechte Seite definiert ist. Wenn $f(x_k) \notin U$, dann bricht die Folge nach dem k -ten Glied ab.

Lemma 12.4. *Es sei $y \in U$ ein hyperbolischer asymptotisch stabiler Fixpunkt. Dann existiert eine Umgebung $W \subset U$ von y , sodass $f(W) \subset W$ ist. (Es folgt, dass jede Folge mit Startwert in W gegen y konvergiert.*

Beweis. Wegen Satz (12.1) existiert eine offene Umgebung U' von y und eine stetige Funktion $g : U' \rightarrow \mathbb{R}$ mit $g(z) > 0$ für alle $z \in U' \setminus \{y\}$ und eine Konstante $c \in (0, 1)$, sodass $g(F(z)) \leq g(z)$ für alle $z \in U'$ gilt. Man wähle $C > 0$ so, dass g am Rand von U' nicht kleiner als C ist, und setze $W = \{z \mid g(z) < C\}$. \square

Lemma 12.5. *Es sei $n = 1$. Wir nehmen an, dass F monoton steigend ist. Es sei y ein Fixpunkt, sodass $F(z) > z$ gilt für alle $z < y$, $z \in U$. Dann konvergiert jede Folge mit Startwert kleiner als y gegen y .*

Beweis. Es sei $V = \{z \in U \mid z < y\}$. Es sei $(x_k)_k$ eine Folge mit Startwert in V . Mit Induktion nach k zeigt man $x_k \in V$ und insbesondere auch, dass die Folge nicht abbricht. Ausserdem ist sie monoton wachsend. Aus monoton wachsend und beschränkt folgt konvergent. Der Grenzwert ist ein Fixpunkt. In V existieren aber keine Fixpunkte, also kann der Grenzwert nur der Randpunkt y sein. \square

13 Die Poincare-Rückkehrabbildung

Zurück zum kontinuierlichen Fall: es sei $n \in \mathbb{N}$, $X \in \mathbb{R}^n$ offen und zusammenhängend, $F : X \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Es sei $g : X \rightarrow \mathbb{R}$ stetig differenzierbar, sodass es keine offene Teilmenge von X gibt, auf der g konstant ist. Die Nullstellenmenge einer solchen Menge nennen wir *Hyperfläche*. Die Hyperfläche H heißt transversal zu F in einem Punkt $p \in H$, wenn $\delta_F(g)(p) \neq 0$ ist; geometrisch bedeutet das, dass der Vektor $F(p)$ nicht in der Tangentialhyperebene bei p liegt. Wenn H transversal zu F bei p ist, dann ist H auch transversal in einer Umgebung von x , weil die Transversalität durch eine Ungleichung mit einer stetigen Funktion definiert ist.

Transversale Hyperebenen sind schon einmal vorgekommen, nämlich bei der Konstruktion einer Transformation eines Vektorfelds bei einem Nicht-Equilibrium zu einem konstanten Vektorfeld. Es sei $p_0 \in H$ ein Punkt der transversalen Hyperfläche. Wir haben eine Umgebung U von p_0 konstruiert und einen Homöomorphismus $\tau : U \rightarrow V$ von einer Umgebung U nach $V \subset \mathbb{R}^n$ offen, der H in eine Hyperebene und das Vektorfeld in ein konstantes Vektorfeld transformiert. Wir können ohne Beschränkung der Allgemeinheit annehmen, dass die Hyperebene gegeben die erste Koordinatenhyperebene $E := \{(0, \dots, x_n) \in U \mid x_i \in \mathbb{R}\}$ ist, und dass das Vektorfeld konstanten Wert $(1, 0, \dots, 0)$ hat.

Es sei $U \subset X$ offen und $H \subset U$ eine Hyperebene, die in jedem Punkt transversal ist. Die Poincare-Rückkehrabbildung ordnet jedem Punkt p in H dem Punkt $p' \in H$ zu, bei dem der Orbit durch p das erste Mal wieder auf H trifft:

$$r := \min\{t > 0 \mid \varphi(t, p) \in H, \Phi(p) = \varphi(r, p)\}.$$

Wenn der Orbit durch p nicht mehr auf H trifft, dann ist $\Phi(p)$ nicht definiert.

Satz 13.1. *Der Definitionsbereich der Poincare-Rückkehrabbildung ist offen. Die Abbildung ist stetig differenzierbar. Ihre Ableitung hat in jedem Punkt positive Determinante.*

Proof. Es sei $p \in H$ ein Punkt, auf dem Φ definiert ist, und $r := \min\{t > 0 \mid \phi(t, p) \in H\}$. Es sei $\tau : U \rightarrow V$ eine Transformation, die H auf eine horizontale Hyperebene E und F auf ein vertikales Vektorfeld transformiert (für die Definition muss U eventuell durch eine kleinere Umgebung ersetzt werden; das macht aber nichts aus, weil die Aussagen des Satzes

nicht von der Wahl der Umgebung abhängen). Die Flussabbildung φ_r ist stetig differenzierbar und definiert in einer offenen Umgebung von U_1 von x . Wir setzen $V := \tau(U_1)$. Wir definieren die Abbildung $P : V_1 \rightarrow V$ als $\tau \circ \phi_r \circ \tau^{-1}$. Die Poincare-Abbildung steckt als Teilfunktion in P drin: wenn wir P schreiben als $(x_1, x_{2\dots n}) \rightarrow (P_1(x_1, x_{2\dots n}), P_{2\dots n}(x_1, x_{2\dots n}))$, dann ist $\Phi = \tau^{-1} \circ P_2|_E \circ \tau$. Es folgt, dass Φ auf einer offenen Umgebung definiert ist und fort stetig differenzierbar ist.

