Dynamische Systeme

May 29, 2015

1 Einleitung

In diesem Text bedeutet der Ausdruck ‘dynamisches System’ nichts anderes als ein System von gewöhnlichen Differentialgleichungen. Warum wird also dieser Ausdruck hier verwendet? Es geht darum zu betonen, dass die gewöhnlichen Differentialgleichungen und ihre Lösungen hier auf eine bestimmte Art und Weise betrachtet werden sollen. Eine mathematische Aufgabe, die die gewöhnlichen Differentialgleichungen uns stellen ist, sie explizit zu lösen. Dabei sucht man nach Formeln für die Lösungen, wobei es erlaubt ist, Kombinationen von bestimmten ‘elementaren Funktionen’ wie Potenzen, der Exponentialfunktion, Sinus usw. zu verwenden. Es ist bekannt, das dies nicht bei allen Gleichungen geht und um weiter zu kommen kann man neue Funktionen einführen (z. B. elliptische Funktionen) die ihrerseits nichts anderes sind als Lösungen von bestimmten Differentialgleichungen, die man schon genauer untersucht hat. Es ist aber so, dass man die meisten gewöhnlichen Differentialgleichungen in keinem nützlichen Sinne ‘explizit’ lösen kann.

Was gibt es für Alternativen? Manche denken (insbesondere manche Wissenschaftler, die keine Mathematiker sind) dass es nur zwei Möglichkeiten gibt. Die erste ist, dass man die strenge Mathematik hinter sich lässt und zu heuristischen Näherungsverfahren übergeht. Z. B. glaubt man, dass in einer gegebenen Anwendung eine bestimmte Größe klein ist und man geht zu einer neuen Gleichung über, in dem man diese Größe in der Ausgangsgleichung gleich Null setzt. Die zweite Methode ist, dass man die Gleichungen diskretisiert und die diskreten Gleichungen die dadurch entstehen auf dem Computer löst. In beiden Fällen hat man eine Näherung eingeführt. Aus mathematischer Sicht hat man ein System von Gleichungen durch eine zweites ersetzt und es ist zu klären, was die Lösungen der beiden Systeme miteinander zu tun haben. Diese Methoden können oft zu guten Ergebnissen führen. Was hier betont werden soll ist, dass es auch eine andere Alternative gibt. (Oft ist es die beste Strategie, alle drei Methoden miteinander zu kombinieren.)

Was ist die dritte Alternative? Es geht darum, mathematisch strenge Aussagen über das qualitative Verhalten von Lösungen zu beweisen. In dem Zusammenhang ist es oft besser, anstatt Lösungen einzeln zu untersuchen, die Beziehungen zwischen verschiedenen Lösungen zu betrachten. Es ist auch hilfreich, das

Problem geometrisch zu betrachten. Ein System von gewöhnlichen Differentialgleichungen besteht aus Gleichungen der Form

$$\frac{dx_i}{dt} = f_i(t, x_j) \tag{1}$$

für reellwertige Funktionen $x_i(t), 1 \leq i \leq m$ und reellwertige Funktionen f_i von $m + 1$ Variablen. Zusammen definieren die f_i eine Abbildung mit Werten im $\mathbb{R}^m$, die wir mit f bezeichnen. Wenn f nicht von t abhängt nennt man das System autonom. In dieser Vorlesung werden wir uns hauptsächlich auf den Fall der autonomen Systeme konzentrieren, weil die Sichtweise der dynamischen Systeme in dem Fall besonders hilfreich ist. Wenn man will kann man den nichtautonomen Fall auf den autonomen reduzieren, in dem man das erweiterte System

$$\frac{dx_i}{dt} = f(y, x_j), \quad \frac{dy}{dt} = 1 \tag{2}$$

betrachtet. Die Existenz einer Lösung $x(t)$ des ursprünglichen Systems mit $x(t_0) = x_0$ ist äquivalent zur Existenz einer Lösung $(x(t), y(t))$ des erweiterten Systems mit $x(t_0) = x_0$ und $y(t_0) = t_0$. Mit dem geometrischen Standpunkt betrachtet man die x_i als Koordinaten eines Punktes im $\mathbb{R}^m$ und f_i im autonomen Fall als die Komponenten eines Vektorfeldes. Die Lösungen der Gleichungen sind die Integralkurven des Vektorfeldes. Es ist dann natürlich, die Gleichungen (1) im autonomen Fall als die vektorwertige Gleichung

$$\frac{dx}{dt} = f(x) \tag{3}$$

zu schreiben. Die Funktion f ist auf einer offenen Teilmenge G des $\mathbb{R}^m$ definiert. Anfangsbedingungen der Form $x_i(t_0) = a_i$ für eine Lösung $x_i(t)$ bedeuten, dass die Integralkurve sich zum Zeitpunkt t_0 im Punkt mit den Koordinaten a_i befindet. Diese Art, die Dinge zu sehen, ist typisch für die Denkweise der dynamischen Systeme. Die mathematischen Objekte sind die gleichen. Es ist nur so, dass man etwas anders über sie redet und denkt, weil man dadurch eine andere geometrische Intuition mobilisieren möchte. Es werden hier nur Systeme erster Ordnung betrachtet, weil es einfach ist, ein System der Ordnung k auf ein System erster Ordnung zu reduzieren in dem man die Ableitungen bis zur Ordnung $k - 1$ als neue Variablen einführt.

Um mit Lösungen umzugehen, die nicht explizit sind, muss man festlegen können, um welche Lösung es sich handelt. Das heißt, dass man wissen muss wie viele Lösungen es gibt und wie man sie parametrisieren kann. Die übliche Art dies zu tun ist das Anfangswertproblem und deshalb befasst sich der nächste Abschnitt mit diesem Thema.

Der Begriff 'dynamisches System' wird oft in einem breiteren Sinn verwendet, wobei andere Arten von Evolutionsgleichungen zugelassen werden. Es könnten z. B. bestimmte partielle Differentialgleichungen sein, oder Gleichungen mit Verzögerung. Dabei verwendet man eine Analogie zwischen den gewöhnlichen Differentialgleichungen und diesen anderen Gleichungen, wobei der euklidische

Raum durch einen unendlichdimensionalen Funktionenraum ersetzt wird. Es gibt große Unterschiede zwischen diesen Klassen von Gleichungen, wodurch die Analogien gefährlich sein können aber auch viele Ähnlichkeiten, wodurch die Analogien sehr nützlich sein können. Der folgende Text beschränkt sich auf den Fall von gewöhnlichen Differentialgleichungen.

2 Das Anfangswertproblem

Das Anfangswertproblem für gewöhnliche Differentialgleichungen gehört zum Stoff der Vorlesung Analysis II. Dieses Thema wird trotzdem hier behandelt, um die Methoden in Erinnerung zu rufen und einen Ausgangspunkt für Verallgemeinerungen bereitzustellen. Das fundamentale Ergebnis über lokale Existenz und Eindeutigkeit ist

Satz 1 Sei f eine stetige Funktion auf der Menge

$$\{(t,x) \in \mathbb{R} \times \mathbb{R}^m : t_0 \leq t \leq t_0 + a, |x - x_0| \leq b\} \tag{4}$$

mit Werten im $\mathbb{R}^m$, die eine Lipschitz-Bedingung bezüglich x erfüllt. Sei M eine Schranke für $|f|$ und $\alpha = \min\{a, b/M\}$. Dann hat die Gleichung $\frac{dx}{dt} = f(t,x)$ eine eindeutige Lösung auf dem Intervall $[t_0, t_0 + \alpha]$ mit $x(t_0) = x_0$.

Beweis Es wird eine Folge definiert, die letzten Endes gegen die gewünschte Lösung konvergiert. Sei $x_0(t) = x_0$. Wenn eine stetige Funktion x_n auf $[t_0, t_0+\alpha]$ definiert ist und $|x_n(t) - x_0| \leq b$ erfüllt sei

$$x_{n+1}(t) = x_0 + \int_{t_0}^t f(s, x_n(s))ds. \tag{5}$$

Diese Bedingungen definieren eine Folge $\{x_n\}$ von stetigen Funktionen auf dem Intervall $[t_0, t_0 + \alpha]$. Sie erfüllen die Ungleichungen

$$|x_{n+1}(t) - x_0| \leq \int_{t_0}^t |f(s, x_n(s))|ds \leq M\alpha \leq b. \tag{6}$$

Es kann durch Induktion gezeigt werden, dass

$$|x_{n+1}(t) - x_n(t)| \leq \frac{MK^n(t-t_0)^{n+1}}{(n+1)!} \tag{7}$$

für alle n, wo K eine Lipschitz-Konstante für f ist. Der Induktionsanfang ist klar. Für den Induktionsschritt benutzen wir, dass für $n \geq 1$

$$\begin{aligned} |x_{n+1}(t) - x_n(t)| &\leq \int_{t_0}^t |f(s, x_n(s)) - f(s, x_{n-1}(s)|ds \\ &\leq K \int_{t_0}^t |x_n(s) - x_{n-1}(s)|ds \leq \frac{MK^n(t-t_0)^{n+1}}{(n+1)!}. \end{aligned} \tag{8}$$

Es folgt, dass die Reihe $x_0+\sum_{n=0}^{\infty}(x_{n+1}(t)-x_n(t))$ gleichmässig konvergiert und wir definieren $x(t)$ als diese Summe. Es ist dann möglich in der Integralgleichung zum Limes $n\to\infty$ überzugehen, mit dem Ergebnis, dass

$$x(t)=x_0+\int_{t_0}^{t}f(s,x(s))ds. \tag{9}$$

Deshalb ist $x(t)$ eine Lösung der Differentialgleichung mit dem gewünschten Anfangswert. Um Eindeutigkeit zu zeigen, sei $y(t)$ eine beliebige Lösung der Gleichung mit dem gegebenen Anfangswert. Dann gilt, durch Induktion,

$$|x_n(t)-y(t)|\leq\frac{MK^n(t-t_0)^{n+1}}{(n+1)!}. \tag{10}$$

Für $n\to\infty$ konvergiert die rechte Seite dieser Ungleichung gegen Null, woraus folgt, dass $x(t)=y(t)$.

Es ist jetzt gezeigt worden, dass die Lösung auf dem Intervall auf dem sie im Satz konstruiert wurde eindeutig ist. In Wirklichkeit ist die Lösung der Gleichung (3) mit gegebenem Anfangswert auf jedem Intervall eindeutig, wo sie definiert ist, wenn f lokal Lipschitz ist. Wir betrachten zwei Lösungen x und y auf einem Intervall $[t_0,t_1)$ mit $x(t_0)=y(t_0)$. Hier darf t_1 auch unendlich sein. Sei t_* das Supremum der Zahlen t mit der Eigenschaft, dass $x=y$ auf dem Intervall $[t_0,t)$. Wegen der lokalen Eindeutigkeit ist $t_*>0$. Die Lösungen x und y sind gleich auf dem Intervall $[t_0,t_*)$ und deshalb durch Stetigkeit, sofern $t_*<\infty$, auf dem Intervall $[t_0,t_*]$. Wenn $t_*<\infty$ können wir das Anfangswertproblem mit Anfangzeit t_* und Anfangsbedingung $x(t_*)=y(t_*)$ betrachten. Aus der lokalen Eindeutigkeit kann dann geschlossen werden, dass $x=y$ auf einem Intervall der Form $[t_*,t_*+\alpha]$, ein Widerspruch zur Definition von t_*. Hier wurde nur der Fall $t\geq t_0$ betrachtet, aber ein ähnliches Argument gilt für $t\leq t_0$.

Ohne die Annahme einer lokalen Lipschitz-Bedingung gilt immer noch eine Aussage über die Existenz einer Lösung, die wir aber hier nicht beweisen. Es handelt sich um den Satz von Peano (vgl. [4], Abschnitt II.2). Die Eindeutigkeit gilt aber nicht mehr, wie man durch einfache Beispiele wie $\frac{dx}{dt}=x^\alpha$, $0<\alpha<1$, zeigen kann. Auch wenn eine lokale Lipschitz-Bedingung gilt kann die lokale Lösung im allgemeinen nicht zu einer globalen Lösung fortgesetzt werden, wie man am einfachen Beispiel $\dot{x}=x^2$ sieht. Die Lösung mit $x(0)=1$ ist $\frac{1}{1-t}$ und existiert nur bis $t=1$. Wegen der Eindeutigkeit kann man das maximale Existenzintervall einer Lösung mit einem gegebenen Anfangswert definieren. Dieses Intervall kann man durch ein Fortsetzungskriterium charakterisieren.

Bevor diese Aussage bewiesen wird sollen ein paar metrische Eigenschaften von Teilmengen des $\mathbb{R}^m$ diskutiert werden. Sei A eine abgeschlossene Teilmenge des $\mathbb{R}^m$. Für einen Punkt $x\in\mathbb{R}^m$ sei $d(x,A)=\inf_{y\in A}d(x,y)$. Für eine feste Teilmenge A ist die Funktion $d(x,A)$ stetig, wie jetzt bewiesen wird. Seien x_1 und x_2 Punkte des $\mathbb{R}^m$ und $\epsilon>0$. Es gibt einen Punkt $y\in A$ mit $d(x_2,y)\leq$

$d(x_2, A) + \frac{\epsilon}{2}$. Mit der Dreiecksungleichung folgt, dass

$$d(x_1, y) \leq d(x_1, x_2) + d(x_2, A) + \frac{\epsilon}{2} \leq d(x_2, A) + \epsilon \tag{11}$$

wenn $d(x_1, x_2) \leq \frac{\epsilon}{2}$. In diesem Fall ist also $d(x_1, A) - d(x_2, A) \leq \epsilon$. Mit dem gleichen Argument ist $d(x_2, A) - d(x_1, A) \leq \epsilon$. Damit ist die Stetigkeit von $d(x, A)$ bewiesen. Wenn $x \notin A$ gilt die Ungleichung $d(x, A) > 0$. Weil sonst gäbe es eine Folge $x_n \in A$ mit $d(x_n, x) \to 0$ für $n \to \infty$ und, weil A abgeschlossen ist, wäre $x \in A$, ein Widerspruch. Wenn A_1 und A_2 Teilmengen des $\mathbb{R}^m$ sind sei $d(A_1, A_2) = \inf_{y_1 \in A_1, y_2 \in A_2} d(y_1.y_2)$. Dieser Ausdruck ist symmetrisch in den zwei Argumenten, kann aber auch in der Form $d(A_1, A_2) = \inf_{y \in A_2} d(y, A_1)$ geschrieben werden. Nehmen wir jetzt an, dass A abgeschlossen ist und K kompakt und dass $A \cap K = \emptyset$. Dann ist $d(K, A)$ das Infimum einer stetigen Funktion auf der kompakten Menge K und muss das Minimum sein. Es folgt, dass $d(K, A) > 0$.

Betrachten wir die Gleichung (3), wobei die Funktion f auf einer offenen Teilmenge G des $\mathbb{R}^m$ definiert ist. Sei (t_-, t_+) das maximale Existenzintervall einer Lösung mit Anfangsbedingung $x(t_0) = x_0 \in G$. Wenn t_+ endlich ist, dann verlässt die Lösung jede kompakte Teilmenge K von G für $t \to t_+$. Um diese Aussage zu beweisen, nehmen wir an, es wäre nicht so. Dann gäbe es eine Folge $\{t_n\}$ mit $t_n \to t_+$ und $x(t_n) \in K$ für alle n. Sei $\epsilon = d(K, \mathbb{R}^m \setminus G)$. Dann ist $\epsilon > 0$. Sei G_1 die Menge aller Punkte x mit $d(x, K) < \frac{\epsilon}{2}$. G_1 ist offene Teilmenge von G und ihre abgeschlossene Hülle $\bar{G}_1$ ist kompakt und in G enthalten. Der Abstand zwischen K und dem Komplement von G_1 ist nicht kleiner als $\frac{\epsilon}{2}$. Deshalb ist die abgeschlossene Kugel vom Radius $\frac{\epsilon}{2}$ um $x(t_n)$ in G_1 enthalten für alle n. Andererseits ist f auf G_1 durch eine positive Zahl $M \geq 1$ beschränkt. Als Folge des Satzes über lokale Existenz existiert eine Lösung auf dem Intervall $[t_n, t_n + \epsilon/2M]$, die den gleichen Anfangswert hat wie die ursprüngliche Lösung. Es reicht, n so groß zu wählen, dass $\epsilon/2M > t_n - t_*$, um einen Widerspruch zur Definition von t_* zu bekommen. Dieser Beweis wurde für den Fall eines autonomen Systems geführt. Es gilt auch ein entsprechender Satz für Systeme die nicht notwendigerweise autonom sein müssen, mit einem ähnlichen Beweis.

3 Ein Beispiel: das fundamentale System der Virusdynamik

In dieser Vorlesung wird die allgemeine Theorie durch die Betrachtung von geeigneten Beispielen aus den Naturwissenschaften begleitet. In diesem Abschnitt wird ein solches Beispiel eingeführt. Das dynamische System, das dabei betrachtet wird, das fundamentale System der Virusdynamik, wird benutzt, um die Ausbreitung von Virusinfektionen im Körper zu modellieren. Die dadurch gewonnenen Erkentnisse haben zu bedeutenden Fortschritten in der Medizin beigetragen. Der biologische Hintergrund wird jetzt kurz skizziert. Mehr Einzelheiten finden man in [8]. Die Krankheit AIDS ist Anfang der 1980er Jahre bekannt geworden und nach ein paar Jahren hatte man das Virus, das die Krankheit verursacht, HIV, isoliert. Der anfängliche Optimismus, man könnte die Krankheit bald heilen hat sich als unberechtigt erwiesen. Nach einer Infektion mit HIV und einer kurzen Anfangsphase mit grippeähnlichen Symptomen bleibt die Krankheit meistens lange (etwa 10 Jahre) symptomfrei. Erst danach kommt die Krankheit AIDS zum Ausbruch. Früher hat man geglaubt, dass das Virus in dieser Zeit aus irgendeinem Grund inaktiv wäre aber inzwischen weiss man, dass diese Auffassung falsch war. Man hat um 1995 verstanden, dass in diesen 10 Jahren ein dynamischer Prozess stattfindet, in dem riesige Mengen von Virusteilchen produziert werden. Bei diesem Fortschritt haben mathematische Modelle eine wichtige Rolle gespielt. In diesem Zusammenhang sind die modernen Kombinationstherapien gegen AIDS entstanden. Man kann die Krankheit bis heute nicht heilen aber man kann jetzt mit Medikamenten ihre gefährlichen Auswirkungen beliebig lange unterdrücken.

Das Modell, das jetzt vorgestellt wird hat, obwohl es in der AIDS-Forschung verwendet wurde, keine besondere Beziehung zum HIV und kann benutzt werden, um viele Viruserkrankungen zu modellieren. Es gibt drei Variablen. Die Population von Zellen, die nicht mit dem Virus infiziert ist, heißt x, die Population von infizierten Zellen heißt y und die Anzahl von Virus-Teilchen heißt v. Die Gleichungen lauten

$$\dot{x} = \lambda - dx - \beta xv, \tag{12}$$

$$\dot{y} = \beta xv - ay, \tag{13}$$

$$\dot{v} = ky - uv. \tag{14}$$

Hier bedeutet der Punkt die Ableitung nach t und die Größen λ, d, β, k und u sind positive Konstanten. Wegen ihrer Interpretation sollen die Größen x, y und v positiv sein, obwohl das Gleichungssystem überall im $\mathbb{R}^3$ definiert und regulär ist. Die Phänomene, die den verschiedenen Parametern entsprechen sind folgende. Neue Zellen entstehen im Körper durch Zellteilung (λ), nichtinfizierte Zellen sterben (d), Viren infizieren Zellen (β), infizierte Zellen sterben (a), Viren werden durch infizierte Zellen freigesetzt (k), Viren werden eliminiert (u). Ein Eingreifen des Immunsystems wird nicht berücksichtigt. Bei HIV sind die beteiligten Zellen selbst Immunzellen (weisse Blutkörperchen), aber diese Tatsache spielt im Modell keine Rolle.

Genau genommen bedeutet die Größe v die Anzahl von Virus-Teilchen außerhalb der Zellen. Dieses Modell vernachlässigt die Tatsache, dass bei der Infektion einer Zelle die Anzahl der Virusteilchen außerhalb der Zelle sich um Eins verringert. Deshalb sollte die Gleichung für $\dot{v}$ einen zusätzlichen Term $-\beta xv$ enthalten. Es wird aber argumentiert, dass dieser Term im Vergleich zu anderen Termen in dieser Gleichung klein ist, so dass man ihn vernachlässigen kann. Aus mathematischer Sicht bekommt man auf diese Weise ein neues System, das wir als 'modifiziertes fundamentales Modell der Virusdynamik' bezeichnen. Man kann den Zusatzterm als $-\delta\beta xv$ schreiben, wobei der Parameter δ den Wert Null oder Eins annimmt.

Die Funktionen auf der Rechten Seite von (12)-(14) sind offensichtlich lokal

Lipschitz-stetig und deshalb kann man den Satz über lokale Existenz und Eindeutigkeit auf dieses System anwenden. Wegen der Interpretation der Unbekannten betrachtet man Anfangsdaten die positiv sind (d.h. x, y und v sind positiv) und man erwartet, dass die Lösungen positiv bleiben. Der Beweis dieser Eigenschaft ist nicht offensichtlich und wird jetzt vorgeführt.

Hilfssatz 1 Sei $(x(t), y(t), v(t))$ eine Lösung des Systems (12)-(14) auf einem Intervall $[t_0, t_1)$ mit $x(t_0) = x_0$, $y(t_0) = y_0$ und $v(t_0) = v_0$. Wenn x_0, y_0 und v_0 positiv sind, dann sind $x(t)$, $y(t)$ und $v(t)$ positiv für alle $t \in [t_0, t_1)$.

Beweis Wir nennen die Variablen x_i. Wenn es ein i und ein t gibt mit $x_i(t) = 0$ sei t_* das Infimum solcher t für alle i. Dann ist die Einschränkung der Lösung auf das Intervall $[t_0, t_*)$ positiv und $x_i(t_*) = 0$ für einen bestimmten Wert von i. Die Gleichung für x_i kann in der Form $\dot{x}_i = -x_i f(x) + g(x)$ geschrieben werden, wo $g(x)$ nicht negativ ist. Deshalb ist $\dot{x}_i \geq -x_i f(x)$ und $\frac{d}{dt}(\log x_i) \geq -f(x) \geq -C$, wo C eine positive Konstante ist. Dabei wird verwendet, dass die Lösung in einem Kompaktum bleibt. Es folgt, dass $x_i(t_*) \geq x_i(t_0)e^{-C(t_*-t_0)} > 0$, ein Widerspruch.

Es wird jetzt gezeigt, mit Hilfe des Fortsetzungskriteriums, dass alle Lösungen des Systems (12)-(14) mit positiven Anfangsbedingungen global in der Zukunft existieren. Dazu reicht es zu zeigen, dass alle Variablen auf einem beliebigen endlichen Intervall $[t_0, t)$ nach oben beschränkt sind. Die Summe der ersten zwei Gleichungen ergibt, dass $\frac{d}{dt}(x + y) \leq \lambda$ und dass $x(t) + y(t) \leq x(t_0) + y(t_0) + \lambda(t - t_0)$. Deshalb sind x und y auf jedem endlichen Intervall beschränkt. Die dritte Gleichung zeigt dann, dass $v(t)$ nicht schneller als linear wachsen kann und deshalb auch auf jedem endlichen Intervall beschränkt ist. In der Tat sind alle Variablen in der Zukunft global beschränkt. Man kann zeigen, dass

$$x(t) + y(t) \leq C_1 = \max\left\{x(t_0) + y(t_0), \frac{\lambda}{\min\{a, d\}}\right\}, \tag{15}$$

$$v(t) \leq C_2 = \max\left\{v(t_0), \frac{kC_1}{u}\right\}. \tag{16}$$

Insbesondere ist die Lösung global beschränkt. Diese Aussagen werden jetzt bewiesen. Nehmen wir an, dass es eine Zeit t gibt mit $x(t) + y(t) = \alpha > C_1$. Sei t_* das Infimum der Zeiten, wo diese Ungleichung gilt, mit α fest. Dann ist $\dot{x}(t_*) + \dot{y}(t_*) \geq 0$. Auf der anderen Seite implizieren die Gleichungen (12) und (13), dass diese Größe negativ ist, ein Widerspruch. Damit ist (15) bewiesen. Die Ungleichung (16) kann mir der gleichen Methode bewiesen werden, wobei die schon bekannte Schranke für $x + y$ benutzt wird. Es kann mit den gleichen Methoden gezeigt werden, dass die Lösungen des modifizierten Systems mit ($\delta = 1$) beschränkt sind und dass die Unbekannten die gleichen Schranken erfüllen wie im Fall $\delta = 0$. Wenn die Anfangsdaten die Ungleichungen $x_0 + y_0 \leq \frac{\lambda}{\min\{a,d\}}$ und $v_0 \leq \frac{k\lambda}{u\min\{a,d\}}$ erfüllen dann muss die ganze Lösung diese Ungleichungen erfüllen. Das heißt, wir haben eine invariante Teilmenge identifiziert.

4 Abhängigkeit von Anfangsbedingungen und Parametern

Wir betrachten eine Lösung $x(t) = \phi(t_0, t, x_0)$ der Gleichung $\dot{x} = f(t, x)$ mit Anfangsbedingung $x(t_0) = x_0$. Die Abbildung ϕ wird als Fluss des dynamischen Systems bezeichnet. Im autonomen Fall hängt $\phi(t_0, t, x_0)$ nur von $t - t_0$ und x ab und wir schreiben auch $\phi(t, x)$ als Abkürzung für $\phi(0, t, x)$. Die Lösung ist auf einem maximalen Existenzintervall (t_-, t_+) definiert, wobei t_- und t_+ von t_0 und x_0 abhängen dürfen. In diesem Abschnitt betrachten wir die Frage der Stetigkeit und Differenzierbarkeit der Abbildung ϕ. Mit anderen Worten geht es um die Frage, ob die Lösung des Anfangswertproblems stetig bzw. differenzierbar von den Anfangsdaten abhängt. Wir betrachten auch den allgemeineren Fall einer Gleichung $\dot{x} = f(t, x, z)$ wobei die Koordinaten des Punktes $z \in \mathbb{R}^k$ als Parameter betrachtet werden.

Fragen der Abhängigkeit von Anfangsdaten und Abhängigkeit von Parametern können oft durch Variablentransformationen ineinander überführt werden. Betrachten wir zuerst das Anfangswertproblem ohne Parameter. In dem wir $\tilde{t} = t - t_0$ und $\tilde{x} = x - x_0$ einführen bekommen wir das neue Problem $\frac{d\tilde{x}}{dt} = f(\tilde{t} + t_0, \tilde{x} + x_0)$ mit $x(0) = 0$. Hier werden t_0 und x_0 als Parameter betrachtet und die Anfangsbedingung ist fest. Umgekehrt kann man das Problem mit Parametern durch das Problem

$$\dot{x} = f(t, x, z), \dot{z} = 0; \quad x(t_0) = x_0, z(t_0) = z_0 \tag{17}$$

ohne Parameter in den Gleichungen ersetzen. Aus dem Grund kann man oft Aussagen über Probleme die Parameter enthalten auf Aussagen ohne Parameter zurückführen.

Satz 2 Sei f eine stetige Funktion auf der Menge

$$\{(t, x, z) \in \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^k : t_0 \leq t \leq t_0 + a, |x - x_0| \leq b, |z - z_0| \leq b'\} \tag{18}$$

mit Werten im $\mathbb{R}^m$, die eine Lipschitz-Bedingung bezüglich x erfüllt . Sei M eine Schranke für $|f|$ und $\alpha = \min\{a, b/M\}$. Dann hat die Gleichung $\dot{x} = f(t, x, z)$ eine eindeutige Lösung $\phi(t_0, t, x_0, z)$ auf dem Intervall $[t_0, t_0 + \alpha]$ mit $\phi(t_0, t_0, x_0, z)) = x_0$. Die Abbildung ϕ ist stetig.

Beweis Die Existenz ist schon bekannt aus Satz 1. Die neue Behauptung ist die stetige Abhängigkeit der Abbildung ϕ von t_0, x_0 und z. Wie schon erklärt reicht es, den Fall mit $t_0 = 0$ und $x_0 = 0$ zu beweisen. Der Beweis ist sehr ähnlich wie im Fall ohne Parameter. Wir definieren eine Folge auf dem Produkt des Intervalls $[0, \alpha]$ mit der Kugel mit Radius b' um z_0 durch $x_0(t, z) = 0$ und

$$x_{n+1}(t, z) = \int_0^t f(s, x_n(s), z) ds. \tag{19}$$

Die Funktionen x_n sind stetig. Sie erfüllen die gleichen Abschätzungen wie im Fall ohne Parameter und die Folge konvergiert gleichmässig gegen die gewünschte Lösung. Diese ist als gleichmässiger Grenzwert von stetigen Funktionen, stetig.

Wenn wir das entsprechende globale Problem betrachten wollen, müssen wir nur beachten, wie das maximale Existenzintervall von x_0 und z abhängt. Es kann passieren, dass $t_- = -\infty$ oder $t_+ = +\infty$. Deshalb ist es nützlich bei der Beschreibung solcher Größen die erweiterte Zahlengerade $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ zu verwenden. Die Vereinigung der maximalen Existenzintervalle für verschiedene Werte von t_0, x_0 und z ist eine offene Teilmenge. Es folgt insbesondere, dass der obere Endpunkt des maximalen Existenzintervalls, als Funktion mit Werten in $\bar{\mathbb{R}}$ betrachtet, eine unterhalbstetige Funktion von x_0 und z ist. Wir erinnern an die Definition.

Definition Eine Funktion F mit Werten in $\bar{\mathbb{R}}$ auf einem topologischen Raum X heißt unterhalbstetig wenn die Bedingung $F(x_1) > a$ für einen Punkt $x_1 \in X$ und eine reelle Zahl a impliziert, dass es eine Umgebung U von x_1 gibt, so dass $F(x) \geq a$ für alle $x \in U$. Eine Funktion F heißt oberhalbstetig wenn $-F$ unterhalbstetig ist.

Der untere Endpunkt des Existenzintervalls ist oberhalbstetig. Intuitiv bedeutet diese Aussage, dass wenn man Anfangsbedingungen und Parameter variiert die Existenzzeit nicht plötzlich schrumpfen kann. Die Existenzzeit kann aber plötzlich wachsen (t_+ ist nicht oberhalbstetig). Ein Beispiel ist $\dot{x} = x^2 - zx^3$ mit $x(0) = 1$. Für $z = 0$ ist $t_+ = 1$ aber für $z > 0$ ist $t_+ = \infty$. In dem Fall ist eine Lösung die einmal Null ist immer Null. Deshalb kann eine Lösung die positiv anfängt nie negativ werden. Auf der anderen Seite ist $\dot{x} < 0$ wenn $x > z^{-1}$, so dass die Lösung von oben beschränkt ist. Deshalb folgt globale Existenz in der Zukunft für $z > 0$ aus dem Fortsetzungskriterium.

Wir haben jetzt stetige Abhängigkeit für kontinuierliche Parameter z bewiesen. Entsprechende Ergebnisse gelten für einen diskreten Parameter, mit einem sehr ähnlichen Beweis. Sei $f_n(t, x)$ eine Folge von stetigen Funktionen, die für $t_0 \leq t \leq t_0 + a$ und $|x - x_0| \leq b$ definiert sind und eine Lipschitz-Bedingung bezüglich x erfüllen wobei die Lipschitz-Konstante K nicht von n abhängt und die Folge gleichmässig gegen eine Funktion $f(t, x)$ konvergiert. Die Funktion f ist stetig und erfüllt eine Lipschitz-Bedingung bezüglich x mit der gleichen Lipschitz-Konstante K. Sei Z die Teilmenge von $\mathbb{R}$, die aus den Punkten 0 und $\frac{1}{n}$ für natürliche Zahlen n besteht. Z ist kompakt. Wir bezeichnen einen Punkt von Z durch z. Sei $\tilde{f}(t, x, 0) = f(t, x)$ und $\tilde{f}(t, x, 1/n) = f_n(t, x)$. Dadurch wird eine Funktion $\tilde{f}(t, x, z)$ auf $[t_0, t_0 + a] \times \bar{B}_b(x_0) \times Z$ definiert. Sie ist stetig und erfüllt eine Lipschitz-Bedingung bezüglich x mit Konstante K. Wir betrachten jetzt eine Iteration wie im Beweis von Satz 2.

$$\tilde{x}_{l+1}(t, z) = \int_0^t \tilde{f}(s, \tilde{x}_l(s), z) ds. \tag{20}$$

Wir bekommen eine Folge $\tilde{x}_l(t, z)$ auf $[t_0, t_0 + a] \times Z$ die gleichmässig gegen einen stetigen Grenzwert $\tilde{x}(t, z)$ konvergiert. Da der Definitionsbereich kompakt ist, ist $\tilde{x}$ gleichmässig stetig. Wenn wir $x_n(t) = \tilde{x}(t, 1/n)$ und $x(t) = \tilde{x}(t, 0)$ definieren dann erfüllen diese Funktionen die Gleichungen $\dot{x}_n = f_n(t, x)$ und $\dot{x} = f(t, x)$ mit $x_n(t_0) = x(t_0) = x_0$. Wegen der gleichmässigen Stetigkeit von

$\tilde{x}$ konvergiert x_n gleichmässig gegen x. Aus diesem Ergebnis folgt durch eine geeignete Variablentransformation eine entsprechende Aussage für Folgen von Anfangsdaten. Seien $x_n(t)$ die Lösungen von $\dot{x} = f(x)$ mit $x(t_0) = x_{0,n}$ und $x_{0,n} \to x_0$ für $n \to \infty$. Dann konvergiert $x_n(t)$ gegen die Lösung von $\dot{x} = f(x)$ mit $x(t_0) = x_0$.

Mit den Aussagen über stetige Abhängigkeit können wir ein paar allgemeine Begriffe einführen, die nützlich sind bei der Untersuchung der globalen Eigenschaften von Lösungen eines dynamischen Systems. Betrachten wir ein dynamisches System, das auf einer Teilmenge G des $\mathbb{R}^m$ definiert ist und eine Lösung $x(t)$ dieses Systems, die auf dem Intervall $[t_0, \infty)$ definiert ist. Wenn es eine Folge $\{t_n\}$ in $[t_0, \infty)$ mit $t_n \to \infty$ und $x(t_n) \to y \in \mathbb{R}^m$ für $n \to \infty$, dann heißt y ω-Limespunkt der Lösung x. Die Menge aller ω-Limespunkte der Lösung x heißt ω-Limesmenge von x. Wenn eine Lösung $x(t)$ auf dem Intervall $(-\infty, t_0]$ definiert ist, es eine Folge $\{t_n\}$ gibt mit $t_n \to -\infty$ und $x(t_n) \to y \in \mathbb{R}^m$ dann heißt y α-Limespunkt der Lösung x. Die Menge aller α-Limespunkt der Lösung x heißt α-Limesmenge von x. Zu jeder Aussage über ω-Limespunkte gibt es eine entsprechende Aussage für α-Limespunkte. Um diese zu erhalten betrachtet man die Lösung $\tilde{x}(t) = x(-t)$. Aus diesem Grund werden wir nur die Aussagen über ω-Limespunkte beweisen und es dem Leser überlassen, diese in Aussagen für α-Limespunkte zu übertragen.

Die ω-Limesmenge ist abgeschlossen. Sei $\{y_n\}$ eine Folge von ω-Limespunkten einer Lösung x die gegen $y \in \mathbb{R}^m$ konvergiert. Dann gibt es Folgen $\{t_{n,k}\}$ mit $t_{n,k} \to \infty$ für $k \to \infty$ und $x(t_{n,k}) \to y_n$. Wenn $\epsilon > 0$ gibt es N mit der Eigenschaft, dass $|y_n - y| < \frac{\epsilon}{2}$ für alle $n \geq N$. Außerdem gibt es $k(n) \geq n$ mit der Eigenschaft, dass $|x(t_{n,k(n)}) - y_n| < \frac{\epsilon}{2}$. Deshalb ist $|x(t_{n,k(n)}) - y| < \epsilon$ für alle $n \geq N$ und $t_{n,k(n)} \to \infty$ für $n \to \infty$. Es folgt, dass y der ω-Limesmenge von x gehört und dass diese Menge abgeschlossen ist. Wenn die Lösung x beschränkt ist, dann ist die ω-Limesmenge kompakt. Weil wenn x beschränkt ist, ist die ω-Limesmenge von x auch beschränkt. Da sie abgeschlossen ist, ist sie kompakt. Wenn die Lösung x beschränkt ist, dann ist die ω-Limesmenge von x nicht leer und sie ist zusammenhängend. Sei $\{t_n\}$ eine beliebige Folge in $[t_0, \infty)$ mit $t_n \to \infty$ für $n \to \infty$. Weil diese Folge beschränkt ist, hat sie eine konvergente Teilfolge, die gegen $y \in \mathbb{R}^m$ konvergiert. Der Punkt y ist in der ω-Limesmenge von x und deshalb ist diese Menge nicht leer. Ein topologischer Raum X heißt zusammenhängend, wenn die Eigenschaften $X = U \cup V$ und $U \cap V = \emptyset$ für offene Teilmengen U und V von X implizieren, dass entweder U oder V leer ist. Sei jetzt X die ω-Limesmenge von x und nehmen wir an, dass U und V Teilmengen mit den soeben erwähnten Eigenschaften sind. Wir nehmen an, dass keine der beiden leer ist und bekommen einen Widerspruch. Sei also $y_1 \in U$ und $y_2 \in V$. Es existieren t_{2n} mit $|x(t_{2n}) - y_1| < \frac{1}{n}$, t_{2n+1} mit $|x(t_{2n}) - y_2| < \frac{1}{n}$ und $t_{2n} \leq t_{2n+1} \leq t_{2n+2}$ für alle n. Für n groß genug ist $t_{2n} \in U$ und $t_{2n+1} \in V$. Sei $t'_n \leq t_{2n+1}$ die erste Zeit nach t_{2n} für die $x(t'_n)$ nicht in U liegt. Eine solche Zeit existiert. $x(t'_n)$ kann nicht in V liegen, da in dem Fall $x(t)$ in V liegen würde für alle t in einer offenen Umgebung von t'_n. Auf diese Weise bekommen wir also eine Folge t'_n mit der Eigenschaft, dass $x(t'_n)$ im Komplement von $U \cup V$ liegt. Diese Folge hat eine Teilfolge, die nach $y \in \mathbb{R}^m$ konvergiert. Der Punkt y

ist ein ω-Limespunkt von x und liegt nicht in X, ein Widerspruch.

Wenn ein dynamisches System auf G definiert ist heißt eine Menge A vorwärts-invariant, wenn aus $x(t_0) \in A$ folgt, dass $x(t) \in A$ für alle $t \geq t_0$. Wenn $x_0 \in G$ ist die vorwärts-Integralkurve durch x_0 das Bild der Lösung auf $[t_0, t_+)$ mit $x(t_0) = x_0$. Eine Teilmenge A ist genau dann vorwärts-invariant wenn die vorwärts-Integralkurve durch jeden Punkt von A in A enthalten ist. Manchmal wird das Wort 'vorwärts' in diesen Begriffen weggelassen, wenn die Bedeutung aus dem Zusammenhang hervorgeht. Wenn x eine Lösung ist und die ω-Limesmenge von x in G liegt, dann ist diese ω-Limesmenge invariant. Um diese Aussage zu beweisen, sei y_0 ein Punkt in der ω-Limesmenge einer Lösung x. Sei $\{t_n\}$ eine Folge mit $x(t_n) \to y$. Sei $\tilde{x}_n(t) = x(t - t_n)$. Dann konvergiert $\tilde{x}_n(0)$ gegen y_0. Sei y die Lösung mit $y(0) = y_0$. Als Folge der stetigen Abhängigkeit der Lösungen von Anfangsdaten konvergiert $\tilde{x}_n(t)$ gegen $y(t)$. Da $x(t + t_n) = \tilde{x}_n(t)$ liegt $y(t)$ in der ω-Limesmenge von x für alle t für die diese Größe definiert ist.

Jetzt soll gezeigt werden, dass Lösungen auch differenzierbar von den Anfangsdaten und Parametern abhängen.

Satz 3 Sei die Funktion f in Satz 2 stetig differenzierbar. Dann ist der Fluss $\phi(t_0, t, x_0, z)$ auch stetig differenzierbar. Die Ableitungen erfüllen die linearen Differentialgleichungen

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial x_{0,j}}\right) = \frac{\partial f_i}{\partial x_k}\frac{\partial x_k}{\partial x_{0,j}}, \tag{21}$$

$$\frac{d}{dt}\left(\frac{\partial x_i}{\partial z_j}\right) = \frac{\partial f_i}{\partial x_k}\frac{\partial x_k}{\partial z_j} + \frac{\partial f_i}{\partial z_j}. \tag{22}$$

mit Anfangsdaten δ^i_j bzw. 0. Außerdem gilt

$$\frac{\partial x_i}{\partial t_0} = -\frac{\partial x_i}{\partial x_{0,j}} f_j. \tag{23}$$

In diesen Gleichungen, und an vielen Stellen im folgendem Text, wird die Einsteinsche Summenkonvention verwendet: über wiederholte Indizes wird summiert. Um diesen Satz zu beweisen benutzt man folgende Strategie. Nehmen wir vorläufig an, dass die Differenzierbarkeit gilt und dass das Ableiten nach den Anfangsdaten und Parametern mit dem Ableiten nach der Zeit vertauscht werden kann. Dann können wir mit der Kettenregel Differentialgleichungen herleiten, die die Ableitungen erfüllen. Wir nennen diese Gleichungen die Variationsgleichungen. Die Lösungen der Variationsgleichungen können benutzt werden, um die Diffenzierbarkeit zu beweisen. Man sagt, dass man die Gleichungen zuerst formal differenziert. Das Ergebnis benutzt man dann, um zu zeigen, dass man tatsächlich differenzieren darf. Die letzte Formel im Satz bekommt man in dem man die Identität

$$\phi(t_0, t, \phi(t, t_0, x_0, z), z) = x_0 \tag{24}$$

formal nach t_0 differenziert. Um den Satz zu beweisen benutzen wir folgende Form des Mittelwertsatzes. Eine Teilmenge G von $\mathbb{R}^m$ heißt konvex wenn $x \in G$, $y \in G$ und $0 \leq s \leq 1$ zur Folge haben, dass $sx + (1 - s)y \in G$.

Hilfssatz 2 Sei f eine stetige Funktion auf $(a, b) \times G$ mit G eine konvexe Teilmenge des $\mathbb{R}^m$ die stetige partielle Ableitungen bezüglich der Komponenten des zweiten Arguments besitzen. Dann existieren stetige Funktionen $f_k(t, x_1, x_2)$ auf $(a, b) \times G \times G$ mit den Eigenschaften, dass

$$f_k(t, x, x) = \frac{\partial f(t, x)}{\partial x_k} \tag{25}$$

und

$$f(t, x_2) - f(t, x_1) = f_k(t, x_1, x_2)(x_{2,k} - x_{1,k}). \tag{26}$$

Man kann f_k durch folgende Formel definieren

$$f_k(t, x_1, x_2) = \int_0^1 \frac{\partial f(t, sx_2 + (1-s)x_1)}{\partial x_k} ds. \tag{27}$$

Beweis Sei $F(s) = f(t, sx_2 + (1-s)x_1)$ für $0 \leq s \leq 1$. Da G konvex ist, ist F wohldefiniert. Es gilt

$$\frac{dF}{ds} = \frac{\partial f}{\partial x_k}(t, sx_2 + (1-s)x_1)(x_{2,k} - x_{1,k}). \tag{28}$$

Wenn f_k so definiert wird wie in der Aussage des Theorems, dann ist $F(1) - F(0)$ gleich der rechten Seite von (26). Da $F(1) = f(t, x_2)$ und $F(0) = f(t, x_1)$ ist das Ergebnis damit bewiesen.

Beweis von Satz 3 Zuerst werden die Ableitungen nach x_0 betrachtet. Zu diesem Zweck können wir die Parameterabhängigkeit durch eine Variablentransformation eliminieren. In diesem Fall haben wir eine Lösung $\phi(t_0, t, x_0)$, die durch Satz 1 geliefert wird. Die Aussage ist rein lokaler Natur und wir brauchen sie also nur für eine kleine Umgebung eines Punktes zu beweisen. Sei h eine reelle Zahl und e_k der $k$te Koordinatenbasisvektor im $\mathbb{R}^m$. Sei $x_h(t) = \phi(t_0, t, x_0 + he_k)$. Aus Satz 1 folgt, dass x_h gleichmässig gegen x_0 konvergiert für $h \to 0$. Es gilt

$$\frac{d}{dt}[x_h(t) - x_0(t)] = f(t, x_h(t)) - f(t, x_0(t)). \tag{29}$$

Der Hilfssatz liefert

$$\frac{d}{dt}[x_h(t) - x_0(t)] = f_k(t, x_0(t), x_h(t))(x_{h,k}(t) - x_{0,k}(t)). \tag{30}$$

Sei $y_h = h^{-1}[x_h(t) - x_0(t)]$ für $h \neq 0$. Die Existenz der partiellen Ableitung ist damit äquivalent, dass der Grenzwert von y_h für $h \to 0$ existiert. Die Anfangsbedingung hat zur Folge, dass $x_h(t_0) = x_0 + he_k$ und $y_h(t_0) = e_k$. Die Funktion y_h erfüllt die Gleichung

$$\dot{y}_h = f_k(t, x_0(t), x_h(t))y_{h,k}. \tag{31}$$

Die Größe $f_k(t, x_0(t), x_h(t))$ konvergiert gegen $\frac{\partial f}{\partial y_k}(t, x_0(t))$ für $h \to 0$. Wir haben eine Schar von Gleichungen für die Funktionen y_h, die stetig vom Parameter h abhängt. Diese Gleichungen haben Lösungen mit $y_h(t_0) = e_k$, die stetig von h abhängen, auch bei $h = 0$. Der Grenzwert existiert und ist eine Lösung der Gleichung (21). Die partielle Ableitung ist die Lösung einer Gleichung, die stetig von Parametern abhängt und ist also selbst stetig.

Jetzt wird die Ableitung nach t_0 betrachtet. Sei diesmal

$$y_h(t) = \frac{\phi(t_0 + h, t, x_0) - \phi(t_0, t, x_0)}{h}, \quad h \neq 0. \tag{32}$$

Wir benutzen die Identität

$$\phi(t_0 + h, t, x_0) = \phi(t_0, t, \phi(t_0 + h, t_0, x_0)), \tag{33}$$

Es folgt, dass

$$hy_h(t) = \phi(t_0, t, \phi(t_0 + h, t_0, x_0)) - \phi(t_0, t, x_0). \tag{34}$$

Für $h \to 0$ gilt $\phi(t_0 + h, t_0, x_0) \to x_0$ und der Hilfssatz liefert

$$hy_h(t) = \left[\frac{\partial x}{\partial x_{0,k}} + o(1)\right] (\phi_k(t_0 + h, t_0, x_0) - x_{0,k}). \tag{35}$$

In dem wir die Beziehung $\phi(t_0 + h, t_0 + h, x_0) = x_0$ und den Mittelwertsatz verwenden bekommen wir

$$\phi_k(t_0 + h, t_0, x_0) - x_{0,k} = -\frac{\partial \phi_k}{\partial t}(t_0 + \theta h, t_0 + h, y_0) \tag{36}$$

für ein θ aus $(0, 1)$. Mit der Differentialgleichung kann die zeitliche Ableitung ersetzt werden. Es folgt, dass

$$y_h(t) = -\left[\frac{\partial x}{\partial x_{0,k}} + o(1)\right] [f_k(t_0, x_0) + o(1)]. \tag{37}$$

Wir sehen dass $\frac{\partial \phi}{\partial t_0}$ existiert und die Beziehung erfüllt, die im Satz angegeben wurde.

In dem man die Ergebnisse über die Existenz und Stetigkeit der Ableitungen nach den Anfangsdaten und die Transformation benutzt, die Parameter durch Anfangsdaten ersetzt bekommt man die Aussage über die differenzierbare Abhängigkeit der Lösungen von Parametern. Man bekommt auch die Evolutionsgleichung für die Ableitungen nach den Variablen z_i.

Betrachten wir jetzt eine Gleichung der Form $\dot{x} = f(t, x, z, z^*)$, wo die partielle Ableitungen erster Ordnung bezüglich x und z der stetigen Funktion f existieren und stetig sind. Diese Gleichung hat eine eindeutige Lösung $x = \phi(t_0, t, x_0, z, z^*)$ mit $x(t_0) = x_0$. Die Lösung hat partielle Ableitungen erster Ordnung bezüglich t, t_0, x_0 und z und diese Ableitungen sind stetig als Funktionen von (t_0, t, x_0, z, z^*). Diese Aussagen können so bewiesen werden wie im Beweis von Satz 3, weil die Variablen z^* keine wesentliche Rolle spielen.

Mit diesen Ergebnissen ist es einfach, die Existenz und Stetigkeit von höheren Ableitungen von x zu zeigen wenn die Existenz and Stetigkeit von entsprechenden Ableitungen von f bekannt ist. Diese Aussage kann man durch Induktion beweisen. Um konkret zu sein, betrachten wir eine partielle Ableitung von $\phi(t_0, t, x_0, z, z^*)$ nach den Variablen x_0 und z der Ordnung n. Die Ableitung von ϕ nach x_0 erfüllt die Gleichung (21) und die Koeffizienten dieser Gleichung haben stetige Ableitungen nach x_0 und z der Ordnungen bis $n-1$. Deshalb besitzt die Lösung dieser Gleichung auch solche Ableitungen. Wir können auch entsprechend mit der Ableitung von ϕ nach z verfahren, mit Hilfe der Gleichung (22). Es kann daraus geschlossen werden, dass alle Ableitungen von ϕ nach x und z bis zur Ordnung n existieren und stetig sind. Aussagen über Ableitungen nach t_0 können erhalten werden in dem die bereits bekannten Aussagen in die Gleichung (23) eingesetzt werden.

Die Aussagen über Stetigkeit und Differenzierbarkeit, die jetzt bewiesen wurden können benutzt werden, um etwas über das qualitative Verhalten von Lösungen im allereinfachsten Fall zu beweisen. Es geht um die Natur eines Flusses in der Nähe eines Punktes, wo das Vektorfeld das ihn erzeugt nicht verschwindet.

Satz 4 (Satz über den Flusskasten). Sei $\dot{x} = f(x)$ ein autonomes dynamisches System wo f eine stetig differenzierbare Funktion auf einer Teilmenge G des $\mathbb{R}^m$ ist. Sei $x_0 \in G$ ein Punkt mit $f(x_0) \neq 0$. Dann gibt es eine offene Umgebung V von 0 im $\mathbb{R}^{m-1}$, eine positive Zahl ϵ und einen Diffeomorphismus F von $[-\epsilon, \epsilon] \times V$ auf eine offene Umgebung U von x_0 mit $F(0) = x_0$, so dass der Fluss ϕ des Systems die Beziehung

$$\phi(t, F(y))) = F(T_t(y)) \tag{38}$$

erfüllt. Hier ist T_t die Translation um den Abstand t in die Richtung x_1, d.h. $(x_1, x_2, \ldots, x_m) \mapsto (x_1 + t, x_2, \ldots, x_m)$.

Beweis Wir können ohne Beschränkung der Allgemeinheit annehmen, dass $x_0 = 0$ und $f(x_0) = e_1$. Wir führen $\bar{x}$ ein als Abkürzung für $(x_2, \ldots, x_n)$. Sei V eine offene Umgebung von 0 im $\mathbb{R}^{m-1}$ mit der Eigenschaft, dass $(0, \bar{y}) \in G$ für alle $\bar{y} \in V$. Sei ϵ so klein, dass $\phi(t, (0, \bar{x})) \in G$ für alle $\bar{x} \in V$ und $|t| \leq \epsilon$. Sei $F(y) = \phi(y_1, (0, \bar{y}))$. Die Ableitung von F im Ursprung ist die Identität. Mit dem Satz über die Umkehrabbildung schließen wir dass es eine Umgebung W des Ursprungs gibt so dass die Einschränkung von F auf W ein Diffeomorphismus auf ihr Bild ist. Wir reduzieren jetzt die Größe von V und ϵ wenn nötig, damit $[-\epsilon, \epsilon] \times V \subset W$. Die Abbildung F erfüllt (38) weil beide Seiten dieser Gleichung gleich $\phi(t + y_1, (0, \bar{y}))$ sind.

Dieses Ergebnis sagt, dass man ein beliebiges Vektorfeld in der Nähe eines Punktes wo es nicht verschwindet durch einen Diffeomorphismus zum einfachen Vektorfeld mit Komponenten $(1, 0, \ldots, 0)$ transformieren kann. Intuitiv ausgedrückt, in der Nähe eines Punktes wo es nicht verschwindet hat ein Vektorfeld keine Struktur. Wenn zwei Vektorfelder mit Flüssen ϕ und ψ eine Beziehung der Form $F(\phi(t, x) = \psi(t, F(x))$ für einen C^1-Diffeomorphismus F erfüllen, dann heißen sie C^1-konjugiert. Es folgt aus dem Satz über den Flusskasten, dass wenn $f(x_0)$ und $g(y_0)$ nicht verschwinden die Einschränkungen von f und g auf geeignete Umgebungen von x_0 und y_0 C^1-konjugiert sind. Wenn die Abbildung F nur stetig ist, dann heißen f und g topologisch konjugiert. In diesem Fall werden die Integralkurven von f auf die von g abgebildet und die Zeitrichtung bleibt erhalten. Wenn diese Bedingungen erfüllt sind, aber die Zeitkoordinaten auf den Integralkurven die durch F verbunden werden nicht notwendigerweise gleich sind heißen f und g topologisch äquivalent.

5 Stationäre Lösungen und ihre Stabilität

Wir haben gesehen, dass in der Nähe von Punkten wo ein Vektorfeld nicht verschwindet, der dazugehörige Fluss ein sehr einfaches qualitatives Verhalten hat. Wo das Vektorfeld in einem Punkt verschwindet können die Dinge viel komplizierter werden. Wenn $f(x_0) = 0$ dann ist $x(t) = x_0$, eine zeitunabhängige Lösung, eine stationäre Lösung. Die Gleichung $f(x) = 0$ ist im allgemeinen schwer zu lösen. Es ist schon schwer zu sagen, wie viele Lösungen sie hat. Jetzt wird ein Ergebnis präsentiert, das die Existenz einer Lösung unter schwachen Voraussetzungen garantiert. Zuerst brauchen wir ein Ergebnis aus der Topologie.

Satz (Brouwersche Fixpunktsatz) Sei A ein topologischer Raum der homöomorph ist zu einer abgeschlossenen Kugel im $\mathbb{R}^m$ und sei $\psi : A \to A$ eine stetige Abbildung. Dann existiert ein Punkt $x \in A$ mit $\psi(x) = x$.

Im nächsten Beweis spielen periodische Lösungen eine Rolle. Eine Lösung $x(t)$ heißt periodisch wenn es eine Zahl $T > 0$ gibt mit $x(T) = x(0)$. Durch Eindeutigkeit folgt dann, dass $x(t + T) = x(t)$ für alle t. Die Zahl T heißt Periode.

Satz 5 Sei ein dynamisches System auf einer offenen Teilmenge des $\mathbb{R}^m$ gegeben und sei A eine invariant Teilmenge, die homöomorph ist zu einer abgeschlossenen Kugel im $\mathbb{R}^m$. Dann liegt mindestens eine stationäre Lösung in A.