Weil τ ein Diffeomorphismus ist, ist die Jacobi-Determinante von τ im ganzen Definitionsbereich entweder positiv oder negativ. Die differenzierbare Abbildung ϕ_r hat Jacobi-Determinante ungleich 0, weil sie invertierbar ist (die inverse Abbildung ist ϕ_{-r}). Für beliebige t ist die Jacobi-determinante von ϕ_t ungleich 0. Die Funktion $t \mapsto \det(\phi'_t)(x)$ hat also keine Nullstellen, ist stetig und hat bei $t = 0$ den Wert 1 – also ist die Jacobi-Determinante bei von ϕ_r positiv. Daher ist die Jacobideterminante von P . Die erste Zeile von P' ist aber gleich dem transformierten Vektor $(1, 0, \dots, 0)$. Daher ist auch die Jacobideterminante von Φ positiv. \square

Bemerkung 13.1. Die folgende Umkehrung von Satz 13.1 gilt ebenfalls: wenn $V \subset V \subset \mathbb{R}^n$ offen und einfach zusammenhängend sind (d.h. eine Löcher haben), und $G : V \rightarrow U$ eine stetig differenzierbare Abbildung mit positiver Jacobi-Determinante ist, dann existiert ein Vektorfeld und eine transversale Hyperebene, sodass G die Poincare-Abbildung ist.

Die periodischen Orbits von Φ sind gleichzeitig die periodischen Orbits von F , die H schneiden. Jeder k -Zykel von F entspricht einer geschlossenen Integralkurve, die H k mal schneidet. Wir können also nun auch von stabilen / asymptotisch stabilen / hyperbolischen k -Zyklen sprechen. Hyperbolische asymptotisch stabile Zyklen sind sogar strukturell stabil: wenn man ein Vektorfeld mit einem hyperbolischen asymptotisch stabilen Zyklus geringfügig stört, bekommt man wieder einen hyperbolischen asymptotisch stabilen Zyklus, der zum ursprünglichen topologisch äquivalent ist.

Bemerkung 13.2. Die Tatsache, dass asymptotisch stabile Zyklen strukturell stabil sind, ist eine mathematische Erklärung dafür, dass in vielen dynamischen Systemen, die von so vielen Parametern abhängen, dass man immer von Störungen ausgehen muss, periodische Phänomene auftreten.

Im Gegensatz dazu ist das Modell von Lotka/Volterra in Beispiel 11.1 nicht strukturell stabil. Wenn man hier die Differentialgleichung nur geringfügig stört, hat man nicht mehr das gleiche Verhalten (dass fast jeder Orbit stabil ist).

13.1 Der Satz von Poincare/Bendixsohn

In diesem Kapitelabschnitt nehmen wir $n = 2$ an. Dieser Fall ist speziell, weil die Poincare-Rückkehrabbildung streng monoton steigend ist. Durch die Monotonie kann man auch dann Stabilität von Fixpunkten nachweisen, wenn man nichts über die Jacobi-Matrix von Φ weiß.

Satz 13.2. *Es sei $U \subset \mathbb{R}^2$ offen, $F : U \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld, und $K \subset U$ ein kompakte Teilmenge. Wir nehmen an, dass F keine Equilibrien in K*

hat. Es sei $p \in K$ und $f : \mathbb{R} \rightarrow K$ die Lösung des Anfangswertproblems mit Anfangswert $f(0) = p$; wir nehmen an, dass die Lösung den Rand von K nicht erreicht. Dann existiert ein periodischer Orbit Z , sodass für jedes $\epsilon > 0$ ein $t > 0$ existiert, sodass der Abstand von $f(t)$ und z kleiner als ϵ ist.

Proof. Es sei $\delta > 0$. Wir überdecken K durch offene Mengen U_i und wählen in jeder offenen Teilmenge U_i eine transversale Hyperfläche H_i , sodass folgendes gilt:

- Für jedes i existiert eine Transformation $\tau_i : U_i \rightarrow V_i$, die H_i in eine horizontale Ebene E_i und F in das konstante vertikale Vektorfeld transformiert.
- die vertikale Höhe von V_i ist kleiner oder gleich δ .
- Jede senkrechte Gerade, die V_i schneidet, schneidet sie auch in E_i .

Wegen der Kompaktheit von K können wir annehmen, dass die Überdeckung endlich ist. Es sei N die Anzahl der Teilmengen.

Aus der zweiten Bedingung für die offenen Mengen U_i folgt, dass die Bahn höchstens eine Zeit δ in U_i bleiben kann. Es folgt, dass es zumindest eine offene Teilmenge – sagen wir U_1 – geben muss, die unendlich oft von C aufgesucht wird, und für die der Limes inferior der Umlaufzeiten, also der Zeitspannen, die benötigt werden, um wieder auf H_1 zu treffen, kleiner als $(N + 1)\delta$ ist. Also gibt es auch eine unendliche Folge $(p_k)_k$, die eine Rekursionsgleichung mit der Poincare-Abbildung erfüllt. Die Folge ist beschränkt und monoton (steigend oder fallend) bezüglich einer Totalordnung auf H , die von einer Parametrisierung induziert ist. Also ist die Folge konvergent. Wir definieren q als den Grenzwert und t_i so, dass $f(t_k) = p_k$ ist.