Beweis Nehmen wir an, es gäbe keine stationären Lösungen in A. Für eine positive Zahl T und $x \in A$ sei $\psi_T(x) = \phi(T, x)$. Da A kompakt ist, ist ψ_T wohldefiniert. Diese Abbildung ist stetig und bildet A in sich ab. Nach dem Brouwerschen Fixpunktsatz existiert ein Punkt z_T mit $\psi_T(z_T) = z_T$. Die Lösung mit Anfangswert z_T ist periodisch mit Periode T. In dem wir verschiedene Werte von T wählen können wir periodische Lösungen bekommen mit Perioden $1/n$ durch Punkte $z_{1/n}$ für alle natürlichen Zahlen n. Weil A kompakt ist hat diese Folge eine konvergente Teilfolge. Nennen wir sie y_n und ihr Grenzwert y. Es gibt also periodische Lösungen, die beliebig nahe bei y anfangen und eine beliebig kurze Zeit brauchen, um zu ihrem Anfangspunkt zurückzukommen. Durch unsere Annahme sind sie nicht stationär. Sei K ein Flusskasten für y. Auf A ist f durch eine Konstante M beschränkt. Eine periodische Lösung mit Periode T kann niemals Punkte erreichen die von ihren Anfangspunkten weiter weg sind als MT. Für ϵ klein genug liegt die offene Kugel vom Radius 2ϵ um y in K. Wenn eine Lösung in der offenen Kugel mit Radius ϵ um y startet und ihre Periode nicht größer ist als ϵ/M dann kann sie den Flusskasten nie verlassen. Aber im Flusskasten gibt es keine periodischen

Lösungen, ein Widerspruch.

Man kann diesen Satz auf das fundamentale System der Virusdynamik anwenden, wobei A das invariante Gebiet für dieses System ist, dass wir schon gefunden haben. Daraus folgt, dass das System für beliebige positive Werte der Parameter mindestens eine nichtnegative stationäre Lösung besitzt. Für dieses System sind es die positiven Lösungen, die am interessantesten sind. Trotzdem haben andere Lösungen die nicht negativ sind ein gewisses Interesse. Man kann sie unter Umständen benutzen, um etwas über das asymptotische Verhalten von positiven Lösungen herauszufinden. Sie können auch interessante Grenzfälle des ursprünglichen Systems sein. Zum Beispiel entspricht eine Lösung des fundamentalen Systems der Virusdynamik mit $y = 0$ und $v = 0$ dem Zustand einer gesunden Person (keine freien Virus-Teichen, keine infizierte Zellen).

Das fundamentale System der Virusdynamik ist einfach genug, dass man die stationären Lösungen explizit ausrechnen kann. Die Koordinaten einer stationären dieses Systems werden mit einem Stern versehen. Die dritte Gleichung liefert $y^* = \frac{u}{k}v^*$. Wenn man diese Beziehung in die zweite Gleichung einsetzt haben beide Seiten einen Faktor v^*. Für eine positive Lösung können wir diesen gemeinsamen Faktor kürzen, mit dem Ergebnis $x^* = \frac{au}{\beta k}$. Wenn wir diese Beziehung in die erste Gleichung einsetzen, bekommen wir $v^* = \frac{d}{\beta}\left(\frac{\beta k\lambda}{adu} - 1\right)$. Sei $R_0 = \frac{\beta k\lambda}{adu}$. Dieses Objekt heißt bei den Biologen fundamentales Vermehrungsverhältnis. Wir sehen, dass es eine positive stationäre Lösung nur dann geben kann, wenn $R_0 > 1$. Wenn man die Gleichung für x^* in die zweite Gleichung für stationäre Lösungen einsetzt sieht man, dass $y^* = \frac{u}{k}v$ und dass $y^* = \frac{du}{\beta k}(R_0 - 1)$. Jetzt haben wir zwei Dinge bewiesen. Wenn $R_0 > 1$ gibt es genau eine positive stationäre Lösung, die wir explizit berechnet haben. Wenn $R_0 \leq 1$ gibt es keine positive stationäre Lösung. Wenn wir nichtnegative Lösungen zulassen, dann ist $v^* = 0$ auch eine Möglichkeit. Dann ist auch $y^* = 0$. Die verbleibende Gleichung sagt dass in diesem Fall $x^* = \frac{\lambda}{d}$.

Die Bedeutung von stationären Lösungen hängt von ihrer Stabilität ab. Eine stationäre Lösung x^* heißt stabil wenn es zu jeder Umgebung U von x^* eine Umgebung V von x^* gibt, so dass für jede Lösung $x(t)$ die Bedingung $x(t_0) \in V$ impliziert $x(t) \in U$ für alle $t \geq t_0$. In Worten, jede Lösung, die in V startet bleibt in U so lange sie existiert. Die stationäre Lösung x^* heißt asymptotisch stabil wenn sie stabil ist und es eine Umgebung U von x^* gibt so dass $x(t_0) \in U$ impliziert $x(t) \to x^*$ für $t \to \infty$. Die erste Bedingung in dieser Definition folgt nicht aus der zweiten, wie das folgende Beispiel zeigt (cf. [9]).

$$\dot{x} = x - rx - ry + xy, \tag{39}$$

$$\dot{y} = y - ry + rx - x^2. \tag{40}$$

Dieses System ist C^1. Das qualitative Verhalten ist einfacher zu sehen in Polarkoordinaten, wo das System die Form

$$\dot{r} = r(1 - r), \tag{41}$$

$$\dot{\theta} = r(1 - \cos\theta) \tag{42}$$

hat. Es gibt stationäre Punkte bei $(0,0)$ und $(1,0)$. Alle Lösungen außer der stationären Lösung im Ursprung konvergieren gegen $(0,1)$ für $t \to \infty$ aber dieser Punkt ist nicht stabil. Diese Aussagen werden hier nicht bewiesen, da die dazu notwendigen Techniken noch nicht eingeführt wurden.

Eine Art, die Stabilität einer stationären Lösung x^* zu untersuchen ist das System um x^* zu linearisieren. Wir betrachten die Taylor-Entwicklung von f um x^*, $f_i(x) = \frac{\partial f_i}{\partial x_j}(x^*)(x_j - x_j^*) + o(|x - x^*|)$. Wir bezeichnen die Matrix $\frac{\partial f}{\partial x}(x^*)$ mit A. Die linearisierte Gleichung um x^* bekommt man in dem man den Restterm in der Taylor-Entwicklung, der klein sein soll, weglässt. Damit bekommt man die Gleichung $\frac{d\hat{x}}{dt} = A\hat{x}$. Die Hoffnung ist, dass unter geeigneten Bedingungen Lösungen des linearisierten Systems Lösungen des ursprünglichen Systems in der Nähe von x^* approximieren. Wenn wir das qualitative Verhalten der Lösungen in der Nähe eines Punktes betrachten unterscheiden wir nicht zwischen Vektorfeldern, die topologisch konjugiert sind durch eine Abbildung die x^* festhält. Deshalb interessieren wir uns in diesem Zusammenhang nur für Eigenschaften von A die unter Ähnlichkeitstransformationen invariant sind. Da jede Matrix mit ihrer Jordanschen Normalform ähnlich ist kann es sich dabei nur um die Eigenschaften der Normalform handeln. Für eine detaillierte Diskussion von linearen Differentialgleichungen wird der Leser auf das erste Kapitel des Buches von Perko [9] verwiesen. Die Reduktion auf die Normalform ist in allgemeinen nur mit Hilfe der komplexen Zahlen möglich, weil die Eigenwerte komplex sein könnnen. Deshalb werden wir hier gelegentlich gezwungen, komplexe lineare gewöhnliche Differentialgleichungen zu betrachten, obwohl wir uns am Ende nur für reelle Lösungen von Gleichungen mit reellen Koeffizienten interessieren.

Für eine komplexe Matrix A mit einem Eigenwert λ heißen die Vektoren x, die $(A - \lambda I)^k x = 0$ für eine natürliche Zahl k erfüllen die entsprechenden verallgemeinerten Eigenvektoren. Sie bilden einen Vektorraum V_λ. Wenn eine Matrix in Normalform ist, gehören die nichtverschwindenden Einträge einer Folge von Blöcken auf der Diagonale, den Jordan-Blöcken. In jedem Block sind die diagonalen Elemente gleich einer Zahl λ, die ein Eigenwert der Matrix ist. Die Einträge unmittelbar oberhalb dieser diagonalen Einträge sind gleich Eins und alle anderen Einträge verschwinden. Nehmen wir an, dass die Blöcke die Größen n_i haben. Die Vektoren, wo nur die ersten n_1 Komponenten von Null verschieden sind, sind verallgemeinerte Eigenvektoren, die zum ersten Eigenwert λ_1 gehören. Die Vektoren, wo nur die Komponenten von $n_1 + 1$ bis $n_1 + n_2$ von Null verschieden sind gehören zum zweiten Eigenwert λ_2 und so weiter. Wenn A eine allgemeine reelle Matrix ist, und λ ein reeller Eigenwert von A, dann wird der Raum V_λ genauso wie im komplexen Fall definiert. Im Fall, dass λ ein komplexer Eigenwert ist, ist die Definition ein wenig komplizierter. Dann ist V_λ die Menge der Realteile der komplexen Lösungen von $(A - \lambda I)^k x = 0$. Weil A reell ist, ist $\bar{\lambda}$ auch ein Eigenwert und $V_{\bar{\lambda}} = V_\lambda$. Der ganze Raum ist die direkte Summe der unterschiedlichen verallgemeinerten Eigenräume.

Wenn wir die Gleichung $\dot{x} = Ax$ lösen wollen, dann können wir A in Jordan-Form bringen, die Gleichung lösen und die Lösung dann zurücktransformieren.

Die Teilräume von verallgemeinerten Eigenvektoren, die zu den Eigenwerten gehören sind invariant unter dem Fluss des linearisierten Systems. Deshalb reicht es, den Fall zu betrachten, wo es nur einen Jordan-Block gibt. Wenn dieser Block von der Größe Eins ist, mit Eigenvektor λ, dann ist die Lösung von der Form $ce^{\lambda t}$ für eine Konstante c. Im allgemeinen ist die Lösung das Produkt der Funktion $e^{\lambda t}$ mit einer Matrix, deren Elemente jeweils eine Konstante mal eine Potenz von t sind. Es handelt sich im allgemeinen um komplexe Exponentialfunktionen. Wenn wir Real- und Imaginärteile bilden, um die Lösungen der reellen Gleichung zu bekommen, dann ist jeder Eintrag eine Linearkombination von Ausdrücken der Form t^k, $t^k e^{at}\cos bt$ oder $t^k e^{at}\sin bt$, wo $\lambda = a + bi$. Wenn λ reell und positiv ist, dann wächst die Lösung, wenn sie nicht identisch Null ist, mindestens so schnell wie $e^{\lambda t}$ wenn t zunimmt. Wenn λ komplex ist, und $a > 0$, dann wächst die Lösung längs geeigneten Folgen, die nach $+\infty$ streben, mindestens so schnell wie e^{at}. Wenn λ reell und negativ ist, dann klingt die Lösung mindestens so schnell ab wenn t zunimmt wie $e^{(\lambda+\epsilon)t}$, wo $\epsilon > 0$ beliebig ist. Wenn λ komplex ist, und $a < 0$, dann klingt sie mindestens so schnell wie $e^{(a+\epsilon)t}$ ab. Es können natürlich ähnliche Aussagen für die andere Zeitrichtung gemacht werden.

Bei der Untersuchung der Eigenschaften von linearen gewöhnlichen Differentialgleichungen ist der Begriff der Exponentialfunktion einer Matrix sehr nützlich. Man definiert

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{43}$$

Für eine beliebige komplexe Matrix A konvergiert diese Reihe gleichmässig auf jeder kompakten Teilmenge in dem Sinne, dass die Elemente der entsprechenden Matrizen es tun. Die Relevanz dieser Definition zu den gewöhnlichen Differentialgleichungen ist dass $e^{tA}x_0$ die Gleichung $\dot{x} = Ax$ löst mit $x(0) = x_0$. Wenn die Matrix A in Jordan-Form ist, dann ist e^{tA} die direkte Summe der Ausdrücke für die einzelnen Jordan-Blöcke. Dadurch bekommt man Abschätzungen für e^{tA} die den Abschätzungen für Lösungen linearer Gleichungen, die oben besprochen wurden entsprechen.

Wir sehen, dass für lineare Systeme Eigenwerte mit positivem Realteil mit Instabilität zu tun haben und Eigenwerte mit negativem Realteil mit Stabilität. Diese Beobachtung motiviert die folgenden Definitionen. Der Raum V_+, der durch alle verallgemeinerten Eigenvektoren aufgespannt wird, die zu Eigenwerten mit positivem Realteil gehören heißt instabiler Teilraum. Der Raum V_-, der durch alle verallgemeinerten Eigenvektoren aufgespannt wird, die zu Eigenwerten mit negativem Realteil gehören heißt stabiler Teilraum. Der Raum V_c der durch alle verallgemeinerten Eigenvektoren aufgespannt wird, die zu Eigenwerten mit verschwindendem Realteil gehören heißt Zentrumsteilraum. Der ganze Raum ist die direkte Summe $V_- \oplus V_c \oplus V_+$. Diese Teiläume sind invariant unter dem Fluss. Für ein lineares System gelten auf Grund der oben angegebenen Tatsachen, folgende Aussagen. Wenn alle Eigenwerte negativen Realteil haben dann ist der Ursprung asymptotisch stabil. Wenn mindestens ein Eigenwert einen positiven Realteil hat, dann ist der Ursprung nicht stabil.

Im folgenden werden wir analoge Aussagen für eine stationäre Lösung eines allgemeinen nichtlinearen Systems beweisen. Dazu müssen wir aber einige weitere Ideen einführen.

Bevor wir das tun befassen wir uns mit der Linearisierung des fundamentalen Modells der Virusdynamik um die zwei stationären Lösungen. Die Linearisierung um einen beliebigen Punkt ist

$$\frac{d\hat{x}}{dt} = (-d - \beta v)\hat{x} - \beta x \hat{v}, \tag{44}$$

$$\frac{d\hat{y}}{dt} = \beta v \hat{x} - a\hat{y} + \beta x \hat{v}, \tag{45}$$

$$\frac{d\hat{v}}{dt} = k\hat{y} - u\hat{v}. \tag{46}$$

Im Fall des stationären Punktes mit $v = 0$ vereinfacht sich dieser Ausdruck wesentlich und man sieht sofort, dass $-d$ ein Eigenwert ist. Die anderen beiden können dann bestimmt werden, in dem man eine quadratische Gleichung löst. Das Ergebnis ist

$$\mu = \frac{1}{2}[-(a+u) \pm \sqrt{(a+u)^2 - 4au(1-R_0)}]. \tag{47}$$

Wir sehen, dass es immer mindestens zwei negative Eigenwerte gibt und dass der dritte positiv bzw. Null bzw. negativ ist, wenn $R_0 > 1$ bzw. $R_0 = 1$ bzw. $R_0 < 1$. Nach den Stabilitätskriterien, die wir noch nicht bewiesen haben, ist dieser Punkt asymptotisch stabil für $R_0 < 1$ und nicht stabil für $R_0 > 1$. Intuitiv haben diese Aussagen folgende Bedeutung. Der Zustand einer gesunden Person wird durch diese stationäre Lösung dargestellt. Wenn eine Infektion stattfindet, wird dieser Zustand ein wenig gestört. Für $R_0 < 1$ wird die Infektion automatisch eliminiert. Für $R_0 > 1$ kann das Virus im Körper Fuss fassen. Die Linearisierung im anderen stationären Punkt ist

$$\frac{d\hat{x}}{dt} = -dR_0\hat{x} - \frac{au}{k}\hat{v}, \tag{48}$$

$$\frac{d\hat{y}}{dt} = d(R_0 - 1)\hat{x} - a\hat{y} + \frac{au}{k}\hat{v}, \tag{49}$$

$$\frac{d\hat{v}}{dt} = k\hat{y} - u\hat{v}. \tag{50}$$

Daraus folgt die Eigenwertgleichung

$$\mu^3 + (a + u + dR_0)\mu^2 + dR_0(a+u)\mu + adu(R_0 - 1) = 0. \tag{51}$$

Es ist nur der Fall $R_0 > 1$ von Interesse, da die stationäre Lösung nur dann positiv ist. In diesem Fall sind alle Koeffizienten im Polynom positiv. Um Informationen über die Eigenwerte zu bekommen benutzen wir das Kriterium von Routh und Hurwitz. Für eine Gleichung dritten Grades

$$\mu^3 + a_1\mu^2 + a_2\mu + a_3 = 0. \tag{52}$$

sagt dieses Kriterium, dass alle Lösungen negativen Realteil haben wenn und nur wenn $a_1 > 0$, $a_3 > 0$ und $a_1 a_2 - a_3 > 0$. In unserem Beispiel haben also alle Eigenwerte negative Realteile falls

$$(a + u + dR_0)dR_0(a + u) > adu(R_0 - 1). \tag{53}$$

Wenn wir diese Ungleichung ausmultiplizieren und die Terme nach Potenzen von R_0 ordnen, dann bekommen wir

$$d^2(a + u)R_0^2 + R_0[d(a^2 + ad + u^2)] + adu > 0, \tag{54}$$

eine Bedingung, die offenbar gilt. Wenn $R_0 = 1$ sind die Eigenwerte 0,$-(a + u)$ and $-d$.

Um Aussagen über Stabilität zu beweisen, benutzen wir Ideen, die auf Ljapunow zurückgehen. Sei $\dot{x} = f(x)$ ein autonomes dynamisches System. Wenn V eine stetig differenzierbare Funktion ist, sei $\dot{V} = \frac{\partial V}{\partial x_i} f_i(x)$. Durch die Kettenregel ist $\dot{V} = \frac{d}{dt}(V(x(t))$. Eine Funktion, die $\dot{V} \leq 0$ erfüllt heißt Ljapunow-Funktion.

Satz 6 Sei G eine offene Umgebung eines Punktes x_0. Sei f ein Vektorfeld der Klasse C^1 mit $f(x_0) = 0$. Sei V eine stetig differenzierbare Funktion mit $V(x_0) = 0$ und $V(x) > 0$ für $x \neq x_0$. Wenn $\dot{V}(x) \leq 0$ für alle $x \in G$ dann ist x_0 stabil. Wenn $\dot{V}(x) < 0$ für alle $x \in G$ außer x_0 dann ist x_0 asymptotisch stabil. Wenn $\dot{V}(x) > 0$ für alle $x \in G$ außer x_0 dann ist x_0 instabil.

Beweis Wir können o.B.d.A. annehmen, dass $x_0 = 0$. Sei $\epsilon > 0$ klein genug, dass $\bar{B}_\epsilon(0) \subset G$ und sei m_ϵ das Minimum der stetigen Funktion V auf der Sphäre S_ϵ vom Radius ϵ um den Ursprung. Dann ist $m_\epsilon > 0$. Da V stetig ist und $V(0) = 0$ existiert $\delta > 0$ so dass $V(x) < m_\epsilon$ für $|x| < \delta$. Weil $\dot{V} \leq 0$ kann V längs der Integralkurven des Vektorfeldes nicht zunehmen. Deshalb erfüllt der Fluss ϕ von f die Beziehung

$$V(\phi(t, x_0)) \leq V(x_0) < m_\epsilon \tag{55}$$

für alle $x_0 \in B_\delta(0)$. Nehmen wir an, dass für $|x_0| < \delta$ es existiert t_1 mit $\phi(t_1, x_0) \in S_\epsilon$. In diesem Fall wäre $V(\phi(t_1, x_0)) \geq m_\epsilon$, ein Widerspruch. Deshalb impliziert $|x_0| < \delta$ dass $|\phi(t, x_0)| < \epsilon$ für $t \geq 0$.

Wir betrachten jetzt den Fall, dass $\dot{V}(x) < 0$ für alle $x \in G$ außer x_0. Dann nimmt V strikt ab entlang der Integralkurven on f. Sei $x_0 \in B_\delta(0)$, wo δ wie soeben definiert wird. Dann ist $\phi(t, x_0) \in B_\epsilon(0)$ für alle $t \geq 0$. Sei $\{t_k\}$ eine Folge mit $t_k \to \infty$. Da $\bar{B}_\epsilon(0)$ kompakt ist existiert eine Teilfolge, so dass $\phi(t_k, x_0)$ gegen einen Punkt y_0 von $\bar{B}_\epsilon(0)$ konvergiert. Wir werden zeigen, dass für jede solche Folge der Grenzwert Null sein muss. Es folgt dann, dass $\phi(t_k, x_0)$ entlang jeder Teilfolge gegen den Ursprung konvergiert. Damit konvergiert $\phi(t, x)$ gegen den Ursprung . Es bleibt zu zeigen, dass wenn $\phi(t, x_0) \to y_0$ dann ist $y_0 = 0$. V nimmt entlang einer Integralkurve echt ab und erfüllt $V(\phi(t, x_0) \to V(y_0)$. Deshalb ist $V(\phi(t, x_0) > V(y_0)$ für alle $t > 0$. Wenn

$y_0 \neq 0$ dann ist $V(\phi(s, y_0)) < V(y_0)$ für $s > 0$. Durch Stetigkeit ist dann auch $V(\phi(s, y)) < V(y_0)$ für y nahe genug bei y_0. Aber dann ist für n groß genug $V(\phi(s + t_n, x_0)) < V(y_0)$, ein Widerspruch.

Zum Schluss betrachten wir den Fall, dass $\dot{V}(x) > 0$ für alle $x \in G$ außer x_0. Sei M das Maximum von V auf der Menge $\bar{B}_\epsilon(0)$. In diesem Fall nimmt V entlang der Integralkurven echt zu. Für $\delta > 0$ beliebig und $x_0 \neq 0$ in $B_\delta(0)$ gilt $V(\phi(t, x_0)) > V(x_0) > 0$ für alle $t > 0$. Die Menge wo $V(x) < V(x_0)$ ist offen und die Teilmenge von $\bar{B}_\epsilon(0)$ wo $V(x) \geq V(x_0)$ ist kompakt. Dort ist $\dot{V}$ positiv und hat ein positives Minimum m. Wir haben $\inf_{t \geq 0} \dot{V}(\phi(t, x_0)) \geq m > 0$. Deshalb ist $V(\phi(t, x_0) \geq V(x_0) + mt > M$ für t hinreichend groß. Damit ist die Instabilität bewiesen.

Die Aussagen über die Stabilität oder Instabilität von stationären Lösungen werden mit Hilfe von Ljapunow-Funktionen bewiesen, die zu diesem Zweck konstruiert werden. Hier folgen wir der Behandlung dieses Themas in [3].

Hilfssatz 3 Sei A eine reelle $n \times n$-Matrix. Die Matrixgleichung $A^T B + BA = -C$ hat eine Lösung für jede positiv definite Matrix C wenn und nur wenn alle Eigenwerte von A negativen Realteil haben.

Beweis Betrachten wir die lineare Gleichung $\dot{x} = Ax$ und die reellwertige Funktion $V(x) = x^T Bx$, wo B eine symmetrische Matrix ist. Dann ist

$$\dot{V}(x) = x^T (A^T B + BA)x. \tag{56}$$

Wenn die Gleichung die in diesem Hilfssatz betrachtet wird gilt, dann ist $\dot{V}(x) < 0$ für $x \neq 0$ und die Lösung der Differentialgleichung konvergiert gegen den Ursprung für $t \to \infty$. Es folgt, dass die Eigenwerte von A negativen Realteil haben. Wenn umgekehrt diese Eigenwerte negativen Realteil haben und C eine positiv definite Matrix ist, dann sei

$$B = \int_0^\infty e^{A^T t} C e^{At} dt. \tag{57}$$

Dieses Integral ist wohldefiniert, da es positive Konstanten K und α gibt mit $\|e^{At}\| \leq Ke^{-\alpha t}$ für $t \geq 0$. Außerdem ist B positiv definit und

$$A^T B + BA = \int_0^\infty \frac{d}{dt}(e^{A^T t} C e^{At}) dt = -C. \tag{58}$$

Es folgt dass wenn der Ursprung für das lineare System asymptotisch stabil ist, dann gibt es eine quadratische Form, die längs der Lösungen dieser Gleichung strikt abnimmt. Jetzt betrachten wir die nichtlineare Gleichung $\dot{x} = Ax + g(x)$, wo g stetig differenzierbar ist und $g(0) = 0$ und $\frac{\partial g_i}{\partial x_j}(0) = 0$ erfüllt. Wenn die Realteile der Eigenwerte von A negativ sind können wir die Matrix B betrachten im Fall $C = I$. Dann ist

$$\dot{V} = -|x|^2 + g^T Bx + x^T Bg = -|x|^2(1 + o(1)). \tag{59}$$

Es folgt, dass der Ursprung asymptotisch stabil ist. Jetzt soll ein Satz über Instabilität für das nichtlineare System bewiesen werden. Dazu ist es nützlich zu bemerken, dass der dritte Teil von Satz 6 verallgemeinert werden kann. Es wird angenommen, das der stationäre Punkt, der untersucht werden soll der Ursprung ist. Wir führen eine offene Teilmenge U ein mit der Eigenschaft, dass der Urspung in der abgeschlossenen Hülle von U liegt. Sei $H = U \cap B_\epsilon(0)$. Wir nehmen an, dass die stetig differenzierbare Funktion V folgende Eigenschaften hat. $V(x) = 0$ auf dem Teile Randes von H, der in $B_\epsilon(0)$ liegt und $V(x) > 0$ in allen anderen Punkten von H. $\dot{V}(x) > 0$ auf H außer im Ursprung. Die Behauptung ist jetzt, dass der Ursprung unter diesen Umständen instabil ist. Wie im Beweis von Theorem 6 sei $\delta > 0$ beliebig. Sei $x_0 \neq 0$ ein belibiger Punkt von $B_\delta(0) \cap H$. Dann ist $V(\phi(t, x_0)) > V(x_0) > 0$ für alle $t > 0$. Deshalb kann die Lösung die Menge H nur durch den Rand von $B_\delta(0)$ verlassen. Die Teilmenge von $\bar{H}$ wo $V(x) \geq V(x_0)$ ist kompakt und $\dot{V}$ hat eine Minimum dort. Damit kann man wie im Beweis von Satz 6 argumentieren, dass die Lösung den Rand von $B_\delta(0)$ nach endlicher Zeit erreicht, und die Instabilität ist bewiesen.

Jetzt soll dieses Ergebnis auf den Fall angewendet werden, wo eine stationäre Lösung einen Eigenwert mit positivem Realteil besitzt. Wir nehmen an, dass die stationäre Lösung im Ursprung ist und dass die Teilräume V_+, V_c und V_- Koordinatenebenen sind. Ein allgemeiner Punkt kann als (x, y, z) dargestellt werden, wobei x, y bzw. z den Teilräumen V_+, V_c bzw. V_- gehören. In der gegebenen Situation ist die Dimension von V_+ positiv. Die Gleichungen sind

$$\frac{dx}{dt} = A_+ x + f(x, y, z), \tag{60}$$

$$\frac{dy}{dt} = A_c y + g(x, y, z), \tag{61}$$

$$\frac{dz}{dt} = A_- z + h(x, y, z) \tag{62}$$

wobei die Eigenwerte von A_+ und $-A_-$ positiven Realteil haben und die Funktionen f, g und h alle $o(\sqrt{|x|^2 + |y|^2 + |z|^2})$ sind in der Nähe des Ursprungs. Durch die Einführung einer neuen Basis im Raum V_c können wir dafür sorgen, dass der symmetrische Anteil von A kleiner ist als jede gegebene positive Zahl ϵ. Um diese Aussage zu beweisen können wir die Matrix zuerst in Jordanscher Normalform schreiben. Anschließend können wir die Jordan-Blöcke einzeln behandeln. Betrachten wir zuerst einen Block mit Eigenwert Null. Sei $e_1, \ldots, e_k$ die Basis die diesen Block reduziert. Eine neue Basis kann durch $\tilde{e}_i = \epsilon^i e_i$ definiert werden und die transformierte Matrix hat die gewünschte Eigenschaft. Bei einem imaginären Eigenwert, der nicht Null ist, können wir zuerst eine Basis $e_1, \ldots, e_k$ wählen wie im vorherigen Fall, wobei die Vektoren diesmal komplex sind. Die Real- und Imaginärteile dieser Vektoren bilden eine Basis in der die einzigen Elemente, die nicht mit ϵ multipliziert sind schiefsymmetrisch sind. Wir haben wieder die gewünschte Eigenschaft. Es folgt, dass $x^T A_c x = \frac{1}{2} x^T (A_c^T + A) x = O(\epsilon x^T x)$. Nach dem Hilfssatz gibt es positiv definite Matrizen B_+ und B_-, so dass $A_+^T B_+ + B_+ A_+ = I$ und $A_-^T B_- + B_- A_- = -I$.