Da der Limes inferior der Folge Umlaufzeiten $(t_{i+1} - t_i)_i$ beschränkt ist, existiert eine konvergente Teilfolge gegeben durch $(i_j)_j$. Es sei $S := \lim_{j \rightarrow \infty} (t_{i_j+1} - t_{i_j})$ der Grenzwert der Umlaufzeiten. Dann ist

$$\varphi(S, q) = \lim_{i \rightarrow \infty} \varphi(t_{i_j+1} - t_{i_j}, p_{i_j}) = \lim_{i \rightarrow \infty} p_{i_j+1} = q.$$

Insbesondere ist der Orbit durch q geschlossen. Das ist unser Z . Der Abstand von p_i zu Z geht gegen Null, und der Abstand von jedem anderen Punkt $f(t)$, $t > t_i$ ist beschränkte durch den Abstand von p_i zu Z mal e^{LS} , wobei L eine Lipschitz-Konstante von F ist; also geht auch dieser Abstand gegen Null. \square

Ein Zykel, der die Eigenschaft wie oben hat, dass sich ein anderer Orbit asymptotisch an diesen annähert, heisst Grenzykel. Ein Grenzykel muss nicht asymptotisch stabil sein. Um hier ein Beispiel zu konstruieren, werden wir ein Vektorfeld zu einer vorgegebenen Poincare-Abbildung konstruieren. Nicht ganz, wir geben nur das asymptotische Verhalten von Φ vor – Fixpunkte, und zwischen den Fixpunkten, ob $\Phi(x)$ größer oder kleiner als x ist.

Wir setzen $X = \mathbb{R}^2 \setminus \{(0,0)\}$, $U = (0, \infty) \times \mathbb{R}$, und $H = \{(x,0) \mid x > 0\}$. Es sei $g : (0, \infty) \rightarrow [0, \infty)$ eine Funktion mit vorgegebenen Nullstellen und vorgegebenen Vorzeichen zwischen den Nullstellen. Wir definieren das Richtungsfeld

$$F : X \rightarrow \mathbb{R}^2, (x, y) \mapsto (g(\sqrt{x^2 + y^2})x + y, g(\sqrt{x^2 + y^2})y - x).$$

Die Transformation in Polarkoordinaten $(x, y) = (r \cos(w), r \sin(w))$ ist

$$G : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}^2, (r, w) \mapsto (g(r)r, 1).$$

Die Lösungen haben daher konstante Winkelgeschwindigkeit gegen den Uhrzeigersinn, und der Radius $t \mapsto R(t)$ folgt der Differentialgleichung

$$\forall t : R'(t) = g(R(t))R(t)\forall t.$$

Das Phasendiagramm dieser skalaren Differentialgleichung richtet sich nach den Nullstellen und den Vorzeichen von g . Die Fixpunkte von Φ sind die Equilibrien der skalaren Differentialgleichung.

Um einen Grenzzykel, der nicht asymptotisch stabil ist, zu konstruieren, wählen wir eine Funktion, die eine Nullstelle hat, aber bei der Nullstelle das Vorzeichen nicht ändert: $g : r \mapsto (1 - r^2)^2$. Jeder Startwert innerhalb des Einheitskreises hat dann einen Orbit, der gegen den Grenzzykel – den Einheitskreis – “konvergiert”; Orbits mit Startwerten ausserhalb driften nach aussen und sind unbegrenzt.

Wenn das kompakte Gebiet K Quellen oder Senken enthält, gilt der Satz auch noch. Quellen wirken sich nur insofern aus, dass bei einem Startwert genau an der Quelle eine konstante Lösung herauskommt, aber das verträgt sich mit der Schlussfolgerung des Satzes. Bei Senken kann es vorkommen, dass die Lösung gegen die Senke konvergiert; das ist auch okay. Kritisch sind Sattelpunkte.

Eine auf ganz \mathbb{R} definierte Lösung $f : \mathbb{R} \rightarrow K$ heißt *klin*, wenn $\alpha := \lim_{t \rightarrow -\infty} f(t)$ und $\omega := \lim_{t \rightarrow +\infty} f(t)$ beide existieren, wenn ω nicht asymptotisch stabil ist, und wenn α nicht asymptotisch stabil für das umgedrehte Vektorfeld $-F$ ist. Also Sattelpunkte oder nichthyperbolische Equilibrien, die sich ähnlich wie Sattelpunkte verhalten. Wenn $\alpha = \omega$ ist, nennen wir die Lösung homoklin, und wenn $\alpha \neq \omega$, dann nennen wir sie heteroklin.

Beispiel 13.3. Es sei $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto y^2 - x^2 + x^3$. Wir definieren das Vektorfeld

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto \left(-\frac{\partial g(x, y)}{\partial y}, \frac{\partial g(x, y)}{\partial x} \right).$$

Dann ist $\delta_F(g)(x, y) = 0$ für alle $(x, y) \in \mathbb{R}$. Das bedeutet, der Wert von g ist entlang jeder Lösung konstant. Wir interessieren uns für Lösungen mit Wert 0.

Die Nullstelle Z von g enthält das einzige Equilibrium $p_0 = (0, 0)$; die Jacobi-Matrix bei p_0 zeigt, dass es sich um einen Sattelpunkt handelt. Der Sattelpunkt zerlegt Z in drei Kurven – siehe Figur 3. Nachdem auf diesen Kurven keine Haltestellen mehr sind, sind es drei Orbits – zusammen mit dem konstanten Orbit p_0 sind es also vier. Eine von den drei Kurven ist beschränkt: es ist ein homokliner Orbit, bei dem $\alpha = \omega = p_0$ ist.

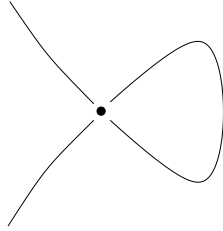


Abbildung 3: Die Nullstelle von g wird durch das Equilibrium in drei Teilkurven zerlegt. Eine davon ist ein homokliner Orbit.

Im obigen Beispiel gibt es unendlich viele Zyklen. Homokline Orbits können aber auch, gemeinsam mit $\alpha = \omega$, eine geschlossene Kurve bilden, an die sich andere Bahnen asymptotisch annähern. Hier ist die stärkste Version des Satzes von Poincaré/Bendixson.

Satz 13.4. *Es sei $U \subset \mathbb{R}^2$ offen, $F : U \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld, und $K \subset U$ kompakt. Es sei $p \in K$ und $f : \mathbb{R} \rightarrow K$ die Lösung des Anfangswertproblems mit Anfangswert $f(0) = p$; wir nehmen an, dass die Lösung den Rand von K nicht erreicht. Dann existiert eine zusammenhängende Menge Z , die entweder eine Vereinigung von endlich vielen Equilibrien, homoklinen und heteroklinen Orbits ist, oder unendlich viele Equilibrien enthält, sodass für jedes $\epsilon > 0$ ein $s > 0$ existiert, sodass für alle $t > s$ der Abstand von $f(t)$ zu Z kleiner als ϵ ist.*

Kline Orbits sind nicht strukturell stabil, sie gehen verloren wenn das Vektorfeld durch kleine Störungen geringfügig verändert wird. Zu jedem Sattelpunkt existieren vier spezielle Orbits, zwei ausgehende und zwei eingehende, alle anderen Orbits sind “hyperbelartig” und gehen an irgendeiner Seite vorbei. Für einen kline Orbit braucht es einen Orbit der gleichzeitig ausgehender Orbit eines Sattelpunktes und eingehender Orbit eines anderen oder auch desselber Sattelpunktes ist. Bei einer kleinen Störung trifft der ausgehende des ersten Sattelpunktes nicht mehr genau auf den zweiten und geht knapp vorbei.

Strukturelle Stabilität ist im zweidimensionalen Fall ein generisches Phänomen: fast alle Vektorfelder sind auf kompakten Teilmengen strukturell stabil. Um diese Aussage mathematisch exakt zu formulieren, bräuchte es eine Topologie auf der Menge aller stetig differenzierbaren Vektorfelder. In Dimension 2 ist ein Vektorfeld genau dann stabil, wenn es folgende Eigenschaften erfüllt.

- Alle Equilibrien sind hyperbolisch.
- Die Fixpunkte der Poincaré-Abbildungen, die periodischen Orbits entsprechen, sind alle hyperbolisch.
- Es gibt keine kline Orbits.

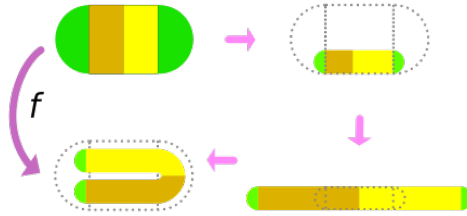


Abbildung 4: Die Hufeisen-Abbildung. Quelle: https://en.wikipedia.org/wiki/Horseshoe_map

13.2 Chaos

Der geometrische Grund für die sehr starken strukturellen Aussagen im zweidimensionalen Fall ist, dass die Poincare-Rückkehrabbildung eine monotone Abbildung ist, und das schränkt die Möglichkeiten stark ein. Bei einem dreidimensionalen Vektorfeld ist die Poincare-Rückkehrabbildung eine stetig differenzierbare Abbildung eines Gebietes in der Ebene auf sich selbst, deren Jacobi-Determinante überall positiv ist; und jede solche Abbildung kann als Poincare-Rückkehrabbildung auftreten. Und es gibt solche Abbildungen, deren Verhalten unter Iteration “chaotisch” ist.

Die Hufeisenabbildung h ist eine Abbildung von einer offenen Menge U , die homöomorph zu einer offenen Kreisscheibe ist, wird durch Abbildung 4 beschrieben werden. Sie ist injektiv und kann beliebig oft iteriert werden. Jedem Punkt können wir eine unendliche Folge von Symbolen aus $\{Y, B, G\}$ zuordnen, die angibt, in welchem der drei Teilgebiete die Iterationen liegen. Das grüne Gebiet ist ein Fangbereich, also sobald eine Folge einmal ein G hat, kommen nur mehr G 's. Das ist aber auch schon die einzige Bedingung: jede andere Folge ist möglich, speziell alle Folgen, die nur die Symbole Y und B enthalten.

Es sei $Z \subset U$ die Menge aller Punkte, deren Iteration immer gelb oder braun (also nie grün) sind, und die ausserdem im Bild von jeder Abbildung der Form $h^k : U \rightarrow U$, $k \in \mathbb{N}$, liegen. Jedem solchen Punkt p kann man eine Funktion $z : \mathbb{Z} \rightarrow \{Y, B\}$ zuordnen: $z(k)$ ist die Farbe des Punktes $h^k(p)$. Man kann nun zeigen, dass jede Funktion $z : \mathbb{Z} \rightarrow \{Y, B\}$ tatsächlich auftritt, und zwar immer nur für einen einzigen Punkt in Z .

Nun ist es eine Übungsaufgabe, die Zyklen mit festgelegter Minimalperiode zu bestimmen. Diese müssen in Z liegen und entsprechen daher $\{Y, B\}$ -Funktionen. Anwendung von h entspricht dem Shift nach rechts.

Die Menge Z ist abgeschlossen, und man kann zeigen, dass die Zyklen in Z dicht liegen. Das ist eines der drei Eigenschaften der Definition von Chaos (Dichtheit periodischer Orbits). Ein konkretes Vektorfeld, das ebenfalls chaotische Merkmale hat, ist die Lorenz-Gleichung, gegeben durch das Vektorfeld

$$L : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto (-s(x - y), rx - y - xz, xy - bz), \quad r = 28, s = 10, b = \frac{8}{3}.$$

Um einen Fangbereich zu bestimmen, definieren wir

$$g : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto rx^2 + sy^2 + s(z - 2r)^2.$$

Ihre vektorielle Ableitung ist

$$\partial_L(g)(x, y, z) = -2s((rx^2 + y^2 + b(z - r)^2 - br^2).$$

Es sei $E := \{p \mid \partial_L(g)(p) \geq 0\}$. Das ist das ein Voll-Ellipsoid, daher kompakt. Wir wählen M so, dass $K := \{p \mid g(p) \leq M\}$ die Menge E enthält – also $M := \max\{g(p) \mid p \in E\}$. Für jede Lösung f mit Startwert in K ist $f(t) \in K$. Um diese Behauptung zu beweisen, nehmen wir indirekt an, dass ein $t_1 > 0$ existiert, sodass $f(t_1) \notin K$ gilt. Die Funktion $h := g \circ f : [0, t_1] \rightarrow \mathbb{R}$ hätte dann folgende Eigenschaften:

- Wenn $h'(t) \geq 0$ ist, dann folgt $f(t) \in E \subset K$ und daher $h(t) \leq M$.
- $h(0) \leq M$ und $h(t_1) > M$.