Als Funktion V nehmen wir $x^T B_+ x - y^T y - z^T B_- z$. Dann ist

$$\dot{V} = x^T x + z^T z - y^T (A_c^T + A) y + o(x^T x + y^T y + z^T z). \tag{63}$$

Die Menge U wird durch die Bedingung $V > 0$ definiert. Dort ist $y^T y \leq C x^T x$ für eine positive Konstante C und deshalb ist auf diesem Gebiet $\dot{V} = -\frac{1}{2}(x^T x + z^T z) + o(|x|^2 + |z|^2)$. Es folgt, dass für ϵ klein genug alle Bedingungen erfüllt sind und der Ursprung ist instabil.

Es folgt aus den soeben bewiesenen Aussagen, dass für $R_0 > 1$ die stationaräre Lösung des fundamentalen Modells der Virusdynamik mit $v > 0$ asymptotisch stabil ist und die Lösung mit $v = 0$ instabil. Dagegen ist für $R_0 < 1$ die Lösung mit $v = 0$, die in diesem Fall die einzige nichtnegative stationäre Lösung ist, stabil. Diese Aussagen betreffen das Verhalten der Lösungen in einer Umgebung der stationären Lösungen. Eine Ljapunow-Funktion kann auch helfen, globale Aussagen zu beweisen. Wenn ein dynamisches System auf einer offenen Teilmenge U definiert ist, und V die Ungleichung $\dot{V} \leq 0$ erfüllt, sind die ω-Limespunkte einer Lösung $x(t)$ in U die in U liegen Punkte wo $\dot{V} = 0$. Die Funktion $V(x(t))$ ist monoton fallend und nichtnegativ. Sie konvergiert also gegen eine Konstante V_∞. Es folgt, dass $V(y) = V_\infty$ für jeden ω-Limespunkt y von $x(t)$. Deshalb ist V auf der ω-Limesmenge konstant. Wenn y ein ω-Limespunkt ist, dann liegt die Lösung mit Anfangswert y in der ω-Limesmenge. Die Funktion V ist längs dieser Lösung konstant und deshalb ist $\dot{V}(y) = 0$. Aus dieser Überlegung folgt, dass wenn $\dot{V} < 0$ auf U dann gibt es keine ω-Limespunkte in U.

Korobeinikov [5] hat Ljapunow-Funktionen benutzt, um das globale qualitative Verhalten der Lösungen des fundamentalen Systems der Virusdynamik zu bestimmen. Er bezeichnet die stationäre Lösung mit $v > 0$ mit (x^*, y^*, v^*) und die stationäre Lösung mit $v = 0$ mit $(x_0, 0, 0)$. Betrachten wir zuerst die Funktion

$$V(x, y, v) = x^* \left(\frac{x}{x^*} - \log \frac{x}{x^*} \right) + y^* \left(\frac{y}{y^*} - \log \frac{y}{y^*} \right) + \frac{a}{k} v^* \left(\frac{v}{v^*} - \log \frac{v}{v^*} \right) \tag{64}$$

Diese Funktion hat ein Minimum im Punkt (x^*, y^*, v^*). Die Ableitung ist

$$\begin{aligned}
\dot{V} &= \left(1 - \frac{x^*}{x}\right) \dot{x} + \left(1 - \frac{y^*}{y}\right) \dot{y} + \frac{a}{k} \left(1 - \frac{v^*}{v}\right) \dot{v} && (65) \\
&= \lambda - dx - \frac{au}{k} v - \lambda \frac{x^*}{x} + \beta x^* v + dx^* \\
&\quad - \beta x v \frac{y^*}{y} + a y^* - a y \frac{v^*}{v} + \frac{au}{k} v^* \\
&= \lambda + dx^* + ay^* + \frac{au}{k} v^* - dx + \left(\beta x^* - \frac{au}{k}\right) v \\
&\quad - \lambda \frac{x^*}{x} - \beta x v \frac{y^*}{y} - a y \frac{v^*}{v} \\
&= dx^* \left(2 - \frac{x}{x^*} - \frac{x^*}{x}\right) + a y^* \left(3 - \frac{x^*}{x} - \frac{x v y^*}{x^* v^* y} - \frac{y v^*}{y^* v}\right). && (66)
\end{aligned}$$

Es kann gezeigt werden, dass $\dot{V} \leq 0$, in dem man die Ungleichung zwischen den arithmetischen und geometrischen Mitteln verwendet. Diese sagt, dass, für positive Zahlen $a_1, \ldots, a_n$, $(\prod_{i=1}^n a_i)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n a_i$, und dass die Gleichheit nur gilt, wenn alle a_i gleich sind. Es folgt, dass $\dot{V}$ nur Null sein kann, wenn $x = x^*$. Wenn es einen positiven ω-Limespunkt gibt muss $\dot{V} = 0$ dort. Deshalb ist $x = x^*$ auf der ganzen Lösung, die an diesem Punkt startet und $\dot{x} = 0$. Wenn diese Information in die Gleichung für x eingesetzt wird, sieht man, dass v konstant ist. Die Gleichung für v zeigt dann, dass y konstant ist. Wir sehen, dass jeder positive ω-Limespunkt eine stationäre Lösung ist. Es kann nur der Punkt (x^*, y^*, v^*) sein. Andererseits sind ω-Limespunkte, wo eine der Variablen verschwindet unmöglich weil V bei Annäherung eines solchen Punktes gegen $+\infty$ strebt. Deshalb kann es keine ω-Limespunkte außer der stationären Lösung (x^*, y^*, v^*) geben. Damit ist gezeigt, dass im Fall $R_0 > 1$ jede positive Lösung gegen die positive stationäre Lösung konvergiert für $t \to \infty$.

Um den Fall $R_0 \leq 1$ zu verstehen betrachten wir die Funktion

$$U(x, y, v) = x_0\left(\frac{x}{x_0} - \log\frac{x}{x_0}\right) + y + \frac{a}{k}v. \tag{67}$$

Im Bereich wo alle Variablen nicht negativ sind und x positiv hat diese Funktion ein Minimum im Punkt $(x_0, 0, 0)$. Die Ableitung ist

$$\begin{aligned}\dot{U} &= \left(1 - \frac{x_0}{x}\right)\dot{x} + \dot{y} + \frac{a}{k}\dot{v} \\ &= \lambda\left(2 - \frac{x}{x_0} - \frac{x_0}{x}\right) + \frac{au}{k}(R_0 - 1)v. \end{aligned} \tag{68}$$

Diese Größe ist nicht negativ und verschwindet nur wenn $R_0 = 1$ und $x = x_0$. Ein ω-Limespunkt einer positiven Lösung kann nicht $x = 0$ erfüllen, da U gegen unendlich geht für $x \to 0$. Wie im Fall $R_0 > 1$ kann man dann argumentieren, dass ein ω-Limespunkt mit $x > 0$ eine stationäre Lösung sein muss. Deshalb konvergiert jede Lösung gegen die eindeutige stationäre Lösung für $t \to \infty$.

Dieses Beispiel zeigt, wie eine Ljapunow-Funktion helfen kann, das asymptotische Verhalten eines dynamischen Systems zu untersuchen. Leider gibt es keine allgemeine Methode, um Ljapunow-Funktionen zu finden. Dabei handelt es sich eher um eine Kunst als eine Wissenschaft. Wie hat Korobeinikov seine Ljapunow-Funktion gefunden? In seiner Arbeit sagt es nicht sehr viel dazu, aber er erwähnt eine Beziehung zu Modellen aus der Epidemiologie. Dieser Spur wollen wir ein wenig nachgehen. Dabei können wir auch mit einigen wichtigen Modellen aus der mathematischen Biologie Bekanntschaft machen. Die Modelle um die es sich handelt wurden von Kermack und McKendrick im Jahr 1927 eingeführt. Betrachten wir eine Population von Menschen (oder Tieren) die einer Infektionskrankrankheit ausgesetzt ist. Sei S der Anteil der Menschen die anfällig sind für die Krankheit (englisch 'susceptible'), I der Anteil die infiziert sind (englisch 'infected' oder 'infectious') und R der Anteil die genesen oder entfernt sind (english 'recovered' oder 'removed'). Dann ist $S + I + R = 1$. Nehmen wir an, dass die Population konstant ist, so dass S, I und R proportional der Zahlen in den verschiedenen Gruppen sind. Im einfachsten Modell sind die Gleichungen

$$\dot{S} = -\beta SI, \tag{69}$$
$$\dot{I} = \beta SI - \alpha I, \tag{70}$$
$$\dot{R} = \alpha I. \tag{71}$$

Dieses Modell ist als SIR-Modell bekannt. Man sieht sofort, dass $S + I + R$ konstant ist. Da wir R aus den anderen Variablen berechnen können, können wir die dritte Gleichung weglassen. In diesem Modell wird angenommen, dass jemand der infiziert ist sofort andere infizieren kann, was nicht sehr realistisch ist für viele Krankheiten. Später werden wir eine andere Alternative kennenlernen. Der Übergang von I zu R kann durch Genesung mit anschließender Immunität passieren, durch räumliche Entfernung (Quarantäne während einer Epidemie) oder durch den Tod. In diesem Modell werden Geburten nicht berücksichtigt und Todesfälle, die nicht durch die Krankheit verursacht werden, auch nicht. Die Vorstellung ist, dass das Modell nur für Zeiträume gültig sein soll, wo diese letzten Effekte keine Rolle spielen. Immunität nach einer Erkrankung entsteht bei der Ansteckung mit vielen Viren, allerdings nicht beim HIV.

Um Lösungen des SIR-Modells zu verstehen, das nunmehr zweidimensional ist, können wir folgendermassen vorgehen. Auf einem Zeitintervall auf dem S monoton ist können wir I als Funktion von S betrachten und die Gleichung

$$\frac{dI}{dS} = -1 + \frac{\alpha}{\beta S} \tag{72}$$

herleiten. In der Tat ist S immer monoton fallend, da $\dot{S} < 0$. Eine Integration zeigt, dass $I + S - \frac{\alpha}{\beta} \log S$ konstant ist längs der Integralkurven. Dadurch sieht man, dass jede Lösung gegen einen Punkt mit $I = 0$ konvergiert für $t \to \infty$. Diese Erhaltungsgröße ist nützlich für die Untersuchung des SIR-Modells. Hier soll darauf aufmerksam gemacht werden, dass sie eine formale Ähnlichkeit mit den Ljapunow-Funktion analysiert wurde. Im allgemeinen wird die Frage, ob eine Krankheit sich in einer Population behaupten kann, durch einen Parameter R_0 bestimmt. Impfung von Kindern kann benutzt werden, um den effektiven Wert von R_0 zu senken und dadurch das Fortschreiten der Krankheit zu bekämpfen. Wenn diese Massnahme zu $R_0 < 1$ führt spricht man von Herdenimmunität. Dieser Zustand ist nicht leicht zu erreichen. Für die Masern schätzt man, dass in den Industrieländern eine Impfrate zwischen 85 und 90 prozent in ländlichen Gebieten und wesentlich mehr als 90 prozent in Städten notwendig ist. In Entwicklungsländern sieht es ganz anders aus, da die Masern dort oft tödlich verlaufen.

Die Gleichungen, die hier betrachtet wurden haben auch eine Ähnlichkeit mit den berühmten Lotka-Volterra-Gleichungen für Räuber-Beute-Systeme. In dem Fall sind die Gleichungen

$$\dot{x} = x(\lambda - by) \tag{73}$$
$$\dot{y} = y(-\mu + cx) \tag{74}$$

und es gibt die Erhaltungsgröße

$$cx - \mu \log x + by - \lambda \log y. \tag{75}$$

Hier ist die Interpretation, dass x die Population der Beute ist (z.B. Hasen) und y die Population der Räuber (z. B. Luchse). In diesem Fall sind die Lösungen periodisch.

6 Der Satz von Arzela-Ascoli

Im nächsten Abschnitt wird die Existenz von invarianten Mannigfaltigkeiten bewiesen. Zu diesem Zweck brauchen wir den Satz von Arzela-Ascoli. Da wir eine Bekanntschaft mit diesem Satz nicht voraussetzen wollen, wird er hier bewiesen. Dieser Satz ist auch bei vielen anderen Fragen in der Theorie der dynamischen Systeme eine wertvolle Hilfe. Das Theorem wird zuerst in einem allgemeinen Zusammenhang bewiesen.

Definition Sei X ein metrischer Raum mit Metrik ρ und $\mathcal{F}$ eine Menge von reellwertigen Funktionen auf X. $\mathcal{F}$ heißt *gleichgradig stetig* wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass $|f(x) - f(y)| < \epsilon$ für alle $f \in \mathcal{F}$ und alle x und y mit $\rho(x, y) < \delta$. (Insbesondere ist jede Funktion $f \in \mathcal{F}$ gleichmässig stetig.)

$\mathcal{F}$ heißt punktweise beschränkt wenn es zu jedem $x \in X$ ein $M(x)$ gibt mit $|f(x)| \leq M(x)$ für alle $f \in \mathcal{F}$.

Satz (Arzela-Ascoli) Sei $\mathcal{F}$ eine punktweise beschränkte gleichgradig stetige Menge von reellwertigen Funktionen auf einem metrischen Raum X und es existiere eine abzählbare dichte Teilmenge $E \subset X$. Dann hat jede Folge $\{f_n\}$ von Funktionen aus $\mathcal{F}$ eine Teilfolge die gleichmässig auf jeder kompakten Teilmenge von X konvergiert.

Beweis Sei $x_1, x_2, x_3, \ldots$ eine Aufzählung der Punkte von E. Sei S_0 die Menge der natürlichen Zahlen. Sei $k \geq 1$ und nehmen wir an, dass eine unendliche Teilmenge S_{k-1} von S_0 gewählt worden ist. Da $\{f_n(x_k) : n \in S_{k-1}\}$ eine beschränkte Folge von reellen Zahlen ist, hat sie eine konvergente Teilfolge. Mit anderen Worten existiert eine unendliche Menge $S_k \subset S_{k-1}$, mit der Eigenschaft, dass $\lim_{n\to\infty} f_n(x_k)$ existiert für $n \in S_k$. Auf diese Weise bekommen wir unendliche Mengen $S_0 \supset S_1 \supset S_2 \supset \ldots$ mit der Eigenschaft, dass der Grenzwert von $f_n(x_k)$, für $1 \leq j \leq k$, wenn $n \to \infty$ innerhalb von S_k existiert. Sei r_k das $k$te Element von S_k und sei S die Folge $r_1, r_2, r_3, \ldots$. Für jeden Wert von k gibt es höchstens $k - 1$ Elemente von S die nicht in S_k sind. Deshalb existiert der Grenzwert $\lim f_n(x)$ für alle $x \in E$ wenn $n \to \infty$ innerhalb von S.

Sei $K \subset X$ kompakt und $\epsilon > 0$. Weil $\mathcal{F}$ gleichgradig stetig ist existiert ein $\delta > 0$ so dass $\rho(p, q) < \delta$ die Ungleichung $|f_n(p) - f_n(q)| < \epsilon$ impliziert für alle n. Es gibt eine Überdeckung von K durch endlich viele offene Kugeln $B_1, B_2, \ldots, B_M$ mit Radius $\frac{\delta}{2}$. Da E in X dicht liegt gibt es einen Punkt $p_i \in E$ in jedem B_i, $1 \leq i \leq M$. Der Grenzwert von $f_n(p_i)$ innerhalb von S existiert. Deshalb gibt es ein N mit der Eigenschaft, dass $|f_m(p_i) - f_n(p_i)| < \epsilon$ für $1 \leq i \leq M$ wenn $m > N$, $n > N$ und m und n in S liegen. Sei jetzt $x \in K$.

Dann gibt es ein i mit $x \in B_i$ und $\rho(x, p_i) < \delta$. Auf Grund der Wahl von N und δ gilt

$$\begin{aligned}&|f_m(x) - f_n(x)| \leq |f_m(x) - f_m(p_i)| + |f_m(p_i) - f_n(p_i)| + |f_n(p_i) - f_n(x)| \\ &< \epsilon + \epsilon + \epsilon = 3\epsilon \qquad (76)\end{aligned}$$

wenn $m > N$, $n > N$, $m \in S$ und $n \in S$.

Mit der gleichen Methode kann man die analoge Aussage für Funktionen mit Werten im $\mathbb{R}^k$ beweisen. Der Fall, der uns im Folgenden am meisten interessieren wird ist der wo $X = \mathbb{R}^m$ ist, mit der euklidischen Metrik. In diesem Fall kann man die Teilmenge E als die Menge der Punkte mit rationalen Koordinaten definieren.

7 Invariante Mannigfaltigkeiten

Im letzten Abschnitt haben wir Ergebnisse über das Verhalten der Lösungen eines dynamischen Systems in der Nähe einer stationären Lösung unter bestimmten Bedingungen bewiesen. Wenn alle Eigenwerte des linearisierten Systems negativen Realteil haben konvergieren Lösungen, die in der Nähe der stationären Lösung starten gegen die stationäre Lösung für $t \to \infty$. In dem wir t durch $-t$ ersetzen, bekommen wir analoge Ergebnisse im Fall, dass alle Eigenwerte positiven Realteil haben. Wenn die Realteile der Eigenwerte beide Vorzeichen haben, haben wir noch wenige Informationen. In diesem Abschnitt wollen wir mehr erfahren, in dem wir invariante Mannigfaltigkeiten betrachten. Eine invariante Mannigfaltigkeit ist eine Untermannigfaltigkeit die invariant ist unter dem Fluss. In diesem Zusammenhang ist es hilfreich die Einschränkung des Flusses $\phi(t, x)$ für einen festen Wert von t zu betrachten. Es handelt sich dabei um einen lokalen Diffeomorphismus zwischen Teilmengen des $\mathbb{R}^m$. Eine invariante Mannigfaltigkeit einer solchen Abbildung ist eine Untermannigfaltigkeit die durch den Diffeomorphismus in sich abgebildet wird.

Für eine reelle Zahl t sei T^t eine stetige Abbildung einer Umgebung G_t des Ursprungs im $\mathbb{R}^m$ auf eine Umgebung des Ursprungs im gleichen Raum mit $T^t(0) = 0$. Eine Menge S heißt invariant bezüglich $\{T^t\}$ wenn $T^t(G_t \cap S) \subset S$ für alle t. Sie heißt lokal invariant wenn es $\epsilon > 0$ gibt mit der Eigenschaft, dass wenn $x \in S$ und $|T^t(0)| < \epsilon$ für $0 \leq t_0$ dann ist $|T^t(x)| \in S$. Wir betrachten jetzt ein lineares System $\dot{x} = Ax$ und das gestörte System $\dot{x} = Ax + F(x)$ wo F eine stetig differenzierbare Abbildung ist die $o(|x|)$ ist für $x \to 0$. Die zweite Eigenschaft ist damit äquivalent, dass $F(0) = 0$ und $\frac{\partial F_i}{\partial x_j}(0) = 0$. Das erste System ist die Linearisierung des zweiten Systems im Ursprung. Wir definieren $T^t(x) = \phi(t, x)$ auf dem Gebiet, wo dieser Ausdruck existiert. Hier ist ϕ der Fluss des nichtlinearen Systems.

Wenn S invariant ist, dann ist der Durchschnitt von S mit einer Kugel lokal invariant. Wenn umgekehrt S lokal invariant, dann ist die Vereiningung der Mengen $T^t(S \cap G_t)$ invariant. Es gibt also enge Beziehungen zwischen invarianten und lokal invarianten Mengen. Dies ist aus folgendem Grund nützlich.

Wenn man F außerhalb einer kleinen Kugel verändert und eine invariante Menge für die neue Gleichung findet, dann bekommt man eine lokal invariante Menge für die ursprüngliche Gleichung. Ein anderer Vorteil von lokal invarianten Mengen ist, dass man erwarten kann, dass sie einfacher sind als global invariante Mengen. Zum Beispiel kann es Lösungen geben, die für $t \to \infty$ und $t \to -\infty$ gegen den Ursprung konvergieren und dazwischen schleifen machen.

Nehmen wir an, dass die Jacobi-Matrix in der Form $A = [P, Q]$ geschrieben werden kann, wo die Eigenwerte p_j von P die Ungleichung Re $p_j \leq \alpha < 0$ erfüllen und die Eigenwerte q_k von Q die Ungleichung Re $q_k \geq \beta > \alpha$. Koordinaten in den Teilräumen, die dieser Zerlegung entsprechen, werden mit y_j und z_k bezeichnet. Der Teilraum wo die Koordinaten z_k verschwinden ist die Vereinigung aller Lösungen, deren Abstand zum Ursprung für t groß durch $e^{(\beta-\epsilon)t}$ beschränkt ist für ein $\epsilon > 0$. Es stellt sich die Frage, ob eine ähnliche Aussage für ein nichtlineares System gilt. Wir betrachten ein System

$$\dot{y} = Py + F_1(y, z), \quad \dot{z} = Qy + F_2(y, z) \tag{77}$$

wo $F = (F_1, F_2)$ die gleichen Eigenschaften hat wie vorher. Gibt es eine lokal invariante Mannigfaltigkeit S von der Form $z = g(y)$ die aus allen Lösungen besteht, die für t groß durch $e^{(\beta-\epsilon)t}$ beschränkt sind für ein $\epsilon > 0$? Wir werden zeigen, dass die Antwort auf diese Frage positiv ist.

Die Lösung mit Anfangswert Null is identisch Null und existiert global in der Zeit. Deshalb ist es so dass für beliebig kleine Anfangswerte die Lösung beliebig lang existiert. Die Identität $T^{t_1+t_2} = T^{t_1}T^{t_2}$ gilt wo immer beide Seiten definiert sind. Diese Eigenschaft folgt aus der Tatsache, dass die Lösungen eindeutig durch ihre Anfangsdaten bestimmt sind. Der Fluss ϕ ist C^1 und seine Jacobi-Matrix $H(t, x)$ bezüglich x erfüllt die lineare Gleichung

$$\dot{H}(t, x) = \left[A + \frac{\partial F}{\partial x}\right] H(t, x), \quad H(0, x) = I. \tag{78}$$

Insbesondere gelten $\dot{H}(t, 0) = AH(t, 0)$ und $H(0, 0) = I$. Deshalb ist $H(t, 0) = e^{At}$. Es folgt, dass

$$\phi(t, x) = e^{At}x + \Xi(t, x), \tag{79}$$

wo

$$\Xi(t, 0) = 0 \quad \text{und} \quad \frac{\partial \Xi}{\partial x}(t, 0) = 0. \tag{80}$$

Es können bei der Konstruktion einer invarianten Mannigfaltigkeit technische Schwierigkeiten dadurch auftreten, dass die Lösungen nicht global in der Zeit definiert sind. Um dem vorzubeugen ersetzen wir die Funktion F durch eine andere die mit F identisch ist für x klein, z. B. für $|x| \leq \frac{1}{2}s$, und für $|x| \geq s$ verschwindet. Wenn wir die neue Funktion auch mit F bezeichnen dann ist der Fluss unserer Gleichung global definiert. Dann ist die Schar T^t eine Gruppe.

Hilfssatz 4 Sei $F(x)$ eine vektorwertige Funktion der Klasse C^1, die für $|x|$ klein definiert ist und $F(0) = 0$ und $\frac{\partial F}{\partial x}(0) = 0$ erfüllt. Sei $\theta > 0$ beliebig. Dann existiert eine Zahl $s = s(\theta) > 0$ (die mit θ gegen Null strebt) und eine

Funktion $G(x)$ der Klasse C^1 die für alle x definiert ist mit den Eigenschaften, dass $G(x) = F(x)$ für $|x| \leq \frac{1}{2}s$, $G(x) = 0$ für $|x| \geq s$ und $\|\frac{\partial G}{\partial x}\| \leq \theta$ für alle x.

Beweis Sei $s > 0$ so klein, dass $\|\partial F/\partial x\| \leq \theta/8$ und deshalb $|F(x)| \leq \theta\|x\|/8$ für $|x| \leq s$. Sei $\psi(t)$ eine glatte reellwertige Funktion von t für $t \geq 0$ mit $\psi(t) = 1$ für $t \leq (\frac{1}{2}s)^2$, $0 < \psi(t) < 1$ für $(\frac{1}{2}s)^2 < t < s^2$, $\psi(t) = 0$ für $t > s^2$ und $0 \leq -\frac{d\psi}{dt} \leq \frac{2}{s^2}$ für alle $t \geq 0$. Sei $G(x) = F(x)\psi(|x|^2)$ oder $G(x) = 0$, je nachdem, ob $|x| \leq s$ oder $|x| \geq s$. Dann ist $\partial G/\partial x = 0$ für $|x| \geq s$. Für $|x| \leq s$ ist $\frac{\partial G_i}{\partial x_j} = \frac{\partial F_i}{\partial x_j}\psi + 2F_i x_j \frac{d\psi}{dt}$ und deshalb ist $|\partial G/\partial x| \leq \theta/8 + 2(\theta\|x\|^2/8)(2/s^2) \leq \theta$. Damit ist der Hilfssatz bewiesen.

Auf Grund dieser Aussage ist es so, dass wenn man Lösungen in der Nähe des Ursprungs betrachtet man o.B.d.A. annehmen kann, dass F global definiert und C^1 ist, $\|\partial F/\partial x\| \leq \theta$ für alle x und $F(x) = 0$ für $|x| \geq s$. Hier darf s von θ abhängen.

Als nächstes soll gezeigt werden, dass es $s_0 = s_0(s,\theta) > 0$ gibt und $\theta_0 = \theta_0(s,\theta)$ so dass s_0 und θ_0 gegen Null streben wenn s und θ es tun und dass wenn ϕ wie in (79) geschrieben wird

$$\Xi(t, x_0) = 0, \qquad 0 \leq t \leq 1,\ |x_0| \geq s_0 \tag{81}$$

$$\|\partial\Xi/\partial x_0(t, x_0)\| \leq \theta_0, \qquad 0 \leq t \leq 1, x_0 \text{ beliebig.} \tag{82}$$

Um dies zu beweisen kann man zunächst die Tatsache verwenden, dass die Bedingung an der Ableitung von F impliziert, dass $|F(x)| \leq \theta|x|$ und dass deshalb die Lösung der Differentialgleichung $|\dot{x}| \leq c_0|x|$ erfüllt, wo $c_0 = \|A\| + \theta$. Deshalb ist

$$\frac{d}{dt}(e^{2c_0 t}|x(t)|^2) \geq 0 \tag{83}$$

und $|x(t)| \geq |x_0|e^{-c_0 t}$. Wenn also $s_0 = se^{c_0}$ und $|x_0| \geq s_0$ dann ist $|x(t)| \geq s$ für $0 \leq t \leq 1$. In diesem Fall reduziert sich die Gleichung auf dem Intervall $[0,1]$ auf $\dot{x} = Ax$. Die Lösung mit $x(0) = x_0$ ist also $x(t) = e^{At}x_0$. Deshalb ist $\Xi(t, x_0) = 0$ für $0 \leq t \leq 1$ und $|x_0| \geq s_0$.

Die Beziehung $\Xi(t, x_0) = \phi(t, x_0) - e^{At}x_0$ impliziert, dass $\frac{\partial\Xi}{\partial x_0} = H(t, x_0) - e^{At}$ oder

$$\frac{\partial\Xi}{\partial x_0} = e^{At}[K(t, x_0) - I], \tag{84}$$

wo $K(t, x_0) = e^{-At}H(t, x_0)$. Die Ableitung von K ist

$$\dot{K} = e^{-At}(\dot{H} - AH) = e^{-At}\frac{\partial\phi}{\partial x_0}e^{At}K. \tag{85}$$

Außerdem is $K(0, x_0) = I$. Die Größe $\|e^{-At}(\partial\phi/\partial x_0)e^{At}\|$ ist durch $c_1\theta$ beschränkt, wo $c_1 = (e^{\|A\|})^2$. Deshalb kann $\|K(t, x_0)\|$ durch $e^{c_1\theta}$ beschränkt werden für $0 \leq t \leq 1$. Es folgt, dass $\|\dot{K}(t, x_0)\| \leq c_1\theta e^{c_1\theta}$ und dass $\|K(t, x_0) - I\| \leq c_1\theta e^{c_1\theta}$ für $0 \leq t \leq 1$. Die Ungleichungen führen auf

$$\|\partial\Xi/\partial x_0\| \leq e^{\|A\|}c_1\theta e^{c_1\theta}. \tag{86}$$

Damit haben wir die gewünschte Bedingung mit $\theta_0 = e^{\|A\|}c_1\theta e^{c_1\theta}$.