So eine Funktion kann es nicht geben (warum nicht?). Also ist K ein Fangbereich, aus dem die Lösungen nicht mehr entkommen.

Das Lorenz-Feld hat in K drei Equilibrien. Alle drei sind Sattelpunkte. Der Nullpunkt hat den Typ $(2, 1)$, und zwei andere haben den Typ $(1, 2)$. Die beiden letzteren haben komplexe Eigenwerte; Orbits in der Nähe kreisen fast in einer Ebene spiralig nach außen, bis zu vom jeweils anderen Equilibrium “eingefangen werden”. Die Menge $Z := \bigcap_{t>0} \varphi(t, K)$ ist eine zusammenhängende abgeschlossene fraktale Menge, die im Abschluss von jedem Orbit liegt.

14 Ausblick: Differentialgeometrie

Differentialgeometrie könnte man definieren als das Teilgebiet der Geometrie, die Funktionen/Eigenschaften untersucht, die unter Diffeomorphismen erhalten bleiben. Ein wesentlicher Grundsatz in der Geometrie (den auch die theoretische Physik übernommen hat) ist: die geometrischen Sätze und Definitionen sind unabhängig von der Wahl des Koordinatensystems. Je nachdem welche Koordinatentransformationen bekommt man verschiedene Teilgebiete der Geometrie: euklidische Geometrie, affine Geometrie (bekannt aus LA-Vorlesung), projektive Geometrie, Topologie, algebraische Geometrie, Differentialgeometrie. Die Mitspieler bei der Differentialgeometrie sind: offene Teilmengen des \mathbb{R}^n , differenzierbare Abbildungen, glatte Hyperflächen oder Teilmannigfaltigkeiten, und vor allem: Vektorfelder.

Equilibrien, Zyklen, Stabilität etc. sind alles differentialgeometrische Eigenschaften von Vektorfeldern, die bei Koordinatentransformation erhalten bleiben. Auch die vektorielle Ableitung einer Funktion wird wieder in die vektorielle Ableitung der transformierten Funktion bezüglich des transformierten Vektorfelds transformiert. Also ist die vektorielle Ableitung ein Element der Differentialgeometrie. Hier ist ein erstes einfaches Beispiel für einen allgemeinen differentialgeometrischen Satz:

Proposition 14.1. *Für jedes Vektorfeld ist die Abbildung ∂_F linear über \mathbb{R} und erfüllt die Leibniz-Regel*

$$\partial_F(fg) = f\partial_F(g) + \partial_F(f)g.$$

für alle stetig differenzierbaren Funktionen f, g .

Diese Proposition läßt sich auch umkehren: wenn eine Operation, die jeder stetig differenzierbaren Funktion eine stetige Funktion zuordnet, linear ist, die Leibniz-Regel erfüllt, und stetig in Parametern von g ist, dann ist sie eine vektorielle Ableitung eines eindeutig definierten Vektorfeldes.

Als Prinzip gilt: man kann alles ableiten, was Werte in einem Vektorraum hat. Insbesondere gibt es die vektorielle Ableitung von Vektorfeldern: wenn F, G zwei stetig differenzierbare Vektorfelder sind, dann ist $\partial_F(G)$ wieder ein Vektorfeld - in der Regel nur stetig, nicht mehr differenzierbar. Um die Schwierigkeit zu vermeiden, dass man bei jedem Differenzieren eine Ordnung in der Differenzierbarkeit verliert, ist es üblich, von vornherein bei allen Funktionen und Vektorfeldern anzunehmen, dass die beliebig oft differenzierbar sind. Statt $\partial_F(G)$ schreibt man auch $[F, G]$.

Proposition 14.2. *Die vektorielle Ableitung von Vektorfeldern ist bilinear und erfüllt die Regeln*

$$[F, G] = -[G, F], \quad [F, [G, H]] + [G, [H, F]] + [H, [F, G]] = 0,$$

$$[G, fH] = f[G, H] + \partial_G(f)H$$

für alle Vektorfelder F, G, H und Funktionen f .

Wenn V ein Vektorraum ist und $[\cdot, \cdot] : V^2 \rightarrow V$ eine bilineare Abbildung ist, die die erste Zeile erfüllt, dann heißt $(V, [\cdot, \cdot])$ eine *Lie-Algebra*. Die Abbildung nennt man auch *Lie-Klammer*. Wenn G konstant ist, dann muss $[F, G] = \partial_F(G)$ nicht unbedingt Null sein (eine Transformation von F entlang G kann ja nicht-konstante Felder ergeben). Wenn F und G beide konstant sind, dann ist allerdings $[F, G] = 0$. Wir nennen zwei Vektorfelder F, G transversal, wenn für jeden Punkt p die beiden Vektoren $F(p)$ und $G(p)$ linear unabhängig sind. Hier ist eine Verallgemeinerung der in Kapitel 11 bewiesenen Behauptung, dass jedes Vektorfeld in einem Nicht-Equilibrium lokal equivalent zu einem konstanten Vektorfeld ist.

Satz 14.3. *Zwei transversale Vektorfelder F, G erfüllen genau dann $[F, G] = 0$, wenn es für jeden Punkt eine lokale Transformation gibt, die F und G gleichzeitig in konstante Felder transformiert.*

Nun zum Problem der Gradientenfelder. Wenn $X \subset \mathbb{R}^2$ und $f : X \rightarrow \mathbb{R}$ differenzierbar ist, dann nennen wir

$$\text{grad}(f) : (x, y) \mapsto \left(\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right)$$

das Gradientenfeld von f . Die stabilen Equilibrien sind genau die lokalen Maxima von f - diese Tatsache nützt man sich an der numerischen Optimierung aus, weil stabile Equilibrien leicht zu finden sind. Das Problem mit den Gradientenfeldern ist, dass die Eigenschaft, Gradientenfeld zu sein, nicht unter Koordinatentransformationen erhalten bleibt.

Beispiel 14.4. Es sei $f : (x, y) \mapsto xy$. Dann ist $F(x, y) = \text{grad}(f)(x, y) = (y, x)$. Wir setzen $(u, v) = (2x, y)$. In uv -Koordinaten ist $F^t(u, v) = (2v, u/2)$. Die zweite partielle Ableitung der ersten Komponente ist verschieden von der ersten partiellen Ableitung der zweiten Komponente, also ist F^t kein Gradientenfeld.

Die transformierte Funktion ist $f^t(u, v) = uv/2$. Das Gradientenfeld der transformierten Funktion stimmt nicht mit dem transformiertem Vektorfeld überein.

Eine geometrische Eigenschaft des Gradientenfeldes ist, dass der Vektor in jedem Punkt normal auf die Niveaulinie von f steht. Dass diese Eigenschaft bei Transformationen nicht erhalten bleibt, ist einleuchtend: Transformationen erhalten in der Regel nicht die Normalität.

Des Rätsels Lösung ist, dass die Werte des Gradientenfeldes nicht im Tangentialraum liegen, sondern im Dualraum dazu. Solche "Ko-Vektorfelder" heißen *Differentialformen*. Die Schreibweise für Differentialformen

$$\eta = F_1(x, y)dx + F_2(x, y)dy$$

passt zur Transformationsformel: für x und y ersetze man die entsprechenden Funktionen in u and v , und dx und dy wird ersetzt durch die Gradienten von diesen Funktionen. Vektorfelder und Differentialformen sind beides stetig differenzierbare Formen von X nach \mathbb{R}^2 , aber unterscheiden sich in vielen Eigenschaften und möglichen Operationen. Um diese kurz zu beschreiben, muss man erwähnen, dass die Differentialformen, die formal wie Vektorfelder aussehen, 1-Formen sind; differenzierbare Funktionen nach \mathbb{R} bezeichnet man als Null-Formen, und dann gibts auch noch 2-Formen (bis zur Dimension von X). In xy -Koordinaten läßt sich jede 2-Form schreiben als $f(x, y)dx \wedge dy$, also jede 2-Form ist eindeutig gegeben durch eine Funktion – trotzdem werden 0-Formen und 2-Formen anderes transformiert und müssen daher unterschieden werden.

- Vektorfelder kann man nur entlang Diffeomorphismen transformieren. Die Transformationsformel für Differentialformen braucht nur die inverse Abbildung. Für jede differenzierbare Abbildung $f : Y \rightarrow X$ und jede Differentialform η auf X ist die Transformation $f^*(\eta)$ eine Differentialform auf Y . X und Y müssen nicht einmal die gleiche Dimension haben.
- Die algebraischen Operationen für Vektorfelder sind die Lie-Klammer und die vektorielle Ableitung. Für Differentialformen hat man das Hack-produkt und die äussere Ableitung d . Die äussere Ableitung einer 0-Form ist eine 1-Form (nämlich der Gradient), und die äussere Ableitung einer 1-Form ist eine 2-Form. Die äussere Ableitung einer 2-Form ist immer 0. Die äussere Ableitung erfüllt $d^2 = 0$ und eine Analog zur Leibnizregel mit dem Hack-Produkt.
- Vektorfelder definieren Differentialgleichungen mit Integralkurven als Lösungen. Differentialformen kann man integrieren: für jede glatte Kurve C und Differentialform η existiert das Kurvenintegral $\int_C \eta$, eine reelle Zahl. Die Stokes-sche Integralformel verknüpft die Integration mit der äusseren Ableitung. Wenn η eine k -Form ist, und B eine kompakte Teilvarietät von Dimension $k + 1$ mit glattem Rand, dann gilt

$$\int_{\partial B} \omega = \int_B d\omega.$$

Im Fall $k = 0$ sagt die Stokes-sche Formel folgendes: wenn $C \subset X$ eine Kurve von p nach q ist, und $f : X \rightarrow \mathbb{R}$ differenzierbar, dann ist $\int_C dg = g(q) - g(p)$. Das Integral hängt also nicht vom Integrationsweg ab, falls die Differentialform ein Gradient ist.

Eine positive 2-Form ω definiert ein Maß auf X : jeder kompakten Teilmenge wird eine positive Zahl zugeordnet, nämlich das Integral von ω . Wenn $\omega = d\eta$ für eine 1-Form η ist, kann dieses Mass durch ein Kurvenintegral berechnet werden (Satz von Green, ein weiterer Spezialfall der Stokes-schen Formel).

Um eine bijektive Abbildung zwischen Vektorfelder und 1-Formen herzustellen, braucht man eine Art "lokale Dualität". Dafür gibt es zwei Möglichkeiten: eine Riemannsche Metrik oder eine symplektische Struktur.