Jetzt betrachten wir wieder das System (77). Sei $B = e^P$ und $C = e^Q$. Die Eigenwerte der Matrizen B bzw. C sind e^{p_j} bzw. e^{q_k}. Die Normen von B bzw. C^{-1} können durch $e^{\alpha+\epsilon}$ bzw. $e^{-\beta+\epsilon}$ beschränkt werden. Um dies zu sehen benutzt man die Tatsache, dass es sich um die Lösungen von linearen gewöhnlichen Differentialgleichungen handelt. Wir nehmen an, dass ϵ so klein ist, dass $b = \|B\|$ und $1/c = \|C^{-1}\|$ die Ungleichungen $b < c$ und $b < 1$ erfüllen. Es wird angenommen, dass die Funktionen F_1 und F_2 stetig differenzierbar sind, dass F_1, F_2 und ihre Ableitungen erster Ordnung im Ursprung verschwinden und dass die Normen dieser Ableitungen nicht größer sind als θ für alle x. Außerdem wird angenommen, dass F_1 und F_2 für $|x|^2 \geq s^2 > 0$ verschwinden. Dann definiert der Fluss für jeden Wert von t eine Abbildung T^t von (y_0, z_0) auf (y, z) der Form

$$y(t, y_0, z_0) = e^{Pt}y_0 + Y(t, y_0, z_0), \tag{87}$$

$$z(t, y_0, z_0) = e^{Qt}z_0 + Z(t, y_0, z_0) \tag{88}$$

wo Y, Z und ihre Ableitungen erster Ordnung im Ursprung verschwinden, die Normen der ersten Ableitungen nicht größer sind als θ_0 für $0 \leq t \leq 1$ und Y und Z verschwinden für $|y|^2 + |z|^2 \geq s_0^2$ und $0 \leq t \leq 1$.

Jetzt wird ein Ergebnis bewiesen über invariante Mannigfaltigkeiten einer Abbildung. Es wird dann auf T^t mit $t = 1$ angewendet um eine Aussage für invariante Mannigfaltigkeiten eines Flusses zu bekommen.

Hilfssatz 5 Sei B und C Matrizen mit den Eigenschaften, die oben aufgeführt wurden. Sei T eine Abbildung von (y_0, z_0) auf (y_1, z_1) mit

$$y_1 = By_0 + Y(y_0, z_0), \tag{89}$$

$$z_1 = Cz_0 + Z(y_0, z_0) \tag{90}$$

wo Y und Z stetig differenzierbar sind und die oben aufgeführten Bedingungen erfüllen. Dann existiert eine stetig differenzierbare Abbildung $z = g(y)$ mit $g(0) = 0$, $(\partial g/\partial y)(0) = 0$ so dass die Abbildung R mit

$$R(y, z) = (u, v) = (y, z - g(y)) \tag{91}$$

folgende Eigenschaften besitzt. Die Abbildung RTR^{-1}, die (u_0, v_0) auf (u_1, v_1) abbildet ist von der Form

$$u_1 = Bu_0 + U(u_0, v_0), \tag{92}$$

$$v_1 = Cv_0 + V(u_0, v_0) \tag{93}$$

wobei U, V und ihre Ableitungen erster Ordnung im Ursprung verschwinden und $V(u_0, 0) = 0$. Die letzte Bedingung bedeutet, dass der Teilraum $v_0 = 0$ unter der Abbildung RTR^{-1} invariant ist und dass die Mannigfaltigkeit $z = g(y)$ lokal invariant bezüglich T ist. Wenn wir die Reduktion ausführen, bekommen wir eine global invariante Mannigfaltigkeit für das transformierte System. Es kann angenommen werden, dass $\theta_0 < \min\left(\frac{c-b}{4}, \frac{1-b}{2}\right)$.

Beweis des Hilfssatzes Wenn R existiert dann gelten die Beziehungen

$$u_1 = Bu_0 + Y(u_0, v_0 + g(u_0)), \tag{94}$$

$$v_1 = Cv_0 + Cg(u_0) + Z(u_0, v_0 + g(u_0)) \\ -g(Bu_0 + Y(u_0, v_0 + g(u_0))). \tag{95}$$

Wenn wir für v_1 in der zweiten der Gleichungen einsetzen bekommen wir

$$V(u, v) = Cg(u) + Z(u, v + g(u)) - g(Bu + Y(u, v + g(u))) \tag{96}$$

und $V(u, 0) = 0$ gilt genau wenn

$$g(u) = C^{-1}[g(Bu + Y(u, g(u))) - Z(u, g(u))]. \tag{97}$$

Um den Hilfssatz zu beweisen müssen wir zeigen, dass die Funktionalgleichung (97) eine Lösung der Klasse C^1 besitzt. Zu diesem Zweck wird eine Folge g_n rekursiv definiert. Sei $g_0(u) = 0$ und wenn $g_{n-1}(u)$ schon definiert worden ist, sei

$$g_n(u) = C^{-1}[g_{n-1}(Bu + Y(u, g_{n-1}(u))) - Z(u, g_{n-1}(u))]. \tag{98}$$

Um diesen Ausdruck abzukürzen, seien $g_{n-1} = g_{n-1}(u)$, $Y^0 = Y(u, g_{n-1}(u))$, $g^0_{n-1} = g_{n-1}(Bu+Y^0)$ und $Z^0 = Z(u, g_{n-1}(u))$. Es ist klar, dass g_n wohldefiniert und C^1 ist für alle n. Wenn ∂g_n die Jacobi-Matrix von g_n ist dann gilt

$$\partial g_n = C^{-1}[\partial g^0_{n-1}(B + \partial_y Y^0 + (\partial_z Y^0)\partial g_{n-1}) \\ -(\partial_y Z^0 + (\partial_z Z^0)\partial g_{n-1})]. \tag{99}$$

Sei $\sigma = \frac{\theta_0}{c-b-3\theta_0}$, so dass $0 < \sigma < 1$. Es wird gezeigt durch Induktion, dass $\|\partial g_n(u)\| \leq \sigma$. Für $n = 0$ ist die Aussage offensichtlich. Nehmen wir jetzt an, dass die Aussage gilt, wenn n durch $n - 1$ ersetzt wird.

$$\|\partial g_n\| \leq c^{-1}[\sigma(b + \theta_0 + \theta_0\sigma) + (\theta_0 + \theta_0\sigma)] \tag{100}$$

$$\leq c^{-1}[\sigma(b + 3\theta_0) + \theta_0]. \tag{101}$$

Der letzte Ausdruck ist gleich σ und damit der induktive Schritt bewiesen.

Als nächstes wird gezeigt, dass die ∂g_n gleichgradig stetig sind. Für jede Funktion $f(u)$ oder $f(y, z)$ sei $\Delta f = f(u + \Delta u) - f(u)$ oder $\Delta f = f(y + \Delta y, z + \Delta z)$. Sei $h_1(\delta) = \sup \|\Delta\partial_{y,z}Y, Z\|$ für $\|\Delta y\|, \|\Delta z\| \leq \delta$ wo $\partial_{y,z}Y, Z$ irgendeine der Jacobi-Matrizen $\partial_y Y, \partial_z Y, \partial_y Z, \partial_z Z$ bezeichnet. Es wird gezeigt durch Induktion, dass $\|\Delta\partial g_n\| \leq h(\delta)$ für $\|\Delta u\| \leq \delta < 1$, wo

$$h(\delta) = \frac{4h_1(\delta)}{c - b - 4\theta_0}. \tag{102}$$

Für $n = 0$ ist die Aussage offensichtlich. Nehmen wir jetzt an, dass die Aussage gilt, wenn n durch $n - 1$ ersetzt wird. Es gilt $\|\Delta g_{n-1}(u)\| \leq \sigma\|\Delta u\| \leq \|\Delta u\|$. Daraus folgt, dass

$$\|\Delta\partial_{y,z}Y^0, Z^0\| \leq h_1(\|\Delta u\|) \tag{103}$$

und

$$\|\Delta(Bu + Y(u, g_{n-1}(u)))\| \le (b + 2\theta_0)\|\Delta u\| \le \|\Delta u\|. \tag{104}$$

Jetzt verwenden wir die allgemeine Beziehung

$$\Delta(f_1(u)f_2(u)) = f_1(u + \Delta u)\Delta f_2 + \Delta f_1 f_2(u) \tag{105}$$

auf den Ausdruck für ∂g_n.

$$\begin{aligned}\|\Delta\partial g_n\| &\le c^{-1}[h(\delta)(b + 2\theta_0) + (h_1(\delta) + h_1(\delta) + \theta_0 h(\delta)) \\ &+(h_1(\delta) + h_1(\delta) + \theta_0 h(\delta))] = c^{-1}[h(\delta)(b + 4\theta_0) + 4h_1(\delta)].\end{aligned} \tag{106}$$

Die rechte Seite ist $h(\delta)$. Jetzt wird gezeigt, dass die Folge g_n gleichmässig auf jeder beschränkten Teilmenge konvergiert. Dies ist dann gegeben, wenn es Konstanten M und r gibt mit $0 < r < 1$, so dass für $n \ge 1$

$$\|g_n(u) - g_{n-1}(u)\| \le M\|u\|r^n. \tag{107}$$

Für $n = 1$ gilt die Ungleichung wenn $Mr = \sigma$. Nehmen wir jetzt an, dass sie gilt wenn n durch $n-1$ ersetzt wird. Die Größe $c\|g_n(u)-g_{n-1}(u)\|$ kann beschränkt werden durch

$$\begin{aligned}&\|g_{n-1}(Bu + Y(u, g_{n-1}(u))) - g_{n-2}(Bu + Y(u, g_{n-2}(u)))\| + \\ &\|Z(u, g_{n-1}(u)) - Z(u, g_{n-2}(u))\|.\end{aligned} \tag{108}$$

Der erste Term ist nicht größer als

$$\begin{aligned}&\|g_{n-1}(Bu + Y(u, g_{n-1}(u))) - g_{n-2}(Bu + Y(u, g_{n-1}(u)))\| + \\ &\|g_{n-2}(Bu + Y(u, g_{n-1}(u))) - g_{n-2}(Bu + Y(u, g_{n-2}(u)))\|.\end{aligned} \tag{109}$$

Es folgt, dass $c\|g_n(u) - g_{n-1}(u)\|$ nicht größer ist als

$$M\|Bu + Y(u, g_n)\|r^{n-1} + \sigma\theta_0 M\|u\|r^{n-1} + \theta_0 M\|u\|r^{n-1}, \tag{110}$$

was wiederum nicht größer ist als $Mr^{n-1}\|u\|(b+4\theta_0)$. Mit der Wahl $r = (b+4\theta_0)$ ist der induktive Schritt vollzogen.

Es folgt dass $g_n(u)$ gegen einen Grenzwert $g(u)$ konvergiert, gleichmässig auf jeder kompakten Teilmenge. Man kann dann zum Grenzfall übergehen und sehen, dass g die Funktionalgleichung erfüllt. Da die Folge ∂g_n gleichmässig beschränkt und gleichgradig stetig ist, gibt es nach dem Satz von Arzela-Ascoli eine Teilfolge, die gleichmässig auf jeder kompakten Teilmenge konvergiert. Es folgt, dass g stetig differenzierbar ist. Damit ist der Hilfssatz bewiesen.

Es gibt entsprechende Aussagen wenn C^1 mit C^r ersetzt wird, r endlich oder unendlich. Die Bedingung $b < 1$ ist nicht notwendig. Diese verschärften Aussagen werden hier nicht bewiesen.

Korollar 1 Seien T, $g(y)$ und θ_0 wie im Hilfssatz. Für (x_0, y_0) gegeben definieren wir eine Folge rekursiv durch $(y_{n+1}, z_{n+1}) = T(y_n, z_n)$. Wenn $z_0 = g(y_0)$ dann gilt $\|(y_n, z_n)\| = O((b + \theta_0)^n)$ für $n \to \infty$. Es ist auch so, dass wenn $y_0 \ne 0$

dann ist $y_n \neq 0$ für alle n, $\|z_n\|/\|y_n\| \to 0$ und $\limsup n^{-1} \log \|(y_n, z_n)\| \leq \alpha$. Wenn $z_0 \neq g(y_0)$ dann gilt $(c - 2\theta_0)^n = O(\|(y_n, z_n)\|)$ für $n \to \infty$.

Wenn $c > 1$, so dass $b < 1 < c$ kann man die Punkte (y_0, z_0) der Mannigfaltigkeit $z = g(y)$ durch drei alternative Bedingungen an den Punkten charakterisieren. Die erste Bedingung ist dass, mit $(y_n, z_n) = T^n((y_0, z_0))$, $\|(y_n, z_n)\|$ exponentiell gegen Null konvergiert für $n \to \infty$. Die zweite ist, dass sie gegen Null konvergiert für $n \to \infty$. Die dritte ist, dass sie in einer Umgebung von $(0, 0)$ bleibt. In diesem Fall heißt die Mannigfaltigkeit $z = g(y)$ stabile Mannigfaltigkeit von T. Die entsprechende Mannigfaltigkeit wenn man n durch $-n$ ersetzt heißt instabile Mannigfaltigkeit von T.
Beweis des Korollars $z_0 = g(y_0)$ ist mit der Bedingung $v_0 = 0$ äquivalent. In diesem Fall ist $v_n = 0$ für alle n. Entsprechend ist $u_n = Bu_{n-1} + U(u_{n-1}, 0)$, so dass $\|u_n\| \leq (b+\theta_0)\|u_{n-1}\|$ und $\|u_n\| \leq (b+\theta_0)^n\|u_0\|$. Insbesondere konvergiert u_n gegen den Ursprung. Es folgt, dass es für $\epsilon > 0$ beliebig ein N gibt, so dass $\|u_n\| \leq (b + \epsilon)\|u_{n-1}\|$ für $n \geq N$ und $\|u_{n+N}\| \leq (b + \epsilon)^n\|u_N\|$ für $n \geq 0$. Da $y_n = u_n$ und $z_n = g(u_n)$ gilt $\|(y_n, z_n)\| \leq (1 + \sigma)\|u_n\|$. Außerdem ist $\limsup n^{-1} \log \|(y_n, z_n)\| \leq \log b$. Durch eine geeignete Wahl der Variablen kann $\log b$ beliebig nah an α gebracht werden und deshalb kann in dieser Beziehung $\log b$ durch α ersetzt werden. In dem wir die Gleichung für $V(u, v)$ nach v ableiten bekommen wir

$$\partial_v V(u, v) = \partial_z Z(u, v + g(u)) - \partial g(Bu + Y(u, v + g(u)))\partial_z Y(u, v + g(u)). \quad (111)$$

Deshalb ist $\|\partial_v V\| \leq \theta_0 + \sigma\theta_0 \leq 2\theta_0$ und $\|V(u, v)\| \leq 2\theta_0\|v\|$. Dann folgt aus der Gleichung $v_n = Cv_{n-1} + V(u_n, v_n)$ dass $\|v_n\| \geq (c - 2\theta_0)\|v_{n-1}\|$ und $\|v_n\| \geq (c - 2\theta_0)^n\|v_0\|$. Außerdem gilt

$$\|(y_n, z_n)\| \geq \|(u_n, v_n)\| - \|g(u_n)\| \geq (1 - \sigma)\|(u_n, v_n)\|. \quad (112)$$

Damit ist das Korollar bewiesen.

Satz 7 Sei T eine Abbildung einer Umgebung des Ursprungs im $\mathbb{R}^m$ nach $\mathbb{R}^m$ der Form

$$x_1 = T(x_0) = \Gamma x_0 + \Xi(x_0), \quad (113)$$

wo Ξ stetig differenzierbar ist, $\Xi(0) = 0$, $\partial\Xi/\partial x_0(0) = 0$ und Γ eine Matrix ist mit d bzw. e_0 bzw. e Eigenwerten, deren Eigenwerte betragsmässig < 1 bzw. $= 1$ bzw. > 1 sind. Dann gibt es eine stetig differenzierbare Abbildung R mit nichtverschwindender Jacobi-Determinante so dass RTR^{-1} folgende Form hat

$$\begin{aligned} u_1 &= Au_0 + U(u_0.v_0, w_0), \\ w_1 &= Bw_0 + W(u_0.v_0, w_0), \\ v_1 &= Cv_0 + V(u_0.v_0, w_0). \end{aligned} \quad (114)$$

Hier sind A bzw. B bzw. C Matrizen die $d \times d$ bzw. $e_0 \times e_0$ bzw. $e \times e$ und deren Eigenwerte betragsmässig < 1 bzw. $= 1$ bzw. > 1 sind. U, V und W und ihre Ableitungen erster Ordnung verschwinden im Ursprung und

$$V = 0, W = 0 \quad \text{wenn} \quad v_0 = 0, w_0 = 0 \tag{115}$$

$$U = 0, W = 0 \quad \text{wenn} \quad u_0 = 0, w_0 = 0. \tag{116}$$

Diese Gleichungen bedeuten, dass die Ebenen $v_0 = 0$, $w_0 = 0$ bzw. $u_0 = 0$, $w_0 = 0$ invariante Ebenen der Dimension d bzw. e sind. Wenn $e_0 = 0$ dann fehlen die Variablen w_0 und w_1.

Beweis Der Hilfssatz 5 liefert eine Abbildung R_0, so dass nach der Transformation mit R_0 die erste Bedingung erfüllt ist. Wenn wir die Abbildung T^{-1} betrachten, dann liefert der Hilfssatz 5 eine Abbildung R_1, so dass nach der Transformation mit R_1 beide Bedingungen erfüllt sind.

Korollar 2 Sei T^t eine Gruppe von Abbildungen, die durch die Gleichungen (87)-(88) definiert ist. Sei g die Funktion aus dem Hilfssatz 5 mit $T = T^1$. Dann hat RT^tR^{-1} die Form

$$u(t, u_0, v_0) = e^{Pt}u_0 + U(t, u_0, v_0), \tag{117}$$

$$v(t, u_0, v_0) = e^{Qt}u_0 + V(t, u_0, v_0), \tag{118}$$

wobei $V(t, u_0, 0) = 0$ für alle t und u_0. Außerdem gilt, dass wenn $y_0 \neq 0$ und $z_0 = g(y_0)$ dann ist $z(t) = g(y(t))$ für alle t, $y(t) \neq 0$ für alle t, $\|z(t)\|/\|y(t)\| \to 0$ und $\limsup t^{-1} \log \|y(t)\| \leq \alpha$ für $t \to \infty$. Wenn $z_0 \neq g(y_0)$ dann gilt $(c-2\theta_0)^t = O(\|(y(t), z(t))\|)$ für $t \to \infty$.

Beweis Zuerst wird gezeigt, dass wenn $n \leq t \leq n + 1$ dann gibt es positive Konstanten c_1 und c_2 mit

$$c_1\|(y(n), z(n))\| \leq \|(y(t), z(t))\| \leq c_2\|(y(n), z(n)). \tag{119}$$

Dazu ist zu bemerken, dass $T^t = T^{t-n}T^n$. Deshalb gilt

$$\|y(t) - e^{P(t-n)}y(n)\| \leq \theta_0(\|y(n)\| + \|z(n)\|) \tag{120}$$

und eine ähnliche Ungleichung für $z(t)$. Die Ungleichungen (119) folgen. Sei $z_0 = g(y_0)$. Dann wird das Verhalten von $(y(n), z(n))$ für n groß durch Korollar 1 beschrieben. Man bekommt die Ungleichung $\limsup t^{-1} \log \|(y(t), z(t))\| \leq \alpha$. Wenn $z(t)$ für irgendein t, sagen wir $t = t_0$ nicht gleich $g(y(t))$ wäre, dann wäre $(c - 2\theta_0)^n = O(\|(y(n + t_0), y(n + t_0))\|)$ für $t \to \infty$. Damit hätten wir aber einen Widerspruch. Deshalb ist $z(t) = g(y(t))$ für alle t. Wenn $y(t) = 0$ für irgendein t, dann impliziert $z(t) = g(y(t))$ dass $\|z(t)\| \leq \sigma\|y(t)\| = 0$. Aber dann ist $(y(t), z(t)) = 0$ für alle t auf Grund der Gruppeneigenschaft von T^t.

Nachdem wir die Existenz von invarianten Mannigfaltigkeiten für Abbildungen gezeigt haben, können wir zu den Differentialgleichungen zurückkehren.

Satz 8 Im dynamischen System

$$\dot{x} = Ax + F(x) \tag{121}$$

sei F stetig differenzierbar, $F(0) = 0$ und $\partial F/\partial x(0) = 0$. Nehmen wir an, dass es r Eigenwerte von A gibt mit negativen Realteilen $\alpha_1 < \alpha_2 < \ldots < \alpha_r < 0$ und

dass die anderen Eigenwerte, falls vorhanden, nicht-negative Realteile haben. Sei d die Summe der Dimensionen d_i der Räume der verallgemeinerten Eigenvektoren die den Eigenwerten α_i entsprechen. Wenn $0 < \epsilon < -\alpha_r$ dann gibt es Lösungen, die die Bedingung $\|x(t)\|e^{\epsilon t} \to 0$ für $t \to \infty$ erfüllen und jede solche Lösung erfüllt $\lim t^{-1} \log \|x(t)\| = \alpha_i$ für ein bestimmtes i. Für ein hinreichend kleines $\epsilon > 0$ bilden der Punkt $x = 0$ und die Menge der Punkte x auf Lösungen $x(t)$ mit $\lim t^{-1} \log \|x(t)\| \leq \alpha_i$ für i fest eine stetig differenzierbare lokal invariante Mannigfaltigkeit der Dimension $d_1 + \ldots + d_i$. Die entsprechende Menge, die durch die Bedingung $\limsup t^{-1} \log \|x(t)\| < 0$ definiert wird ist eine stetig differenzierbare Mannigfaltigkeit der Dimension d.
Beweis Wenn lim durch lim sup ersetzt wird, dann folgt der letzte Teil des Satzes aus Korollar 2. Außerdem folgt, dass für $t \to \infty$ die Bedingung $\liminf t^{-1} \log \|x(t)\| < \alpha_{i+1}$ die Bedingung $\limsup t^{-1} \log \|x(t)\| \leq \alpha_i$ impliziert, wobei α_{r+1} als Null interpretiert wird. Deshalb impliziert die Bedingung, dass $\limsup t^{-1} \log \|x(t)\| = \alpha_i$, dass $\liminf = \limsup$.

Man bekommt ähnliche Ergebnisse in dem man t durch $-t$ ersetzt. Die Argumente, die in den Beweisen von Satz 8 und Korollar 2 angewendet werden liefern

Satz 9 Seien A und F wie im letzten Satz. Es existieren außerdem noch e Eigenwerte von A mit positivem Realteil. Sei $x(t)$ die Lösung des Systems mit $x(0) = x_0$ und T^t die entsprechende Abbildung. Sei $\epsilon > 0$. Es gibt eine Abbildung R mit nichtverschwindender Jacobi-Determinante so dass RT^tR^{-1} folgende Form hat.

$$u(t) = e^{Pt}u_0 + U(u_0, v_0, w_0),$$
$$w(t) = e^{P_0 t}w_0 + W(u_0, v_0, w_0), \tag{122}$$
$$v(t) = e^{Qt}v_0 + V(u_0, v_0, w_0). \tag{123}$$

U, V und W und ihre Ableitungen erster Ordnung verschwinden im Ursprung. Wenn $v_0 = w_0 = 0$ dann ist $V = W = 0$ und wenn $u_0 = w_0 = 0$ dann ist $U = W = 0$. Außerdem ist $\|e^P\| < 1$, $\|e^{-Q}\| < 1$ und die Eigenwerte von P_0 haben Betrag Eins. Die Abbildung R transformiert die Gleichung für x in die Gleichung

$$\dot{u} = Pu + F_1(u, v, w), \tag{124}$$
$$\dot{w} = P_0 w + F_2(u, v, w), \tag{125}$$
$$\dot{v} = Qv + F_3(u, v, w), \tag{126}$$

wobei die F_i nicht C^1 sein müssen. Die Ebenen $v_0 = w_0 = 0$ und $u_0 = w_0 = 0$ sind lokal invariante Mannigfaltigkeiten und heißen stabile und instabile Mannigfaltigkeiten. Im Fall dass A keine Eigenwerte mit Realteil Null werden sie Durch die Bedingungen charakterisiert, dass sie aus den Lösungen bestehen, die für $t \to +\infty$ oder $t \to -\infty$ gegen den Ursprung konvergieren.

Um diese allgemeinen Ideen zu illustrieren können wir das fundamentale Modell der Virus-Dynamik mit $R_0 \neq 1$ betrachten. Der einzige Fall in diesem

Beispiel wo eine stationäre Lösung Eigenwerte hat, deren Realteile beide Vorzeichen haben ist die Lösung mit $v = 0$ im Fall $R_0 > 1$. Dort ist die stabile Mannigfaltigkeit zweidimensional und die instabile eindimensional. Die Gerade $v = y = 0$ ist invariant und liegt in der stabilen Mannigfaltigkeit. Die stabile Mannigfaltigkeit kann den positiven Bereich nicht treffen, wegen dem Satz von Korobeinikov. Um etwas über die instabile Mannigfaltigkeit in diesem Fall zu verstehen, betrachten wir den positiven Eigenwert der Linearisierung, nennen wir ihn μ_+. Sei $(\hat{x}, \hat{y}, 1)$ die Komponenten eines entsprechenden Eigenvektors. Dann gelten die Beziehungen $\hat{y} = \frac{\beta x}{\mu_+ + a}$ und $\hat{x} = -\frac{\beta x}{\mu_+ + d}$. Weil die zweite und dritte Komponenten des Vektors positiv sind ist es geometrisch klar, dass die eine Hälfte der instabilen Mannigfaltigkeit im positiven Bereich liegt.

8 Zentrumsmannigfaltigkeiten

Im letzten Abschnitt wurde die Existenz von stabilen und instabilen Mannigfaltigkeiten bewiesen. Beide haben äquivalente Eigenschaften und deshalb konzentrieren wir uns hier auf den Fall der stabilen Mannigfaltigkeit. Es handelt sich um eine invariante Mannigfaltigkeit durch eine stationäre Lösung x_0, deren Tangentenraum in x_0 mit dem stabilen Teilraum in diesem Punkt übereinstimmt. Durch diese Eigenschaften ist sie in einer hinreichend kleinen Umgebung von x_0 eindeutig bestimmt. Wenn das System C^k für eine natürliche Zahl $k \geq 1$ dann ist die Mannigfaltigkeit C^k. Wenn das System C^∞ bzw. analytisch ist, ist die Mannigfaltigkeit C^∞ bzw. analytisch.

Angesichts dieser Tatsachen kann man Fragen, ob es für den Zentrumsteilraum V_c auch ein Analogon der stabilen Mannigfaltigkeit gibt. Mit anderen Worten, gibt es eine invariante Mannigfaltigkeit M_c durch x_0 deren Tangentenraum mit V_c übereinstimmt? Es gibt in der Tat eine solche Mannigfaltigkeit, die Zentrumsmannigfaltigkeit. Es würde zu weit führen, wenn wir diese Aussage in dieser Vorlesung beweisen würden. Stattdessen begrenzen wir uns darauf, Aussagen über solche Mannigfaltigkeiten zu formulieren und zu zeigen wie die Existenzaussage in konkreten Anwendungen nützlich sein kein. Weitere Informationen über Zentrumsmannigfaltigkeiten findet man in [1].

Wenn x_0 eine stationäre Lösung eines dynamischen Systems ist, und V_c nicht trivial ist, dann gibt es eine Zentrumsmannigfaltigkeit. Sie muss aber keineswegs eindeutig sein, wie man schon bei folgendem einfachen System sieht.

$$\dot{x} = -x, \tag{127}$$

$$\dot{y} = y^2. \tag{128}$$

Der Ursprung ist eine stationäre Lösung. Der Zentrumsteilraum ist die Menge $x = 0$ und sie ist auch eine Zentrumsmannigfaltigkeit in diesem Fall, aber nicht die einzige. Alle Lösungen dieser Gleichung außer der stationären Lösung werden durch $x = ae^{-t}$, $y = -1/(t+b)$ gegeben. Wenn wir für y als Funktion von x auflösen, dann bekommen wir $x = Ce^{1/y}$. Für $y > 0$ gibt es invariante Mannigfaltigkeiten die in jeder Ordnung tangential zum Zentrumsteilraum sind. Es ist in der Tat so, dass obwohl M_c nicht eindeutig ist, die Ableitungen im Punkt x_0 immer eindeutig bestimmt sind. Wenn das Systen C^k ist, mit k endlich, dann gibt es auch eine Zentrumsmannigfaltigkeit, die C^k ist. Wenn das System C^∞ ist, dann gibt es nicht immer eine Zentrumsmannigfaltigkeit, die C^∞ ist, obwohl es solche gibt, die C^k sind für jeden festen Wert von k. Für ein analytisches System gibt es nicht immer eine analytische Zentrumsmannigfaltigkeit.