Eine **Riemannsche Metrik** ist im wesentlichen ein lokales Skalarprodukt, positiv definit und symmetrisch. Eine Riemannsche Mannigfaltigkeit ist ein topologischer Raum, der lokal diffeomorph zu einer offenen Menge des \mathbb{R}^n ist und eine stetig differenzierbare Abbildung besitzt, die jedem Punkt eine symmetrische und positiv definite $n \times n$ Matrix zuordnet. Die Länge einer differenzierbaren kompakten Kurve ist eine positive Zahl. Die Länge der kürzesten Verbindung zwischen zwei Punkten erfüllt alle Axiome einer Metrik, und das macht jede Riemannsche Mannigfaltigkeit zu einem metrischen Raum. Die kürzeste Verbindung wird auf Kurven erzielt, die "Geodätische" heißen und gewisse Differentialgleichungen erfüllen müssen. Die Geodätischen in der Ebene sind Geraden, die Geodätischen auf der Kugel die Großkreise. Eine Riemannsche Metrik definiert auch Winkel zwischen Kurven oder Vektorfeldern oder Volumina.

Insbesondere definiert eine Riemannsche Metrik eine Bijektion zwischen Vektorfeldern und 1-Formen. Jeder differenzierbaren Funktion $f : X \rightarrow \mathbb{R}$ kann man dann zunächst eine Differentialform df und dann ein Vektorfeld zuordnen: das Gradientenfeld, wie wir es kennen. Der Satz, dass der Gradient normal auf die Niveauebenen $f = \text{konstant}$ steht, gilt allgemein für Riemannsche Mannigfaltigkeiten. Die stabilen Equilibrien der Gradientenfelds sind die lokalen Maxima von f .

Eine Frage in der Riemannschen Geometrie ist, ob zwei gegebene Riemannsche Mannigfaltigkeiten lokal equivalent sind: ob es einen Diffeomorphismus gibt, der Basispunkt in Basispunkt abbildet und der die erste Riemannsche Metrik in die zweite transformiert. Eine notwendige Bedingung ist, dass der Riemannschen Krümmungsskalar bei den Basispunkten den gleichen Wert hat. Dieser ist in der Ebene gleich 0 und bei einer Kugel vom Radius r gleich $\frac{1}{r^2}$, also sind Ebene und Kugel nicht lokal equivalent als Riemannsche Mannigfaltigkeiten.

In der allgemeinen Relativitätstheorie sind Riemannsche Pseudo-Metriken fundamental. Die sind genauso definiert wie Riemannsche Metriken, bis auf die Bedingung "positiv definit". Die Matrizen müssen vollen Rang haben, dürfen aber indefinit sein. Die "Lorentz-Metrik" lässt sich in einer Basis lokal schreiben als Diagonalmatrix mit Einträgen $-1, -1, -1, 1$. Tangentialvektoren mit positivem Quadrat heißen zeitartig, mit Quadrat Null lichtartig, und mit negativem Quadrat raumartig. Die Bewegung eines relativistischen Teilchens in der vierdimensionalen Raumzeit ist eine Kurve mit zeitartiger Ableitung. Wenn keine Kräfte oder Kollisionen auf das Teilchen wirken, dann erfüllt die Kurve die gleiche Differential-

gleichung wie eine Geodätische und heißt auch Geodätische. Allerdings ist sie nicht die kürzeste Verbindung, sondern die längste. Nun hat die Länge einer Geodätischen eine physikalische Interpretation: die Eigenzeit, also die Zeit, die das Teilchen messen würde, wenn es auf seiner Reise durch die Raumzeit eine Uhr mit sich führt. Daraus lässt sich das “Zwillingsparadoxon” ableiten.

Die zweite Möglichkeit, eine lokale Dualität zu definieren ist eine **symplektische Struktur**. Hier ordnet man jedem Punkt eine schiefsymmetrische und nichtsinguläre Matrix zu. Das geht nur, wenn der Raum X gerade Dimension hat (also etwa in der Ebene, aber nicht im dreidimensionalen Raum). Das Transformationsgesetz ist genau das gleiche wie bei 2-Formen ω , daher kann man auch sagen eine symplektische Struktur ist eine 2-Form mit bestimmten Eigenschaften (Nicht-Degeneriertheit und äußere Ableitung gleich Null). Wie im Fall einer Riemann-Metrik bekommt man eine Bijektion zwischen Vektorfeldern und 1-Formen. Ein Vektorfeld F heißt *hamiltonianisch*, wenn es die symplektische Struktur erhält, also wenn $\partial_F(\omega) = 0$ ist. Ein Satz besagt, dass das genau dann der Fall ist, wenn die entsprechende Differentialform ein Gradient ist, also von der Form df für eine skalare Funktion $f : X \rightarrow \mathbb{R}$. Wenn F ein hamiltonianisches Vektorfeld und f die entsprechende skalare Funktion ist, dann gilt immer $\partial_F(f) = 0$.

In der analytischen Mechanik kennt man viele Beispiele von diesen Paaren Funktion-Vektorfeld. Ein Vektorfeld F , das die symplektische Struktur erhält, wird in diesem Zusammenhang als *Symmetrie* bezeichnet. Für fixes $t \in \mathbb{R}$ ist $\varphi : X \rightarrow X$ ein Diffeomorphismus (mit Umkehrung φ_{-t}), und nachdem $\partial_F(\omega) = 0$ gilt, erhält jeder solche Diffeomorphismus die symplektische Struktur; streng genommen sind also die Flüsse die Symmetrien, aber man kann zeigen, dass jede eindimensionale Gruppe von Symmetrien der Fluss eines Vektorfelds ist; also ist es schon okay, hamiltonianische Vektorfelder und Symmetrien zu identifizieren. E. Nöther formulierte das folgende *physikalische Gesetz*:

Zu jeder Symmetrie läßt gehört eine Erhaltungsgröße; die Erhaltungsgröße ist eindeutig bis auf eine additive Konstante.