Wie kann es sein, dass eine Zentrumsmannigfaltigkeit in Anwendungen nützlich sein kann, wenn man sie nicht explizit kennt und wenn sie nicht einmal eindeutig ist? Wie wir später sehen werden bestimmt die Dynamik der Einschränkung des Systems auf eine Zentrumsmannigfaltigkeit die ganze Dynamik in der Nähe der stationären Lösung. Außerdem kann man manchmal die Dynamik auf der Zentrumsmannigfaltigkeit bestimmen, ohne diese Mannigfaltigkeit zu kennen.

Als ein erstes Beipiel betrachten wir das System

$$\dot{x} = xy + ax^3, \tag{129}$$

$$\dot{y} = -y + cx^2. \tag{130}$$

Die Zentrumsmannigfaltigkeit ist von der Form $y = \phi(x)$, wobei $\phi(x) = O(x^2)$. Aus der Gleichung $\dot{y} = \phi'(x)\dot{x} = O(x^3)$ folgt, dass auf der Zentrumsmannigfaltigkeit $y = cx^2 + O(x^3)$. Dann ist aber $\dot{x} = (a+c)x^3 + O(x^4)$. Für $a + c > 0$ ist der Ursprung instabil. Für $a + c < 0$ nimmt x entlang der Zentrumsmannigfaltigkeit ab. Wir werden später sehen, dass in diesem Fall die asymptotische Stabilität des Ursprungs folgt.

Im Falle des fundamentalen Modells der Virusdynamik mit $R_0 = 1$ hat die stationäre Lösung im Punkt $(\lambda/d, 0, 0)$ eine eindimensionale Zentrumsmannigfaltigkeit. Der Tangentenraum zu dieser Mannigfaltigkeit in diesem Punkt wird durch den Vektor $(-au, du, dk)$ aufgespannt. Entlang der Zentrumsmannigfaltigkeit haben wir

$$x = \frac{\lambda}{d} - \frac{au}{dk}v + \psi_1(v), \tag{131}$$

$$y = \frac{u}{k}v + \psi_2(v). \tag{132}$$

Die Funktionen ψ_1 und ψ_2 sind $O(v^2)$. Wenn wir die Gleichung für y differenzieren bekommen wir $\dot{y} = \left(\frac{u}{k} + \psi_2'\right)\dot{v} = u\psi_2 + O(v^3)$. Auf der anderen Seite liefert die Evolutionsgleichung für y

$$\dot{y} = -\frac{\beta au}{dk}v^2 - a\psi_2 + O(v^3). \tag{133}$$

Deshalb ist $\psi_2(v) = -\frac{\beta au}{(a+u)dk}v^2 + O(v^3)$ und v nimmt entlang der Zentrumsmannigfaltigkeit ab. Wie im letzten Beispiel kann man daraus die asymptotische Stabilität der stationären Lösung schließen.

9 Der Satz von Grobman und Hartman

Eine stationäre Lösung eines dynamischen Systems wo die Linearisierung keine rein imaginären Eigenwerte besitzt heißt *hyperbolisch*. Der Satz 9 zeigt, dass man ein System in der Nähe einer hyperbolischen stationären Lösung durch eine Transformation vereinfachen kann, so dass sie mehr wie das linearisierte System aussieht. Kann aber ein System in dieser Situation durch eine Transformation gänzlich linear gemacht werden? Betrachten wir das System

$$\dot{x} = Ax + F(x) \tag{134}$$

im Fall, dass kein Eigenwert von A verschwindenden Realteil hat. Gibt es einen lokalen Diffeomorphismus R der Klasse C^1 so dass $y = R(x)$ die Gleichung $\dot{y} = Ay$ erfüllt? Im allgemeinen ist die Antwort auf diese Frage negativ, schon im zweidimensionalen Fall. Diese Aussage wird hier nicht bewiesen. Wenn man aber anstatt einer stetig differenzierbaren nur eine stetige Abbildung R verlangt, dann sieht es besser aus. Es gilt nämlich folgender Satz

Satz (Grobman-Hartman) Nehmen wir an, dass im System (134) die Matrix A keine Eigenwerte mit verschwindendem Realteil besitzt, und dass F stetig differenzierbar ist, mit $F(0) = 0$ und $\partial F/\partial x(0) = 0$. Seien ϕ bzw. ψ die Flüsse von (134) bzw. des Systems $\dot{y} = Ay$. Dann gibt es eine stetige injektive Abbildung R einer Umgebung von $x = 0$ in den $\mathbb{R}^m$, so dass $R(\phi(t, x_0)) = \psi(t, R(x_0))$ für x_0 in einer Umgebung des Ursprungs und t klein. Inbesondere bildet R Lösungen auf Lösungen, wobei die Parametrisierung behalten wird. Die Systeme sind also topologisch konjugiert.

Die entsprechende Aussage gilt nicht immer wenn es rein imaginäre Eigenwerte gibt. Eine Verallgemeinerung auf diesen Fall wird später vorgestellt. Wie bei den invarianten Mannigfaltigkeiten wird der Beweis des Satzes über Flüsse durch einen entsprechenden Satz über Abbildungen vorbereitet.

Hilfssatz 6 Seien B bzw. C invertierbare Matrizen die $d \times d$ bzw. $e \times e$ sind und die Ungleichungen $b = \|B\| < 1$ und $c^{-1} = \|C^{-1}\| < 1$ erfüllen. Sei T eine Abbildung der Form

$$T(y_0, z_0) = (By_0 + Y(y_0, z_0), Cz_0 + Z(y_0, z_0)) \tag{135}$$

wo Y und Z Funktionen der Klasse C^1 sind die mit ihren ersten Ableitungen im Ursprung verschwinden. Dann existiert eine stetige injektive Abbildung $R(u, v) = (\Phi(u, v), \Psi(u, v))$ einer Umgebung des Ursprungs im $\mathbb{R}^m$ auf eine ebensolche, die T in die lineare Abbildung A transformiert, $A = RTR^{-1}$, wobei $A(u_0, v_0) = (u_1, v_1) = (Bu_0, Cv_0)$.

Wir können die Funktion F sowie in früheren Beweisen abschneiden und die Abbildung R für das abgeschnittene System existiert dann auf ganz $\mathbb{R}^m$. Bevor wir Hilfssatz 6 beweisen brauchen wir zwei andere Aussagen.

Hilfssatz 7 Sei L eine invertierbare $m \times m$-Matrix und sei $l_1 = \|L^{-1}\|$. Sei S eine Abbildung der Form $x_1 = S(x_0) = Lx_0 + X(x_0)$, wobei X auf ganz $\mathbb{R}^m$ definiert ist und erfüllt eine Lipschitz-Bedingung

$$\|X(x_0 + \Delta x_0) - X(x_0)\| \leq \theta_1 \|\Delta x_0\| \tag{136}$$

wo $\theta_1 l_1 < 1$. Dann ist S injektiv und surjektiv auf $\mathbb{R}^m$. Wenn außerdem $\|X(x_0)\| \leq c_0$ für alle x_0 und die Umkehrabbildung von S von der Form $L^{-1}+X_1$ ist, dann gilt $\|X_1(x_1)\| \leq l_1 c_0\|$ für alle x_1.
Beweis Es gilt

$$\|L^{-1}[X(x_0+\Delta x_0)-X(x_0)]\| \leq l_1\theta_1\|\Delta x_0\| \tag{137}$$

Um zu sehen, dass S injektiv ist reicht es zu beweisen, dass $L^{-1}S$ injektiv ist. Wenn x_0 und $x_0+\Delta x_0$ das gleiche Bild unter $L^{-1}S$ haben dann gilt

$$0 = \|x_0+\Delta x_0+L^{-1}X(x_0+\Delta x_0)-x_0-L^{-1}X(x_0)\| \geq (1-l_1\theta_1)\|\Delta x_0\|, \tag{138}$$

woraus man schließen kann, dass $\Delta x_0 = 0$. Um zu sehen, dass S surjektiv ist reicht es zu sehen, dass $L^{-1}S$ surjektiv ist, d.h. dass für x_1 gegeben es ein x_0 gibt mit $x_1 = x_0 + L^{-1}X(x_0)$. Um die Existenz von x_0 benutzen wir eine Iteration. Sei $x^0 = 0$ und $x^n = x_1 - L^{-1}X(x^{n-1})$ für $n \geq 1$. Für $n \geq 2$ gilt

$$\|x^n - x^{n-1}\| \leq \|L^{-1}[X(x^{n-1})-X(x^{n-2})]\| \leq l_1\theta_1\|x^{n-1}-x^{n-2}\|. \tag{139}$$

Es folgt, dass $\|x^n - x^{n-1}\| \leq (l_1\theta_1)^{n-1}\|x^1 - x^0\|$ für $n \geq 1$. Da $0 < b_1\theta_1 < 1$ hat die Folge $\{x^n\}$ einen Grenzwert, sagen wir x_0 für $n \to \infty$. Damit ist bewiesen, dass S surjektiv ist. Deshalb definiert S eine Bijektion zwischen x_0 und $x_1 = S(x_0)$. Es gilt

$$X^1(x_1) = x_0 - L^{-1}x_1 = L^{-1}(Lx_0 - x_1) = -L^{-1}X(x_0) \tag{140}$$

und damit ist Hilfssatz 7 bewiesen.

Als nächstes brauchen wir etwas Terminologie. Die Matrix B vom Hilfssatz 6 hat eine inverse. Sei $b_1 = \|B^{-1}\|$. Seien a_1, θ_1 und θ Konstanten die die Ungleichungen $a_1 = \max\{b, 1/c\} > 0$, $0 < b_1\theta_1 < 1$ und

$$\theta = \theta_1(1+c) + \max\{b, c\} < 1 \tag{141}$$

erfüllen. Wenn c_0 eine positive Konstante ist, sei $\Omega(\theta_1, c_0)$ die Menge aller Paare von Funktionen $(Y(y_0,z_0), Z(y_0,z_0))$ die für alle (y_0,z_0) folgende Beziehungen erfüllen. $Y(0,0)=0$, $Z(0,0)=0$, $\|Y(y_0,z_0)\|+\|Z(y_0,z_0)\| \leq c_0$,

$$\|\Delta Y\|, \|\Delta Z\| \leq \frac{1}{2}\theta_1(\|\Delta y_0\|+\|\Delta z_0\|), \tag{142}$$

wo Δ wieder eine Differenz bezeichnet. Hilfssatz 6 ist im Fall $Y_1 = 0$, $Z_1 = 0$ des folgenden Ergebnisses erhalten.

Hilfssatz 8 Seien B und C wie im Hilfssatz 6, (Y,Z) und (Y_1,Z_1) ein Paar von Elementen von $\Omega(\theta_1, c_0)$ und

$$(y_1,z_1) = T(y_0,z_0), y_1 = By_0 + Y(y_0,z_0), z_1 = Cz_0 + Z(y_0,z_0), \tag{143}$$
$$(y_1,z_1) = U(y_0,z_0), y_1 = By_0 + Y_1(y_0,z_0), z_1 = Cz_0 + Z_1(y_0,z_0). \tag{144}$$

Dann existiert eine eindeutige stetige Abbildung

$$(u, v) = R_0(y, z), u = y + \Lambda(y, z), v = z + \Theta(y, z), \tag{145}$$

definiert für alle (y, z) mit $\Lambda(0, 0) = 0$, $\Theta(0, 0) = 0$, Λ und Θ beschränkt und $R_0T = UR_0$. Außerdem ist R_0 eine Bijektion.

Beweis Nach dem Hilfssatz 7 hat T eine inverse, definiert für alle (y_1, z_1), sagen wir

$$T^{-1}(y_1, z_1) = (B^{-1}y_1 + Y^1(y_1, z_1), C^{-1}z_1 + Z^1(y_1, z_1)). \tag{146}$$

$\|Y^1(y_1, z_1)\|$ und $\|Z^1(y_1, z_1)\|$ können durch b_1c_0 beschränkt werden. Die Gleichung $R_0T = UR_0$ ist mit den Gleichungen

$$By + Y + \Lambda(By + Y, Cz + Z) = B(y + \Lambda) + Y_1(y + \Lambda, z + \Theta), \tag{147}$$
$$Cz + Z + \Theta(By + Y, Cz + Z) = C(z + \Theta) + Z_1(y + \Lambda, z + \Theta) \tag{148}$$

äquivalent, wo (y, z) das Argument von Y, Z, Λ und Θ ist. Die erste dieser Gleichungen kann umgeschrieben werden als

$$y+\Lambda = B[B^{-1}y+Y^1+\Lambda(T^{-1})]+Y_1(B^{-1}y+Y^1+\Lambda(T^{-1}), C^{-1}z+Z^1+\Theta(T^{-1})). \tag{149}$$

Es handelt sich um die erste Komponente der Gleichung $R_0 = UR_0T^{-1}$. Deshalb ist die Gleichung $R_0T = UR_0$ mit folgenden Gleichungen äquivalent.

$$\Theta = C^{-1}[Z - Z_1(y + \Lambda, z + \Theta) + \Theta(By + Y, Cz + Z)], \tag{150}$$
$$\begin{aligned}\Lambda &= B[Y^1 + \Lambda(T^{-1})] \\ &+Y_1(B^{-1}y + Y^1 + \Lambda(T^{-1}), C^{-1}z + Z^1 + \Theta(T^{-1})).\end{aligned} \tag{151}$$

Die Existenz von R_0 wird dadurch gezeigt, dass die Existenz einer Lösung der letzten zwei Gleichungen mit einer Iteration gezeigt wird. Sei $\Lambda^0 = 0$, $\Theta^0 = 0$ und

$$\Theta^n = C^{-1}[Z - Z_1(y + \Lambda^{n-1}, z + \Theta^{n-1}) + \Theta^{n-1}(By + Y, Cz + Z)], \tag{152}$$
$$\begin{aligned}\Lambda^n &= B[Y^1 + \Lambda^{n-1}(T^{-1})] \\ &+Y_1(B^{-1}y + Y^1 + \Lambda^{n-1}(T^{-1}), C^{-1}z + Z^1 + \Theta^{n-1}(T^{-1}))\end{aligned} \tag{153}$$

für $n \geq 1$. Λ^n und Θ^n sind wohldefiniert und stetig für alle n. Sie sind auch beschränkt. Es ist klar dass Λ^0, Θ^0, Λ^1 und Θ^1 beschränkt sind und mit der Definition

$$r_n = |||\Lambda^n - \Lambda^{n-1}||| + |||\Theta^n - \Theta^{n-1}||| \tag{154}$$

wo $|||\ |||$ das Supremum von $\|\ \|$ bedeutet gilt

$$|||\Theta^n - \Theta^{n-1}||| \leq c[\theta_1 r_{n-1} + |||\Theta^{n-1} - \Theta^{n-2}|||] \tag{155}$$
$$|||\Lambda^n - \Lambda^{n-1}||| \leq [b|||\Lambda^{n-1} - \Lambda^{n-2}||| + \theta_1 r_{n-1}]. \tag{156}$$

Die Summe dieser zwei Ungleichungen gibt

$$r_n \leq [\theta_1(c + 1) + \max\{b, c\}]r_{n-1} = \theta r_{n-1}. \tag{157}$$

Daraus folgt, dass $r_n \le r_1\theta^{n-1}$ für alle $n \ge 1$. Deshalb konvergieren die Folgen Λ^n und Θ^n gleichmässig gegen Grenzwerte Λ und Θ, die stetig und beschränkt sind. Diese Größen erfüllen die Funktionalgleichungen. Die Eindeutigkeit kann mit der üblichen Methode bewiesen werden. Um den Beweis abzuschließen bleibt zu zeigen, dass R_0 bijektiv ist. Wir bezeichnen die eindeutige Lösung R_0 der Gleichung $R_0T = UR_0$ mit R_{TU}, so dass $R_{TU}T = UR_{TU}$. Wenn wir die Rollen von T und U vertauschen sehen wir, dass es eine eindeutige Lösung R_{UT} gibt der Geichung $R_{UT}U = TR_{UT}$. Es folgt, dass

$$R_{TU}R_{UT}U = R_{TU}TR_{UT} = UR_{TU}R_{UT}, \tag{158}$$

$$TR_{UT}R_{TU} = R_{UT}UR_{TU} = R_{UT}R_{TU}T. \tag{159}$$

Durch Eindeutigkeit ist $R_{TU}R_{UT} = R_{UU} = I$ und $R_{UT}R_{TU} = R_{TT} = I$. Deshalb sind R_{TU} und R_{UT} bijektiv.

Beweis des Satzes von Grobman-Hartman Es wird angenommen, dass A hat $d > 0$ Eigenwerte mit positivem Realteil und $e > 0$ Eigenwerte mit positivem Realteil. Der allgemeine Fall ist dann leicht zu erhalten, in dem man künstliche Zusatzkomponenten hinzufügt. Zunächst werden alle Größen normiert wie im Hilfssatz 6. Sei R_0 die Abbildung, die durch den Hilfssatz 6 für T^1 geliefert wird, so dass $R_0T^1R_0^{-1} = L$. Hier ist L die Abbildung, die durch den Fluss für $t = 1$ des linearisierten Systems gegeben wird, d.h. $L = e^A$. Sei

$$R = \int_0^1 L^{-s}R_0T^s ds. \tag{160}$$

Dann ist

$$L^tR = \left(\int_0^1 L^{t-s}R_0T^{s-t}ds\right)T^t. \tag{161}$$

In dem man $s - t$ als neue Integrationsvariable einführt wird das letzte Integral zu

$$\int_{-t}^0 L^{-s}R_0T^s ds + \int_0^{1-t} L^{-s}R_0T^s ds. \tag{162}$$

Im ersten dieser Integrale kann man die Beziehung $L^{-s}R_0T^s = L^{-1-s}R_0T^{s+1}$ nutzen. Deshalb ist

$$L^tR = \left(\int_0^1 L^{-s}R_0T^s ds\right)T^t = RT^t. \tag{163}$$

Um den Beweis abzuschließen reicht es zu zeigen, dass $R = R_0$. Dies folgt aus dem Hilfssatz 8 mit $U = L$.

Die Schwierigkeiten die es gibt, wenn man im Satz von Grobman und Hartman die stetige Abbildung durch eine Abbildung höherer Differenzierbarkeit ersetzen will haben mit dem Phänomen der Resonanzen zu tun. Um diesen Punkt zu illustrieren betrachten wir das einfache System

$$\dot{x} = -x, \tag{164}$$

$$\dot{y} = -2y + x^2. \tag{165}$$

Die allgemeine Lösung dieses Systems ist

$$x(t) = ae^{-t}, \tag{166}$$
$$y(t) = a^2 te^{-2t} + be^{-2t}. \tag{167}$$

Wenn es einen C^2-Diffeomorphismus gäbe, der die Lösung des linearisierten Systems in die Lösung des nichtlinearen Systems transformieren würde, dann wäre die Lösung des nichtlinearen Systems von der Form

$$x(t) = ae^{-t} + be^{-2t} + o(e^{-2t}), \tag{168}$$
$$y(t) = ce^{-2t} + o(e^{-2t}). \tag{169}$$

Dies ist aber nicht der Fall. Das Problem hier ist, dass der Exponent der auf der rechten Seite der zweiten Gleichung entsteht, wenn der Ausdruck für x eingesetzt wird gleich dem Koeffizienten -2 auf der linken Seite ist. Im allgemeinen kann es Probleme geben, wenn ein Eigenwert als Linearkombination mit ganzzahligen Koeffizienten von anderen geschrieben werden kann. Es gibt einen Satz von Sternberg der besagt, dass wenn die Koeffizienten des Systems C^∞ sind und es keine Resonanzen gibt die Abbildung im Satz von Grobman und Hartman auch C^∞ gewählt werden kann. Eine entsprechende Aussage im analytischen Fall (mit gewissen zusätzlichen Annahmen) wurde schon 1879 von Poincaré bewiesen.

Sei x_0 eine hyperbolische stationäre Lösung, wo m_+ Eigenwerte der Linearisierung A positiven Realteil haben und m_- negativen Realteil. Das System $\dot{x} = Ax$ ist dann das Modell für das nichtlineare System, wenn wir topologische äquivalente Systeme betrachten. Man kann aber noch fragen, wann lineare Systeme einander topologisch äquivalent sind. Es stellt sich heraus, dass dies der Fall ist, wenn sie die gleichen Werte von m_+ und m_- haben. Wir können also als Modell den *Standardsattel* nehmen, der definiert ist durch $\dot{x} = x$, $\dot{y} = -y$ für $x \in \mathbb{R}^{m_+}$ und $y \in \mathbb{R}^{m_-}$. Diese Aussage wird hier nicht bewiesen, aber man kann die zentrale Idee verstehen im Fall der 2×2-Matrix $-I$ und einer Matrix mit Eigenwerten $-1 \pm i$. In einem Fall sind die Lösungskurven radial und im anderen Fall sind sie spiralförmig. Es ist nicht schwer zu sehen, dass die Spiralen durch einen Homöomorphismus der verschiedenen Kreise mit Mittelpunkt im Ursprung verschieden rotiert gerade gedreht werden können. Weitere Informationen zu diesem Thema findet man in [7], Kapitel 2.

Mit dem Satz von Grobman und Hartman und das Ergebnis über lineare System, das gerade erwähnt wurde kann man schließen, dass in der Nähe einer hyperbolischen stationären Lösung ein dynamisches System immer mit einem Standardsattel topologisch äquivalent ist. Der Fall, wo alle Eigenwerte positiven Realteil haben heißt hyperbolische Quelle und der Fall wo alle Eigenwerte negativen Realteil haben heißt hyperbolische Senke. Eine Quelle x_0 kann nie ω-Limespunkt einer anderen Lösung sein. Jede Lösung, die nahe genug bei x_0 startet konvergiert gegen x_0 für $t \to -\infty$. Es gibt entsprechende Aussagen für eine Senke. Sei jetzt x_0 eine hyperbolische stationäre Lösung, die weder Quelle noch Senke ist. Wenn x_0 in der ω-Limesmenge einer Lösung $x(t)$, dann hat $x(t)$

auch ω-Limespunkte auf den stabilen und instabilen Mannigfaltigkeiten von x_0. Für eine allgemeine stationäre Lösung haben wir

Satz (Shoshitaishvili) Sei x_0 eine stationäre Lösung eines dynamischen Systems. Dann ist das System in der Nähe von x_0 topologisch äquivalent mit dem Produkt der Einschränkung des Systems auf eine Zentrumsmannigfaltigkeit von x_0 mit einem Standardsattel.

Es folgt insbesondere, dass die Einschränkungen des Systems auf zwei verschiedene Zentrumsmannigfaltigkeiten von x_0 topologisch äquivalent sind. Deshalb ist die fehlende Eindeutigkeit der Zentrumsmannigfaltigkeit kein Problem. Mit diesem Satz kann man die Aussagen über asymptotische Stabilität beweisen, die wir bei der Betrachtung der Beispiele für Zentrumsmannigfaltigkeiten erwähnt haben. Es gibt einen Satz von Takens, der eine gemeinsame Verallgemeinerung der Sätze von Sternberg und Shoshitaishvili ist. Wenn bei einer stationären Lösung die nicht hyperbolisch sein muss es in einem geeigneten Sinne keine Resonanzen gibt, gilt ein Analogon des Satzes von Sternberg.

10 Theorie von Poincaré-Bendixson

Nachdem wir lange das lokale Verhalten von Lösungen in der Nähe einer stationären Lösung besprochen haben, wenden wir uns jetzt globalen Eigenschaften zu. Eindimensionale dynamische Systeme sind leicht zu analysieren. Systeme deren Dimension mindestens drei ist können große Schwierigkeiten bergen. Typische Themen sind Chaos und seltsame Attraktoren. Dazwischen gibt es die Dimension zwei, die auf Grund der Theorie von Poincaré-Bendixson relativ gut beherrschbar ist. Diese Theorie ist das Hauptthema dieses Abschnitts aber bevor wir dazu kommen machen wir ein paar Bemerkungen über den eindimensionalen Fall. Dabei geht es um das System $\dot{x} = f(x)$, wo x eine skalare Größe ist. Die stationären Lösungen sind die Nullstellen von f. Die Menge auf der f ungleich Null ist, ist eine Vereinigung von offenen Intervallen U_i. Auf U_i mit i fest ist jede Lösung strikt monoton. Sei U_i das Intervall (x_-, x_+), wo die Endpunkte unendlich sein dürfen. In jeder Zeitrichtung muss die Lösung gegen einen Grenzwert streben (endlich oder unendlich) und dieser Grenzwert kann nur ein Endpunkt des Intervalls sein. Nehmen wir an, dass die Lösung monoton steigend ist. Dann gibt es nur drei Möglichkeiten für die Asymptotik in der Zukunft. Die Lösung strebt nach endlicher Zeit gegen unendlich, die Lösung existiert global in der Zukunft und strebt gegen unendlich für $t \to \infty$ oder die Lösung existiert global und strebt gegen $x_+ < \infty$ für $t \to \infty$. Für monoton fallende Lösungen gibt es entsprechende Möglichkeiten. Die ω-Limesmenge einer beschränkten Lösung ist immer eine stationäre Lösung. Nehmen wir an, dass ein System auf einem Intervall I definiert ist, dass es zwei stabile stationäre Lösungen x_1 und x_2 gibt mit $x_1 < x_2$ und dass f auf dem Intervall $[x_1, x_2]$ nicht identisch Null ist. Dann gibt es eine instabile stationäre Lösung x_3 mit $x_1 < x_3 < x_2$, wie man folgendermassen beweisen kann. Es gibt einen Punkt

$x_4 \in (x_1, x_2)$ mit $f(x_4) \neq 0$. Wir dürfen annehmen, dass $f(x_4) > 0$ weil wir sonst x durch $-x$ ersetzen könnten. Sei (x_5, x_6) das maximale offene Intervall um x_4 auf dem f positiv ist. Dann können wir $x_3 = x_5$ setzen.

Jetzt kommen wir zu zweidimensionalen Systemen. Hier ist ein zentrales Hilfssmittel der Jordansche Kurvensatz. Eine Jordankurve ist die Menge der Punkte x in der Ebene der Form $x = x(t)$, $a \leq x \leq b$, wo $x(t)$ stetig ist, $x(a) = x(b)$ und $x(s) \neq x(t)$ für $a \leq s < t \leq b$.

Satz (Jordanscher Kurvensatz) Wenn J eine Jordankurve ist, dann ist ihr Komplement in der Ebene die Vereinigung zweier disjunkter offener Mengen E_1 und E_2, mit $\partial E_1 = \partial E_2 = J$. Eines dieser Gebiete ist beschränkt, heißt Inneres von J und ist einfach zusammenhängend.

Ein topologischer Raum X heißt einfach zusammenhängend wenn es zu jeder stetigen Abbildung $\gamma : S^1 \to X$ eine stetige Abbildung $H : [0,1] \times S^1 \to X$ gibt mit $H(0,x) = \gamma(x)$ und $H(1,x) = \gamma(0)$ für alle $x \in S^1$. Anschaulich bedeutet dies, dass jede geschlossene Kurve stetig zu einem Punkt deformiert werden kann.

Betrachten wir eine stetige Abbildung $x : [a,b] \to \mathbb{R}^2$ mit Bild J. Sei $\eta(t)$ eine stetige Abbildung von $[a,b]$ nach $\mathbb{R}^2 \setminus \{0\}$. Intuitiv handelt es sich um ein stetiges Vektorfeld auf J, das nie verschwindet. Für $x \in \mathbb{R}^2 \setminus \{0\}$ sei $\pi(x) = x/\|x\|$. Dadurch wird eine Abbildung $\pi : \mathbb{R}^2 \setminus \{0\} \to S^1$ definiert. Sei $\phi(t)$ der Winkel von der positiven x_1-Richtung zu $\eta(t)$. Dann ist $\cos\phi = \eta_1/\|\eta\|$ und $\sin\phi = \eta_2/\|\eta\|$. Durch diese Formeln ist ϕ bis auf eine ganze Zahl mal 2π bestimmt. Wenn verlangt wird, dass ϕ stetig ist und der Wert in einem Punkt, z. B. a festgelegt wird, dann ist ϕ eindeutig bestimmt. Anders gesagt, ist $\pi \circ \eta$ eine Abbildung von $[a,b]$ nach S^1. Die Vorschrift $\phi \mapsto (\cos\phi, \sin\phi)$ definiert eine stetige Abbildung p von $\mathbb{R}$ nach S^1. Wir suchen also eine Abbildung $\tilde{\eta}$, so dass $p \circ \tilde{\eta} = \pi \circ \eta$. Eine solche Abbildung existiert und ist eindeutig bis auf eine additive Konstante. Sei $j_\eta(J)$ durch $2\pi j_\eta(J) = \phi(b) - \phi(a)$ definiert. Wenn J aus zwei Kurven J_1 und J_2 zusammengesetzt ist, dann ist $j_\eta(J) = j_\eta(J_1) + j_\eta(J_2)$. Wir interessieren und für diese Definition in dem Fall, dass J eine Jordan-Kurve ist. Es werden hier nur solche Jordan-Kurven betrachtet, die stückweise C^1 sind und es wird angenommen, dass sie positiv orientiert sind in dem Sinne, dass $(-dx_2/dt, dx_1/dt)$ immer ins Innere von J weist. Es ist klar, dass $j_\eta(J)$ eine ganze Zahl ist. Sie heißt Index von J.

Satz (Umlaufsatz) Sei J eine positiv orientierte Jordan-Kurve der Klasse C^1 der auf $[0,1]$ definiert ist und $\eta(t)$ das entsprechende Tangentenvektorfeld. Dann ist $j_\eta(J) = 1$.

Beweis Auf dem Dreieck Δ das durch $0 \leq s \leq t \leq 1$ gegeben wird soll eine stetige Funktion η mit Werten in S^1 definiert werden. Wenn $s \neq t$ und $(s,t) \neq (0,1)$ dann sei $\eta(s,t) = [x(t) - x(s)]/\|x(t) - x(s)\|$. Diese Funktion hat eine eindeutige stetige Fortsetzung auf Δ. $\eta(t,t) = \dot{x}(t)/\|\dot{x}(t)\|$ und $\eta(0,1) = -\eta(0,0)$. $\eta(0,t) = -\eta(t,1)$. Nehmen wir an, dass $x(0)$ so gewählt wird, dass die Tangente in diesem Punkt parallel zur x_1-Achse ist und kein

Punkt der Kurve unterhalb der Tangente liegt. Es gibt eine eindeutige stetige Funktion $\tilde{\eta} : \Delta \to \mathbb{R}$ mit $p \circ \tilde{\eta} = \eta$ und $\tilde{\eta}(0,0) = 0$. Wir schreiben auch ϕ für $\tilde{\eta}$. Dann ist $2\pi j_\eta(J) = \phi(1,1) - \phi(0,0)$, was durch die Betrachtung von $\phi(t,t)$ folgt. Jetzt ist $0 \leq \phi(0,t) \leq \pi$ und $\phi(0,1) = k\pi$ wobei k eine ungerade ganze Zahl ist. Deshalb ist $\phi(0,1) = \pi$. Die Zahl $\phi(s,1)$ ist immer zwischen π und 2π und $\phi(1,1) = k\pi$ wobei k eine gerade ganze Zahl ist. Deshalb ist $\phi(1,1) = 2\pi$. Da $2\pi j_\eta(J) = \phi(1,1) - \phi(0,0)$ ist der Satz damit bewiesen.

Der Index ist invariant unter Deformationen des Vektorfeldes

Hilfssatz 9 Sei J eine Jordan-Kurve und $\xi(t)$ und $\eta(t)$ zwei Vektorfelder auf J die ineinander deformiert werden können ohne zu verschwinden. Dann gilt $j_\xi(J) = j_\eta(J)$.

Zu sagen dass man das Vektorfeld deformieren kann bedeutet, dass es ein stetiges Vektorfeld $\eta(s,t)$ gibt für $a \leq t \leq b$ und $0 \leq s \leq 1$ mit $\eta(t,0) = \xi(t)$, $\eta(t,1) = \eta(t)$, $\eta(a,s) = \eta(b,s)$ und $\eta(t,s) \neq 0$.

Beweis Sei $j(s)$ der Index von $\eta(t,s)$ mit s fest. Dann ist $j(s)$ eine stetige Funktion von s. Da aber $j(s)$ auch eine ganze Zahl ist muss sie konstant sein. Insbesondere ist $j(0) = j(1)$.

Als nächstes definieren wir den Index einer stationären Lösung. Sei J eine positiv orientierte Jordankurve auf dem ein Vektorfeld f nie verschwindet. Dann heißt $j_f(J)$ der Index von f bezüglich J, wobei $j_f(J) = j_\eta(J)$ und $\eta(t) = f(x(t))$. Wie im Hilfssatz 9 kann man zeigen, dass wenn J_0 und J_1 zwei Jordankurven sind die ineinander deformiert werden können ohne eine stationäre Lösung zu treffen, dann gilt $j_f(J_0) = j_f(J_1)$. Wir betrachten jetzt ein dynamisches System, das auf einem Gebiet G definiert ist. Sei J eine Jordankurve in G, so dass das Innere von J auch in G liegt und dass das Vektorfeld auf und im Inneren von J nicht verschwindet. Dann ist $j_f(J) = 0$. Da das Innere von J einfach zusammenhängend ist kann J zu einer Kurve J_1 deformiert werden, der ein kleiner Kreis in der Nähe eines Punktes x_0 ist. Da $f(x_0) \neq 0$ ist der Winkel zwischen $f(x)$ und der positiven x_1-Richtung fast konstant. Da $j_f(J)$ eine ganze Zahl ist kann sie nur Null sein. Für einen Punkt x_0 ist der Index gleich für alle Jordankurven J mit der Eigenschaften, dass x_0 im Inneren von J liegt und es keine stationären Punkte im Inneren von J gibt außer eventuell x_0 selbst. Diese Zahl heißt dann der Index von x_0 bezüglich f. Wenn x_0 keine stationäre Lösung ist, dann ist dieser Index Null. Wenn es nur endlich viele stationäre Lösungen im Inneren von J gibt, was wiederum in G liegt dann ist $j_f(J) = j_f(x_1) + \ldots j_f(x_n)$. Wir geben keinen vollständigen Beweis dieser Aussage aber die grundlegende intuitive Idee des Beweises ist leicht zu verstehen. Mann deformiert die Kurve in eine Kurve, die Folgende Eigenschaften hat. Sie Umrundet fast einen stationären Punkt auf einem kleinen Kreis und bewegt sich dann zu einem Kreis um einen anderen stationären Punkt. Auf diese Weise wird jeder stationäre Punkt einmal besucht. Danach werden sie in der umgekehrten Reihenfolge besucht, wobei der Rückweg zwischen zwei Punkten immer nahe beim Hinweg liegt. Schließlich ist die Kurve wieder nahe beim Ausgangspunkt. Die Summanden in der Formel liefern die Kreise. Die Beiträge von dem Hinweg von einem Punkt zum nächsten und dem entsprechenden Rückweg sind fast entgegengesetzt.

Satz 10 Sei f eine stetig differenzierbare Funktion auf einer offenen Menge G und sei $x(t)$ eine periodische Lösung der Gleichung $\dot{x} = f(x)$ mit Periode p. Wenn $x(t)$, $0 \leq t \leq p$ eine Jordan-Kurve ist deren Inneres I in G enthalten ist dann enthält I eine stationäre Lösung.
Beweis Wenn die Jordan-Kurve positiv orientiert ist, ist nach dem Umlaufsatz $j_f(J) = 1 \neq 0$. Es kann also nicht sein, dass es keine stationären Lösungen in I gibt.

Jetzt kommen wir zum Theorem von Poincaré und Bendixson.

Satz 11 (Poincaré-Bendixson) Sei f eine stetig differenzierbare Funktion auf einer offenen Teilmenge G des $\mathbb{R}^2$ und sei $x(t)$ eine Lösung von $\dot{x} = f(x)$ für $t \geq 0$, die in einer kompakten Teilmenge von G enthalten ist und die nicht periodisch ist. Wenn es keine stationären Lösungen in der ω-Limesmenge von $x(t)$ gibt, dann ist die ω-Limesmenge das Bild einer periodischen Lösung $y(t)$.

Aus dem Beweis von Satz 11 wird folgen, dass wenn die Voraussetzungen des Satzes gelten es eine monoton steigende Folge $\{t_n\}$ gibt mit den Eigenschaften, dass $x(t + t_n) \to y(t)$ für $n \to \infty$, gleichmässig auf $[0, p]$ und $t_{n+1} - t_n \to p$ für $n \to \infty$. Hier ist p die minimale Periode von $y(t)$.

Beweis von Satz 11 Eine beschränkte und abgeschlossene Strecke L heißt transversal zur Gleichung $\dot{x} = f(x)$ wenn $f(x) \neq 0$ für alle $x \in L$ und die Richtung von $f(x)$ für keinen Punkt $x \in L$ parallel zu L ist. Dann überquert die Lösung L immer in die gleiche Richtung. Der Beweis wird in fünf Schritte (a)-(e) aufgeteilt.

(a) Sei $x_0 \in G$, $f(x_0) \neq 0$ und L eine Strecke durch x_0, die transversal zu f ist. Es folgt aus dem Satz über lokale Existenz, dass es eine Umgebung G_0 von x_0 und $\epsilon > 0$ gibt so dass für $x_1 \in G_0$ die Lösung mit $x(0) = x_1$ für $|t| \leq \epsilon$ existiert und L nur einmal trifft. Es ist nämlich so, dass für $\delta > 0$ beliebig die Umgebung G_0 und die Zahl ϵ so gewählt werden können, dass der Unterschied zwischen $x(t)$ und $x_1 + tf(x_1)$ nicht größer ist als $\delta|t|$ ist für $|t| \leq \epsilon$. Insbesondere folgt, dass $x(t)$ die Strecke L nur endlich oft treffen kann für t in einem beschränkten Intervall.

(b) Sei L eine Strecke, die transversal zu f ist und den Punkt x_0 enthält. O.B.d.A. können wir annehmen, dass L eine Teilmenge der x_2-Achse ist. Nehmen wir an, dass $x(t)$ die Strecke L trifft für Werte $t_1 < t_2 \ldots$ von t. Dann ist $x_2(t_n)$ eine monotone Funktion von n. Um dies zu sehen, nehmen wir an, dass x_1 bei Überquerungen von L zunimmt. Betrachten wir o.B.d.A. den Fall, dass $x_2(t_1) < x_2(t_2)$. Die Menge die aus der Kurve $y(t)$, $t_1 \leq t \leq t_2$, und der Strecke auf der x_2-Achse mit $x_2(t_1) \leq x_2 \leq x_2(t_2)$ besteht ist eine Jordan-Kurve J.

Für $t > t_2$ ist $x(t)$ immer im Inneren von J oder niemals im Inneren von J. Diese Aussage folgt weil die Lösung L immer nur in eine Richtung überqueren kann. Es ist dann klar, dass $x_2(t_3) > x_2(t_2)$ und das Argument kann wiederholt werden. Die Folge $\{x_2(t_n)\}$ ist monoton steigend.

(c) Es wird jetzt gezeigt, dass die ω-Limesmenge von $x(t)$ höchstens einen Punkt von L enthält. Wenn y ein solcher Punkt ist, dann trifft $x(t)$, nach (a), die Strecke L unendlich oft. Nach (b) konvergieren die Schnittpunkte mit L monoton nach x_0.

(d) Die ω-Limesmenge von $x(t)$ ist nicht leer. Sei y_0 ein Punkt dieser Menge. Die Lösung $y(t)$ mit $y(0) = y_0$ ist in der ω-Limesmenge von $x(t)$ enthalten. Die ω-Limesmenge von $y(t)$ ist auch in der ω-Limesmenge von $x(t)$ enthalten und ist nicht leer. Sei z_0 ein Punkt dieser Menge. Es folgt aus den Annahmen des Satzes, dass z_0 keine stationäre Lösung ist. Es gibt also eine Strecke L_0 durch z_0, die transversal zu f ist. $y(t)$ überquert L_0 unendlich oft. z_0 und jeder Punkt der Überquerung liegen in der ω-Limesmenge von $x(t)$. Es folgt aus (c) dass diese Punkte übereinstimmen. Es existieren also $t_1 < t_2$ gibt mit $y(t_1) = y(t_2) = z_0$. Deshalb hat $\dot{x} = f(x)$ eine periodische Lösung der Periode $p = t_2 - t_1$. Es kann angenommen werden, dass p die minimale Periode ist.

(e) Jetzt soll gezeigt werden, dass die ω-Limesmenge von $x(t)$ mit ihrer Teilmenge, dem Bild von $y(t)$ übereinstimmt. Wenn diese Aussage falsch wäre, dann wäre das Komplement Z des Bildes von $y(t)$ in der ω-Limesmenge von $x(t)$ nicht leer. Das Bild von $y(t)$ muss auch einen Häufungspunkt x_1 von Z enthalten, da die ω-Limesmenge von $x(t)$ zusammenhängend ist. Sei L_1 eine Strecke durch x_1, die transversal zu f ist. Jede kleine Kugel um x_1 enhält einen Punkt $x_2 \in Z$. Sei $w(t)$ die Lösung von $\dot{x} = f(x)$ mit $w(0) = x_2$. Das Bild von w ist in der ω-Limesmenge von $x(t)$ enthalten. Wenn x_2 hinreichend nahe bei x_1 liegt, dann überquert $w(t)$ die Strecke L_1. Diese Überquerung kann nur bei x_1 passieren, als Folge von (c). Da x_2 nicht im Bild von $y(t)$ ist bekommen wir einen Widerspruch.

Jetzt wird die Behauptung nach dem Satz bewiesen. Sei $y(t)$ die periodische lösung mit $y(0) = y_0$ und sei L_0 eine Strecke durch y_0, die transversal zu f ist. Seien $t_1 < t_2 < \ldots$ die aufeinanderfolgenden Überquerungen von L_0 durch die Lösung $x(t)$. Dann konvergiert $x(t_n)$ monoton längs L_0 gegen y_0. Durch stetige Abhängigkeit der Lösungen von den Anfangsdaten konvergiert $x(t + t_n)$ gegen $y(t)$, gleichmässig auf $[0, p]$. Für $n \to \infty$ konvergiert $x(t_n + p)$ gegen $y(p) = y_0$. Deshalb ist es so, dass für $\epsilon > 0$ und n groß $x(t)$ im Intervall $[t_n + p - \epsilon, t_n + p + \epsilon]$ die Strecke L_0 überquert. Deshalb ist $t_{n+1} \leq t_n + p + \epsilon$. Außerdem ist $|x(t_n + t) - y(t)|$ klein für n groß und $0 < \epsilon \leq p$. Deshalb existiert $\delta > 0$ mit der Eigenschaft, dass $\|x(t_n + t) - y_0\| \geq \delta$ für $0 < \epsilon \leq t \leq p - \epsilon$. Insbesondere gibt es keine Überquerung von L_0 für $\epsilon \leq t \leq p - \epsilon$. Deshalb ist $t_{n+1} \geq t_n + p - \epsilon$ für n groß und die Behauptung ist bewiesen.

Satz 12 Seien f und $x(t)$ wie in Satz 11 bis auf die Tatsache, dass es eine endliche Anzahl n von stationären Lösungen in der ω-Limesmenge von $x(t)$ gibt. Wenn $n = 0$ lässt sich Satz 11 anwenden. Wenn $n = 1$ und die ω-Limesmenge von $x(t)$ ein Punkt ist, dann konvergiert die Lösung gegen diesen Punkt für $t \to \infty$. Wenn $n \geq 1$ und die ω-Limesmenge von $x(t)$ mehr als einen Punkt enthält, dann besteht diese Menge aus stationären Lösungen $x_1, \ldots, x_n$ und eine endliche oder unendliche, aber abzählbare Menge von Lösungen $y(t)$ auf $\mathbb{R}$ mit folgender Eigenschaft. Die Lösung $y(t)$ besitzt Grenzwerte für $t \to +\infty$ und $t \to -\infty$ und diese Grenzwerte gehören zu den Punkten x_i.

Beweis Wir betrachten den Fall, dass $n \geq 1$ und die ω-Limesmenge von $x(t)$ mehr als einen Punkt enthält. Da die ω-Limesmenge zusammenhängend ist enthält sie einen Punkt y_0, der keine stationäre Lösung ist. Die Lösung $y(t)$ mit $y(0) = y_0$ liegt in der ω-Limesmenge von $x(t)$. Betrachten wird den Fall, dass die ω-Limesmenge von $y(t)$ einen Punkt z_0 enthält, der keine stationäre Lösung ist. Dann folgt aus dem Teil (d) des Beweises von Satz 11, dass die Lösung $y(t)$ periodisch ist. Außerdem gibt es eine Umgebung von $y(t)$ die keine anderen ω-Limespunkte von $x(t)$ enthält. Da die ω-Limesmenge von $x(t)$ zusammenhängend ist, müsste die Menge nur aus der periodischen Lösung bestehen, ein Widerspruch. Deshalb ist die ω-Limesmenge von $y(t)$ einer der Punkte x_i. Das Argument für α-Limespunkte geht genauso. Es bleibt zu zeigen, dass die Menge der Lösungen $y(t)$ abzählbar ist. Nehmen wir an die Menge wäre überabzählbar. Dann gäbe es Punkte x_i und x_j, nicht notwendigerweise verschieden, die durch eine überabzählbare Menge von Lösungen verbunden werden. Jede dieser Kurven bzw. jedes paar dieser Kurven bildet eine Jordan-Kurve. Wenn eine Menge dieser Kurven so ist, dass für jedes paar das Innere einer Kurve das Innere der anderen nicht überschneidet, dann muss diese Menge abzählbar sein. Deshalb muss es zwei der Jordan-Kurven geben, deren Innere sich überschneiden. Dann muss eine, sagen wir J_1, in der abgeschlossenen Hülle des Inneren der anderen Kurve, sagen wir J_2, liegen. Das Innere von J_1 muss auch das Innere einer Kurve J_3 überschneiden, die unterschiedlich ist zu J_2. Das Bild der Lösung $x(t)$ liegt zwischen J_1 und J_2. Deshalb kann es nicht sein, dass J_3 in der abgeschlossenen Hülle des Inneren von J_1 liegt. Es kann auch nicht sein, dass J_3 im Komplement der abgeschlossenen Hülle des Inneren von J_2 liegt. Deshalb muss J_3 zwischen J_1 und J_2 liegen und das Bild von $x(t)$ zwischen einer dieser zwei Kurven und J_3 liegen. Aber dann kann die andere der beiden nicht in der ω-Limesmenge von $x(t)$ liegen, ein Widerspruch.

Der Satz von Poincaré und Bendixson kann manchmal benutzt werden, um die Existenz von periodischen Lösungen zu beweisen. Es gibt ein einfaches Kriterium, das oft benutzt werden kann, um die Existenz von periodischen Lösungen von zweidimensionalen dynamischen Systemen auszuschließen. Sei $\dot{x} = f(x)$ ein zweidimensionales dynamisches System und sei g eine reellwertige Funktion. Wenn $\mathrm{div}(gf) \geq 0$ und $\mathrm{div}(gf)$ nicht identisch verschwindet heißt g Dulac-Funktion. Wenn eine Dulac-Funktion existiert, dann hat das System keine periodische Lösung, deren Inneres im Definitionsbereich von f liegt. In dem Fall ist das Integral der Komponente von gf in der Normalenrichung um die geschlossene Integralkurve nach dem Satz von Stokes gleich dem Integral über dem Inneren der Kurve von $\mathrm{div}(gf)$. Da das erste Integral verschwindet und das zweite Integral strikt positiv ist bekommt man einen Widerspruch.

11 Oszillatoren

Wenn ein dynamisches System eine periodische Lösung besitzt, wird ein Phänomen, das durch diese Lösung modelliert wird anhaltende Oszillationen aufweisen. Eine solche Situation wird oft als Oszillator bezeichnet. Damit das Verhalten tatsächlich in der Wirklichkeit beobachtet werden kann, sollte sie eine gewisse Stabilität besitzen. Die Definitionen von Stabilität sind ähnlich wie im Fall einer stationären Lösung. Sei $x(t)$ eine periodische Lösung mit Bild γ. Die Lösung heißt orbital stabil, wenn es zu jeder Umgebung U von γ eine Umgebung V von γ gibt, so dass jede Lösung die in V startet für immer in U bleibt. Die Lösung heißt orbital asymptotisch stabil wenn sie orbital stabil ist und wenn es eine Umgebung U von γ gibt mit der Eigenschaft, dass für jede Lösung die in U startet der Abstand von $x(t)$ zu γ gegen Null strebt für $t \to \infty$.

Zunächst sollen einige bekannte Beispiele kurz besprochen werden. In den 1920er Jahren untersuchte Balthasar van der Pol einen elektrischen Schaltkreis mit nichtlinearer Dämpfung der zu Oszillationen führt. Ein dynamisches System, das diesen Schaltkreis beschreibt wird van der Pol-Oszillator genannt. Die ursprüngliche Gleichung ist eine skalare Gleichung zweiter Ordnung. Durch Einführung neuer Variablen bekommt man ein System von zwei Gleichungen erster Ordnung, das zur Klasse der Liénard-Systeme gehört. In diesem System gibt es eine stabile periodische Lösung. Ihr Entdecker hat dieses Phänomen als Relaxations-Oszillation bezeichnet. Auf diesen Begriff kommen wir später zurück. Die Übertragung elektrischer Signale durch Nervenzellen kann durch ein vierdimensionales dynamisches System beschrieben werden. Dieses System spielte eine zentrale Rolle beim Verständnis dieses biologischen Phänomens. Für Ihre Arbeit zu diesem Thema bekamen Hodgkin und Huxley den Nobelpreis für Medizin. Eine vereinfachte Version dieses Modells, die zweidimensional ist und das noch wesentliche qualitative Eigenschaften des vollen Systems behält ist das FitzHugh-Nagumo-Modell. Das FitzHugh-Nagumo-Modell ist eng mit dem van der Pol-Oszillator verwandt und besitzt auch eine stabile periodische Lösung. Man hat lange geglaubt, dass eine chemische Reaktion immer gegen ein Gleichgewicht strebt, so dass anhaltende Oszillationen in chemischen Reaktionen nicht möglich wären. Später hat man diese Vorstellung durch Experimente widerlegt. Es handelt sich um die sogenannte Belousov-Zhabotinski-Reaktion. Die tatsächliche Reaktion ist sehr kompliziert, aber man kann vereinfachte mathematische Modelle dafür konstruieren. Ein bekanntes Beispiel ist das Field-Noyes-Modell, auch Oregonator genannt. Dieses dynamische System ist dreidimensional, so dass die Poincaré-Bendixson-Theorie nicht darauf angewendet werden kann. Eine weitere Vereinfachung ergibt den zweidimensionalen Brüsselator.

Jetzt wird eine konkretes Beispiel untersucht, das aus der Biologie kommt. In unseren Körpern wird Energie freigestezt in dem Zuckermoleküle chemisch verarbeitet werden. Dieser Prozess heißt Glykolyse. Der Mechanismus kann am besten in einfachen Organismen untersucht werden, zum Beipiel in der Bäckerhefe, *Saccharomyces cerevisiae*. Dieser Einzeller gewinnt Energie aus Zucker und produziert dabei Ethanol, was für die Produktion von alkoholischen Getränken verwendet wird. Ein einfaches Experiment sieht folgendermassen aus. Wir haben Hefezellen in einer Lösung und führen Glucose hinzu mit einer Konstanten Rate k_0. Wenn k_0 klein genug ist, wird Ethanol auch mit konstanter Rate produziert. Wenn aber k_0 erhöht wird fängt die Ethanolproduktion an zu oszillieren. Dieses Phänomen wurde von Higgins und Selkov untersucht. Diese Autoren haben zur Beschreibung dieses Experiments ein zweidimensionales dynamisches System eingeführt, das heute unter dem Namen Higgins-Selkov-Oszillator bekannt ist. Wenn man die Zellen aufbricht und die Inhalte extrahiert, dann sieht man die Oszillationen immer noch. Dadurch wird nahegelegt, dass es sich um ein rein chemisches Phänomen handelt, das unabhängig von den komplizierten Strukturen in der Zelle ist.

Nehmen wir an, wir haben eine chemische Reaktion mit einem Substrat S und einem Produkt P. Die Konzentration des Substrats erfüllt die Geichung $\dot{S} = k_0 - k_1 S P^2$. Das Substrat wird mit der konstanten Rate k_0 bereitgestellt und mit einer Rate aufgebraucht, die mit der Konzentration des Produkts zunimmt. Bei der Glykolyse geht es um das Enzym Phosphofructokinase (PFK), das Fructose-6-Phosphat in Fructose-1,6-Bisphosphat verwandelt, wobei ATP zu ADP wird. ATP wird als reichlich vorhanden betrachtet und deshalb nicht im Modell berücksichtigt. ADP erhöht die Aktivität von PFK und führt dadurch zu einer positiven Rückkopplung. Im Modell spielt S die Rolle der Konzentration von Glucose und P die der Konzentration von ADP. Die Gleichung für P ist $\dot{P} = k_1 S P^2 - k_2 P$. Die Konstanten k_i sind alle positiv. Auf Grund ihrer Interpretationen sollen die Größen S und P positiv sein. Wenn sie positiv starten bleiben sie positiv. Der Beweis ist wie beim Hilfssatz 1. Die Summe der Gleichungen liefert $(S+P)\dot{} = k_0 - k_2 P$. Deshalb ist eine Lösung auf jedem endlichen Intervall der Form $[0, t_1)$ beschränkt und die Lösungen existieren global in der Zukunft.

Der Higgins-Selkov-Oszillator hat eine eindeutige stationäre Lösung, die durch $P = \frac{k_0}{k_2}$ und $S = \frac{k_2^2}{k_0 k_1}$ gegeben wird. Die Linearisierung in diesem Punkt ist

$$\frac{d\hat{S}}{dt} = -\frac{k_0^2 k_1}{k_2^2} \hat{S} - 2k_2 \hat{P}, \tag{170}$$

$$\frac{d\hat{P}}{dt} = \frac{k_0^2 k_1}{k_2^2} \hat{S} + k_2 \hat{P}. \tag{171}$$

Die Determinante der Jacobi-Matrix ist $\frac{k_0^2 k_1}{k_2} > 0$ und die Spur ist $-\frac{k_0^2 k_1}{k_2^2} + k_2$, ein Ausdruck der sein Vorzeichen ändert wenn $k_0^2 = \frac{k_2^3}{k_1}$. Wenn beide Eigenwerte reell sind haben sie das gleiche Vorzeichen, und dieses Vorzeichen wird durch die Spur bestimmt. Wenn die Eigenwerte komplex sind haben sie den gleichen Realteil und das Vorzeichen des Realteils wird durch das Vorzeichen der Spur bestimmt. Wir sehen, dass die stationäre Lösung asymptotisch stabil ist wenn

$k_0^2 > \frac{k_2^3}{k_1}$ und instabil wenn $k_0^2 < \frac{k_2^3}{k_1}$.