Beispiel 14.5. Die meisten physikalischen Gesetze sind translations-invariant: die Transformation $x_1 \mapsto x_1 + t$, alle anderen Koordinaten gleichbleibend, ist eine Symmetrie, nämlich der Fluss eines konstanten Vektorfeldes. Nach Nöther ist die dazugehörige Erhaltungsgröße der Impuls, genau genommen die Koordinate des Impulses Richtung x_1 . Für alle Bahnen, die translations-invarianten Gesetzen genügen, gilt, dass dieser Impuls konstant ist.

Die Erklärung im Modell der symplektischen Räume ist: eine Symmetrie ist ein hamiltonianisches Vektorfeld F . Die entsprechende Differentialform ist ein Gradient, also von der Form df , und dabei ist f bis auf eine additive Konstante eindeutig bestimmt. Diese Erklärung ist kein Beweis des Satzes von Nöther, sondern es ist eher umgekehrt: der Nöthersche Satz führte zur Entwicklung der symplektischen Geometrie als mathematisches Modell zur Beschreibung von physikalischen Modellen, die diesen Satz erfüllen.

Zur Beschreibung der zeitliche Entwicklung eines Systems nimmt man zusätzlich zur symplektischen Struktur auch eine sogenannte *Hamilton-Funktion* $h : X \rightarrow \mathbb{R}$ als gegeben an.

Physikalisch entspricht sie der Gesamtenergie, und die wesentlichste Eigenschaft ist, dass sie während der zeitlichen Entwicklung konstant bleibt. Die Hamilton-Funktion ist aber nicht irgendeine Erhaltungsgröße, sondern diejenige Erhaltungsgröße, deren zugehörige Symmetrie die Bewegung der Teilchen beschreibt: die Bahn $x : \mathbb{R} \rightarrow X$ eines Teilchens ist eine Lösung der Differentialgleichung $\forall t : \dot{x}(t) = -H(x(t))$.

Es wird Zeit für ein Beispiel. Wir beschreiben die Bewegung eines Federpendels auf einer geraden (vertikalen) Linie. Der Phasenraum ist zweidimensional, also \mathbb{R}^2 , mit Koordinaten p – Impuls – und q – Ort. Leider sind diese beiden Koordinaten nicht gerade unabhängig: wenn man die Ortsbewegung kennt, dann kennt man auch die Geschwindigkeit und daher auch den Impuls. Wir werden aber diese schon bekannte Beziehung nicht verwenden, wir tun praktisch vorläufig so, als ob die Masse variabel wäre. Die beiden Koordinaten können theoretisch unabhängig voneinander gemessen werden. Den Ort messen wir zum Beispiel durch ein Foto, auf dem sich das Pendel vor einer Skala befindet; und den Impuls kann man messen, indem man ein vergleichsmässig leichtes Teilchen mit dem Federpendel kollidieren lässt und schaut, in welche Richtung es wegfliegt.

Im zweidimensionalen Fall ist eine symplektische Struktur eine Volumensform, die nirgends Null ist. Wir setzen $\omega := dp \wedge dq$; hier liefert die Volumensform den bereits bekannten Flächeninhalt. Ein Vektorfeld $F : (p, q) \mapsto (F_1(p, q), F_2(p, q))$ ist genau dann hamiltonianisch bzw. volumserhaltend, wenn

$$\partial_F(dp \wedge dq) = \left(\frac{\partial F_1(p, q)}{\partial p} + \frac{\partial F_2(p, q)}{\partial q} \right) dp \wedge dq = 0$$

gilt; dieser Ausdruck heißt auch *Divergenz* des Vektorfeldes. Diese Vektorfelder sind Symmetrien in Bezug auf den Flächeninhalt: wenn die Fluss-Abbildung auf eine kompakte, von glatten Kurven begrenzte Teilmenge angewendet, dann erhält man eine kompakte Menge mit gleichem Flächeninhalt.

Es sei $f : \mathbb{R}^2$ eine skalare Funktion. Das dazugehörige Vektorfeld bekommt man durch Anwendung der schiefsymmetrischen Matrix, die zu Ω gehört:

$$F : (p, q) \mapsto \left(\frac{\partial f(p, q)}{\partial q}, -\frac{\partial f(p, q)}{\partial p} \right).$$

Man kann leicht zeigen, dass jedes solche Vektorfeld hamiltonianisch ist, und dass $\partial_F(f) = 0$ gilt.

Wir wählen nun als Hamilton-Funktion

$$h : \mathbb{R}^2 \rightarrow \mathbb{R}, (p, q) \mapsto \frac{p^2}{2m} + \frac{cq^2}{2};$$

die Konstanten m, c heißen *Masse* und *Federkonstante*. Das dazugehörige Vektorfeld ist

$$H : (p, q) \mapsto \left(\frac{\partial h(p, q)}{\partial q}, -\frac{\partial h(p, q)}{\partial p} \right) = \left(-cq, -\frac{p}{m} \right).$$

Wenn man das Vektorfeld mit -1 multipliziert, hat man die Bewegungsgleichungen $p' = -cq, q' = \frac{p}{m}$. Die zweite Gleichung liefert den bereits bekannten Zusammenhang zwischen

Impuls und Geschwindigkeit, den wir vorhin verläufig ignoriert haben. Die erste Gleichung ist equivalent zum Hooke'schen Gesetz. Die Differentialgleichung ist linear und läßt sich leicht lösen. Alle Lösungen außer der konstanten Lösung im Nullpunkt sind periodisch. Der Nullpunkt ist ein Equilibrium, sodass die Eigenwerte der Jacobi-Matrix auf der imaginären Achse liegen.