Wenn wir zeigen könnten dass im Fall des Higgins-Selkov-Oszillators mit Parametern für die die stationäre Lösung instabil ist mindestens eine in der Zukunft beschränkte nicht-stationäre Lösung existiert dann könnten wir aus dem Satz von Poincaré-Bendixson schließen, dass es eine periodische Lösung gibt. Die Beschränktheit ist aber anscheinend schwer zu beweisen. In der Tat besitzt das System für solche Parameter eine periodische Lösung aber sie besitzt auch unbeschränkte Lösungen [12]. Aus diesem Grund betrachten wir jetzt ein anderes verwandtes Modell, das Schnakenberg-Modell [11], das einfacher zu analysieren ist. In diesem anderen Modell wird die Gleichung für P durch die Gleichung $\dot{P} = k_1SP^2 - k_2P + k_3$ ersetzt, wobei angenommen wird, dass $k_3 < k_0$. Die Lösungen existieren global in der Zukunft nach dem gleichen Argument wie im Fall des Higgins-Selkov-Oszillators. Im Fall des Schnakenberg-Modells gibt es eine eindeutige stationäre Lösung mit $P = \frac{k_0 - k_3}{k_2}$ und $S = \frac{k_0 k_2^2}{(k_0 - k_3)^2 k_1}$. Die Linearisierung in diesem Punkt ist

$$\frac{d\hat{S}}{dt} = -\frac{(k_0 - k_3)^2 k_1}{k_2^2}\hat{S} - 2\frac{k_0 k_2}{k_0 - k_3}\hat{P}, \tag{172}$$

$$\frac{d\hat{P}}{dt} = \frac{(k_0 - k_3)^2 k_1}{k_2^2}\hat{S} + \frac{(k_0 + k_3)k_2}{k_0 - k_3}\hat{P}. \tag{173}$$

Die Determinante der Jacobi-Matrix ist $\frac{(k_0 - k_3)^2 k_1}{k_2} > 0$ und die Spur ist von der Form $-\frac{(k_0 - k_3)^2 k_1}{k_2^2} + \frac{(k_0 + k_3)k_2}{k_0 - k_3}$. Wir bekommen ähnliche Aussagen über die Stabilität wie im Fall des Higgins-Selkov-Oszillators, wobei die Grenze zwischen dem stabilen und dem instabilen Fall für das Schnakenberg-Modell durch

$$\frac{(k_0 - k_3)^3}{k_0 + k_3} = \frac{k_2^3}{k_1} \tag{174}$$

gegeben wird.

Jetzt wird gezeigt, dass die Lösungen des Schnakenberg-Modells beschränkt sind. Daraus folgt, dass wenn die stationäre Lösung instabil ist die ω-Limesmenge jeder nicht-stationären Lösung eine periodische Lösung ist. Der erste Schritt ist zu zeigen, dass in jeder Lösung P für t hinreichend groß durch eine positive Konstante P_- nach unten beschränkt ist. Es gilt die Ungleichung $\dot{P} \geq k_3 - k_2P$. Durch Integration bekommt man

$$P(t) \geq \frac{k_3}{k_2} + \left(P(0) - \frac{k_3}{k_2}\right)e^{-k_2 t}. \tag{175}$$

Daraus folgt die Aussage für einen beliebigen Wert von $P_- < \frac{k_3}{k_2}$. Wir können dann die Ungleichung $\dot{S} \leq k_0 - k_1 P_-^2 S$ benutzen, um zu sehen dass S durch eine Konstante S_+ beschränkt ist. Es folgt, dass

$$\frac{d}{dt}(P + S) = -k_2P + k_0 + k_3 \leq -k_2(P + S) + k_2S_+ + k_0 + k_3. \tag{176}$$

Daraus folgt, dass $P+S$ beschränkt ist. Nach Poincaré-Bendixson existieren also periodische Lösungen des Schnakenberg-Modells.

Wir haben jetzt gesehen, wie die Existenz einer periodischen Lösung eines gegebenen Modells bewiesen werden kann. Dadurch bekommen wir allerdings wenig Informationen darüber, wo die Lösung liegt. Es gibt eine Möglichkeit, wie unter Umständen die Lösungen für bestimmte Parameterwerte besser lokalisiert werden können. Hierbei geht es um die schon erwähnten Relaxations-Oszillationen. Diese Ideen werden jetzt im Fall des van der Pol-Oszillators erklärt. Der van der Pol-Oszillator wird durch die Gleichung $\ddot{u}-k(1-u^2)\dot{u}+u=0$ definiert, wo $k>0$ ein Parameter ist. Wir ersetzen diese Gleichung durch das System

$$\epsilon\dot{u} = v - u^3/3 + u = v - G(u), \tag{177}$$

$$\dot{v} = -\epsilon u \tag{178}$$

mit $\epsilon = k^{-1}$. Wenn (u,v) das System (177)-(177) erfüllt, dann erfüllt u die ursprüngliche Gleichung zweiter Ordnung. Für dieses System kann man folgendes anschauliches Bild entwickeln. Wenn ϵ klein ist and eine Lösung weit von der Kurve $v=G(u)$ ist, dann ist die Ableitung von u groß und die von v klein. Deshalb hat die Lösung unter diesen Umständen die Tendenz, horizontale Sprünge zu machen. Ansonsten bewegt sie sich fast auf der Kurve $v=G(u)$. Insbesondere bekommt man ein Bild, wie eine periodische Lösung für ϵ klein aussehen könnte. Dieses Bild kann in einen Beweis umgewandelt werden. Die Funktion G hat die Symmetrie-Eigenschaft $G(-u)=-G(u)$. Sie hat ein eindeutiges Maximum bei $u=-1$ und eine eindeutiges Minimum bei $u=1$. Das Maximum ist $\frac{2}{3}>0$ und das Minimum $-\frac{2}{3}<0$. Die Nullstellen von G sind bei 0 und $\pm\sqrt{3}$. Die einzige stationäre Lösung des Systems ist im Ursprung.

Satz 13 Sei J die Jordan-Kurve, die aus folgenden Teilen besteht: die horizontale Strecke, die auf $v=G(u)$ startet und beim lokalen Maximum dieser Kurve endet, die horizontale Strecke, die auf $v=G(u)$ startet und beim lokalen Minimum dieser Kurve endet und die Teile der Kurve $v=G(u)$, die die Endpunkte dieser Strecken miteinander verbinden. Für ϵ klein genug hat der van der Pol-Oszillator eine periodische Lösung mit der Eigenschaft, dass ihr Bild für $\epsilon\to 0$ gegen J konvergiert.

Beweis Dazu wird ein Gebiet H konstruiert, deren Punkte nicht weiter von J sind als eine vorgegebene Konstante und die die Eigenschaft hat, dass für ϵ klein genug H invariant ist unter dem Fluss. Außerdem enthält dieses Gebiet keine stationären Lösungen. Nach Poincaré-Bendixson enthält das Gebiet eine periodische Lösung und die Definition des Gebiets sorgt für die erwünschte Konvergenz. Sei h eine positive Konstante. Sei x_1 der Punkt $\left(0,\frac{2}{3}+2h\right)$. Sei x_2 der eindeutige Punkt auf $v=G(u)$ mit $v=\frac{2}{3}+2h$. Sei x_3 der eindeutige Punkt auf $v=G(u)-h$ mit der gleichen u-Koordinate wie x_2. Sei x_4 der Punkt von $v=G(u)-h$, wo die Tangente den Punkt $\left(0,-\frac{2}{3}-2h\right)$ trifft. Die u-Koordinate von x_4 ist $\left(\frac{2+3h}{2}\right)^{\frac{1}{3}}$. Wir können die Symmetrie der Kurve $v=G(u)$ verwenden, um aus den Punkten x_1 bis x_4 weitere Punkte x_5 bis x_8 zu bestimmen. Mit diesen Punkten wird jetzt eine Jordan-Kurve J_1 konstruiert. x_1 wird mit x_2 durch eine horizontale Strecke verbunden, x_2 wird mit x_3 durch eine vertikale Strecke verbunden, x_3 wird mit x_4 durch einen Teil der Kurve $v = G(u) - h$ verbunden und es wird x_4 mit x_5 durch eine Strecke verbunden. Verbindungen zwischen den anderen Punkten werden dann durch die Symmetrie bestimmt. Seien x_9 und x_{10} die Punkte $(-1, 2/3)$ und $(0, 2/3)$, sei x_{11} ein noch zu bestimmender Punkt im Inneren von J mit $u > 1$ und $G(u) < v < 2/3$ und sei x_{12} der Punkt auf dem Graphen von $v = G(u)$ mit der gleichen u-Koordinate wie x_{11}. Punkte x_{13} bis x_{16} werden durch Symmetrie bestimmt. Mit diesen Punkten wird eine Jordan-Kurve J_2 konstruiert. Die Punkte x_9 bis x_{12} werden miteinander durch Strecken verbunden und x_{12} wird mit x_{13} durch einen Teil des Graphen von G verbunden. Verbindungen zwischen den anderen Punkten werden dann durch Symmetrie definiert. Die Kurven J_1 und J_2 treffen sich nicht und J_2 liegt im Inneren von J_1. Sei H das abgeschlossene Gebiet zwischen J_1 und J_2. Es bleibt zu zeigen, dass H die gewünschten Eigenschaften besitzt. Der Ursprung liegt im Inneren von J_2 wodurch gewährleistet ist, dass keine stationären Lösungen in H liegen. Für $h \to 0$ konvergiert der maximale Abstand von J eines Punktes von J_1 gegen Null. Wenn wir annehmen, dass der Abstand von x_{11} zu $(G^{-1}(2/3), 2/3)$ beliebig klein ist, dann ist der maximale Abstand von J eines Punktes von J_2 auch beliebig klein. Es bleibt zu zeigen, dass das Vektorfeld überall auf dem Rand von H nach innen weist. Auf den horizontalen und vertikalen Strecken ist diese Bedingung erfüllt, abgesehen von den Endpunkten. An den Endpunkten ist das Vektorfeld tangential zum Rand aber es kann trotzdem keine Lösung durch diese Punkte entkommen. Auf dem Teil von J_2 zwischen x_{12} und x_{13} ist sie auch erfüllt. Es bleiben noch drei Teile die zu überprüfen sind. Betrachten wir zuerst der Teil der Kurve J_1 zwischen x_3 und x_4. Sei $g(u)$ die Steigung dieser Kurve im Punkt u. Wenn für eine Lösung v als Funktion von u geschrieben wird dann gilt $\frac{dv}{du} = -\frac{\epsilon^2 u}{v - G(u)}$. Auf dem Teil der Kurve, den wir im Moment betrachten gilt $v - G(u) = -h$ und deshalb ist $\frac{dv}{du} = \frac{\epsilon^2 u}{h} < \frac{\epsilon^2 u(x_3)}{h}$. Für ϵ klein genug ist diese Größe kleiner als $g(x_4) < g(u)$. Dort ist $\dot{v} < 0$ und deshalb gilt die erwünschte Aussage. Als nächstes betrachten wir den Teil von J_1 zwischen x_4 und x_5. Dort ist $|v - G(u)| > h$ und es folgt, dass $|\frac{dv}{du}| \leq \frac{\epsilon^2 u(x_4)}{h}$. Für ϵ klein genug ist diese Größe kleiner als $g(x_4)$, die Steigung der Strecke. Als letztes wird der Teil von J_2 zwischen x_{10} und x_{11} überprüft. Sei K die Länge der Strecke von x_{11} nach x_{12}. Dann ist $|v - G(u)| > K$ auf dieser Strecke. Dort ist $|\frac{dv}{du}| \leq \frac{\epsilon^2 u(x_{11})}{K}$ und strebt gegen Null für $\epsilon \to 0$. Für ϵ klein genug weist das Vektorfeld dort ins Innere von H weil $\dot{u} > 0$ dort.

Die Theorie von Poincaré und Bendixson lässt sich nur auf zweidimensionale Systeme anwenden. Wie kann man die Existenz von Oszillationen in Systemen höherer Dimension beweisen? Eine Möglichkeit ist die Theorie der monotonen Systeme die jetzt kurz besprochen wird. Das System $\dot{x} = f(x)$ heißt monoton (oder kooperativ) wenn $\frac{\partial f_i}{\partial x_j} > 0$ für alle $i \neq j$. Der Name 'kooperativ' kommt vom Fall, dass die x_i Populationsdichten von verschiedenen Organismen sind. Dann bedeutet die Bedingung, dass eine höhere Populationsdichte einer Art das

Wachstum der Populationsdichte aller anderen Arten erhöht. Häufiger in der Praxis in der Populationsdynamik sind Fälle, wo $\frac{\partial f_i}{\partial x_j} < 0$ für alle $i \neq j$. Dann heißt das System kompetitiv. Durch die Transformation $t \to -t$ kann man ein kooperatives System in ein kompetitives überführen und umgekehrt. Monotone Systeme haben in gewissem Sinne ein einfacheres asymptotisches Verhalten als allgemeine dynamische Systeme. Grob gesagt haben die Lösungen eine größere Tendenz gegen stationäre Lösungen zu konvergieren. Es ist auch so, dass monotone Systeme von n Gleichungen nicht komplizierter sind als allgemeine Systeme von $n-1$ Gleichungen. Eine detaillierte Behandlung dieser Ideen mit präzisen Aussagen würde hier zu weit führen. Es soll aber ein Satz bewiesen werden, der in der Theorie eine zentrale Rolle spielt und den Namen 'monoton' erklärt. Gewisse Eigenschaften der Lösungen von monotonen Systemen lassen sich auf die von kompetitiven Systemen durch die Transformation $t \to -t$ übertragen.

Satz 14 (Müller-Kamke) Sei $\dot{x} = f(x)$ ein kooperatives dynamisches System auf einer konvexen Teilmenge G des $\mathbb{R}^m$ und seien x_0 und $\tilde{x}_0$ Punkte aus G mit $x_{0,i} \leq \tilde{x}_{0,i}$ für alle i. Seien $x(t)$ bzw. $\tilde{x}(t)$ Lösungen mit $x(0) = x_0$ und $\tilde{x}(0) = \tilde{x}_0$ die auf einem gemeinsamen Zeitintervall $[0, t_1)$ existieren. Dann gilt $x_i(t) \leq \tilde{x}_i(t)$ für alle $t \in [0, t_1)$ und alle i.

Beweis Sei $y_\epsilon(t)$ die Lösung von $\dot{y}_\epsilon = f(y_\epsilon) + \epsilon$ mit $y_\epsilon(0) = \tilde{x}_0$. Sei t_* das Supremum aller $t < t_1$, so dass $x_i(t) \leq y_{\epsilon,i}(t)$ für diese Werte von t und alle i. Entweder ist t_* die obere Grenze des maximalen Existenzintervalls der Lösung $y_\epsilon(t)$ oder es existiert mindestens ein j so dass $x_j(t_*) = y_{\epsilon,j}(t_*)$. Im zweiten Fall gilt

$$\frac{d}{dt}(y_{\epsilon,j} - x_j) = f_j(y_\epsilon) - f_j(x) + \epsilon > 0 \tag{179}$$

für $t = t_*$. Denn

$$\begin{aligned} f_j(y_\epsilon) - f_j(x) &= f_j(y_\epsilon) - f_j(x_1, y_{\epsilon,2}, ..., y_{\epsilon,n}) + \dots \\ &+ f_j(x_1, x_2, \dots, x_{n-1}, y_{\epsilon,n}) - f_j(x_1, x_2, \dots, x_{n-1}, x_n). \end{aligned} \tag{180}$$

Es folgt aus dem Fundamentalsatz der Integral- und Differentialrechnung, dass jeder Summand auf der rechten Seite nicht-negativ ist. Dass die Punkte über die man integrieren muss in G enthalten sind folgt aus der Konvexität. Deshalb gilt für solche Werte von j die Ungleichung $x_j(t) < y_{\epsilon,j}(t)$ für t etwas größer als t_* und die entsprechende Ungleichung gilt für die andere Komponenten durch Stetigkeit. Deshalb bekommen wir einen Widerspruch falls $t_* \neq t_1$. Es folgt dass $x_i(t) \leq y_{\epsilon,i}(t)$, so lange die Lösung $y_\epsilon(t)$ existiert. Durch stetige Abhängigkeit der Lösung von Parametern können wir schließen, dass $y_\epsilon(t)$ existiert für $t < t_1$ und ϵ hinreichend klein und dass $y_\epsilon(t)$ gegen $\tilde{x}(t)$ konvergiert für $\epsilon \to 0$. Damit ist der Satz bewiesen.

Bei den Fortschritten im Verständnis von HIV, die in den 1990er Jahre passiert sind hat, wie schon erwähnt, die mathematische Modellierung eine große Rolle gespielt. Damals erschienen um die gleiche Zeit zwei einflussreiche Arbeiten, die verschiedene mathematische Modelle für das gleiche biologische

System verwendet haben. Das eine ist das fundamentale System der Virusdynamik, dessen Asymptotik wir schon untersucht haben. Das andere nennen wir, um einen Namen dafür zu haben, das alternative System. Es unterscheidet sich vom fundamentalen System darin, dass in der Geichung für $\dot{x}$ ein Term $px\left(1-\frac{x}{\bar{x}}\right)$ hinzukommt. In diesem Ausdruck sind p und $\bar{x}$ positive Konstanten. Die Interpretation ist, dass im alternativen System berücksichtigt wird, dass nicht-infizierte T-Zellen sich durch Teilung vermehren können. Durch eine Variablentransformation können diese Systeme als kompetitive Systeme geschrieben werden. Sei $x_1 = x$, $x_2 = -y$ und $x_3 = v$. Dann ist das System für (x_1, x_2, x_3) kompetitiv. Deshalb können Aussagen benutzt werden, die denen der Poincaré-Bendixson-Theorie ähnlich sind. In [2] wurde gezeigt, dass die Asymptotik der Lösungen für $t \to \infty$ durch eine Zahl R_0 bestimmt wird. Wenn $R_0 \leq 1$ konvergieren die Lösungen gegen eine stationäre Lösung am Rand, mit $v = 0$. Wenn $R_0 > 1$ konvergieren alle Lösungen die nicht stationär sind für $t \to \infty$ gegen eine positive stationäre Lösung für bestimmte Werte der Parameter und gegen eine nichttriviale periodische Lösung für andere Werte der Parameter. Diese zwei Klassen werden durch die Stabilität der eindeutigen positiven stationären Lösung unterschieden. Die Frage stellt sich, wie zwei Modelle für das gleiche biologische System unterschiedliche Ergebnisse liefern können. In der Tat ist es so, dass die Parameterwerte die zu periodischen Lösungen führen nicht im biologisch relevanten Bereich liegen.

12 Verzweigungstheorie

Da es bei dynamischen Systeme der Dimension drei oder mehr oft schwierig oder unmöglich ist, das qualitative Verhalten der allgemeinsten Lösungen zu bestimmen ist es sinnvoll nach Methoden zu suchen, die es zumindest erlauben, die globale Dynamik in bestimmten eingeschränkten Fällen zu analysieren. Eine Methode dieser Art benutzt den Begriff 'Verzweigung' (oder Bifurkation). Nehmen wir an, wir haben ein System $\dot{x} = f(x, \lambda)$, das von einem Parameter λ abhängt und dass die Gleichung für $\lambda = 0$ leicht zu analysieren ist. Unter welchen Umständen ist das System für λ klein aber ungleich Null mit dem System für $\lambda = 0$ topologisch äquivalent und wenn dies nicht der Fall ist welche Beziehung besteht zwischen den Äquivalenzklassen der zwei Systeme? Wenn die Systeme nicht äquivalent sind spricht man von einer Verzweigung bei $\lambda = 0$.

Betrachten wir den Fall, dass das System $\dot{x} = f(x, \lambda)$ für $\lambda = 0$ eine stationäre Lösung hat für $x = 0$, d.h. $f(0, 0) = 0$. Ein besonders einfacher Fall ist der, in dem diese stationäre Lösung hyperbolisch ist. Dann ist insbesondere die Ableitung $Df(0)$ invertierbar und wir können den Satz über implizite Funktionen anwenden. Daraus folgt, dass es für x und λ klein genau eine Lösung $x = g(\lambda)$ gibt von $f(x, \lambda) = 0$ mit λ fest. Die Funktion g ist so glatt wie f. Die Matrix $Df(g(\lambda), \lambda)$ hängt glatt von λ ab. Die Eigenwerte davon hängen stetig von λ ab. Deshalb ist diese stationäre Lösung hyperbolisch für λ hinreichend klein und die Anzahl der Eigenwerte mit positivem Realteil bleibt konstant. Eine Senke bleibt eine Senke, eine Quelle bleibt eine Quelle und ein echter Sattelpunkt bleibt ein echter Sattelpunkt. Damit ist bewiesen, dass in diesem Fall $\lambda = 0$ kein Verzweigungspunkt ist.

Jetzt betrachten wir Verzweigungen in eindimensionalen dynamischen Systemen. Die Ergebnisse sind auch für höherdimensionale Systeme nützlich wegen der Möglichkeit einer Reduktion zu einer Zentrumsmannigfaltigkeit. Der Ausgangssystem darf eine beliebige Dimension haben, so lange die Zentrumsmannigfaltigkeit der stationären Lösung eindimensional ist. Sei also $\dot{x} = f(x, \lambda)$ eine Gleichung für eine reellwertige Funktion x. Wir schreiben f' für die Ableitung $\frac{\partial f}{\partial x}$ und benutzen eine analoge Notation für höhere Ableitungen. Der Fall ohne Verzweigung ist der, wo $f(0,0) = 0$ und $f'(0,0) \neq 0$. Dieser Fall ist in folgendem Sinn generisch. Wenn eine solche Funktion f auf einer Menge $[-a, a] \times [-a, a]$ gegeben ist können wir annehmen, in dem wir a gegebenfalls verkleinern, dass f' auf der ganzen Menge ungleich Null ist. Wenn $\epsilon > 0$ gegeben ist betrachten wir alle Funktionen g mit der Eigenschaft, dass

$$\sup(|g(x, \lambda) - f(x, \lambda)| + |g'(x, \lambda) - f'(x, \lambda)|) \tag{181}$$

auf dieser Menge nicht größer ist als ϵ. Für ϵ klein genug gibt es einen eindeutigen Punkt x in $(-a, a)$ mit $g(x, 0) = 0$ und $g'(x, 0) \neq 0$. Die Abwesenheit einer Verzweigung ist also stabil. Auf der anderen Seite kann man für jede Funktion f auf der gegebenen Menge mit $f(0,0) = 0$ eine Funktion g finden, die die Ungleichung erfüllt, so dass $g(0,0) = 0$ und $g'(0,0) \neq 0$. Man kann diese Aussagen auch so formulieren, dass man sagt, dass die Funktionen ohne Verzweigung bezüglich einer geeigneten Topologie eine offene und Dichte Teilmenge bilden. Es sollen aber solche Formulierungen hier nicht weiter verfolgt werden und der Begriff 'generisch' wird ab diesem Punkt auf eine intuitive Art eingesetzt.

Die Möglichkeiten für Verzweigungen werden erkundet in dem man eine gewisse Anzahl von Bedingungen annimmt und dann Fälle betrachtet, die innerhalb dieser Klasse generisch sind. Im einfachsten Fall nehmen wir an, dass $f'(0,0) = 0$, so dass eine Verzweigung vorliegt und dass die Bedingungen $f''(0,0) \neq 0$ und $\partial f / \partial \lambda(0,0) \neq 0$ gelten. Diese Verzweigung heißt Falte (oder auch Sattelknoten). Ein Modell dafür ist $f(x, \lambda) = x^2 - \lambda$. Wir könnten auch den äquivalenten Ausdruck $f(x, \lambda) = -x^2 + \lambda$ nehmen. Der andere ähnliche Ausdruck $f(x, \lambda) = x^2 + \lambda$ ist mit den anderen beiden nicht topologisch äquivalent weil die Definition der topologischen Äquivalenz verlangt, dass die Zeitrichtung erhalten bleibt. In diesem Beipiel gibt es keine stationären Lösungen für $\lambda < 0$, genau eine stationäre Lösung für $\lambda = 0$ und zwei stationäre Lösungen für $\lambda > 0$, davon eine asymptotisch stabil und eine instabil. Diese Verhältnisse hat man immer wenn es eine Falte gibt. Intuitiv kann man die Situation so beschreiben. Wenn λ zunimmt gibt es einen kritischen Parameterwert wo zwei stationäre Lösungen (eine stabil und eine instabil) aus dem Nichts entstehen. Oder man könnte sagen, dass wenn λ abnimmt zwei stationäre Lösungen (eine stabil und eine instabil) zusammenstossen und sich vernichten.

Nehmen wir jetzt an, wir haben ein System der Form $\dot{x} = x^2 - \lambda + O(x^3)$. Dann ist es lokal topologisch äquivalent zum Modellsystem $\dot{x} = x^2 - \lambda$. Für den

Beweis benutzen wir die Tatsache, dass in einer Dimension ein Homöomorphismus, der stationäre Lösungen in stationäre Lösungen abbildet auch die Bilder der Lösungen die sie verbinden aufeinander abbildet. Betrachten wir das System $\dot{y} = y^2 - \lambda + \psi(y,\lambda)$, wobei ψ glatt ist und die Bedingung $\psi(y,\lambda) = O(y^3)$ erfüllt. Nach dem Satz über implizite Funktionen bilden die stationären Lösungen eine Mannigfaltigkeit der Form $\lambda = g(y)$, wo $g(y) = y^2 + O(y^3)$. Für λ hinreichend klein und positiv gibt es also genau zwei stationäre Lösungen nahe bei Null. Als nächstes wird ein parameterabhängiger Homöomorphismus h_λ konstruiert für λ klein, der die Äquivalenz definiert. Für $\lambda < 0$ ist h_λ die Identität. Für $\lambda > 0$ ist h_λ eine lineare Funktion $h_\lambda(x) = a(\lambda) + b(\lambda)x$, wo die Koeffizienten a und b so gewählt werden, dass $-\sqrt{\lambda}$ und $+\sqrt{\lambda}$ auf die beiden stationären Lösungen abgebildet werden.

Die generische Falte, also ein System für das im Ursprung $f = 0$, $f' = 0$, $f'' \neq 0$ und $\partial f/\partial\lambda \neq 0$ gelten ist lokal topologisch äquivalent zum Modellsystem. Um diese Aussage zu beweisen reicht es zu zeigen, dass das gegebene System topologisch äquivalent ist zu einem System der speziellen Form, die gerade behandelt wurde. Eine Taylor-Entwicklung der rechten Seite in x gibt

$$f(x,\lambda) = f_0(\lambda) + f_1(\lambda)x + f_2(\lambda)x^2 + O(x^3). \tag{182}$$

Diese Gleichung kann vereinfacht werden in dem man in x verschiebt, $\xi = x + \delta$, wobei die Verschiebung δ von λ abhängt. Einsetzen in das dynamische System gibt

$$\dot{\xi} = f_0(\lambda) + f_1(\lambda)(\xi - \delta) + f_2(\lambda)(\xi - \delta)^2 + O((\xi - \delta)^3). \tag{183}$$

Sortieren der Terme nach Potenzen von ξ liefert

$$\begin{aligned}\dot{\xi} = {} & [f_0(\lambda) - f_1(\lambda)\delta + f_2(\lambda)\delta^2 + O(\delta^3)] \\ & + [f_1(\lambda) - 2f_2(\lambda)\delta + O(\delta^2)]\xi \\ & + [f_2(\lambda) + O(\delta)]\xi^2 + O(\xi^3).\end{aligned} \tag{184}$$

Die Bedingung, dass der Koeffizient von ξ verschwindet ist dass

$$F(\lambda,\delta) = f_1(\lambda) - 2f_2(\lambda)\delta + \psi(\lambda,\delta)\delta^2 = 0 \tag{185}$$

für eine glatte Funktion ψ. Wir haben $F(0,0) = 0$, $\frac{\partial F}{\partial \delta}(0,0) = -2f_2(0) \neq 0$ und $\frac{\partial F}{\partial \lambda}(0,0) = f_1'(0)$. Der Satz über implizite Funktionen hat zur Folge, dass es eine glatte Funktion $\delta = \delta(\lambda)$ gibt mit $\delta(0) = 0$ und $F(\lambda,\delta(\lambda))(0) = 0$. Es folgt auch, dass $\delta(\lambda) = \frac{f_1'(0)}{2f_2(0)}\lambda + O(\lambda^2)$. Nach der Transformation enthält die Gleichung für $\dot{\xi}$ keine Terme, die linear in ξ sind. Betrachten wir einen neuen Parameter $\mu = \mu(\lambda)$, wo die rechte Seite der Koeffizient von ξ^0 in der Entwickung nach Potenzen von ξ ist. Dann ist $\mu(\lambda) = f_0'(0)\lambda + \lambda^2\phi(\lambda)$ mit einer glatten Funktion ϕ. $\mu(0) = 0$ und $\mu'(0) = f_0'(0) = \frac{\partial f}{\partial \lambda}(0,0)$. Nach dem Satz über die Umkehrabbildung gibt es eine glatte Umkehrabbildung $\lambda = \lambda(\mu)$ mit $\lambda(0) = 0$. Deshalb wird die Gleichung für $\dot{\xi}$ zu $\dot{\xi} = \mu + b(\mu)\xi^2 + O(\xi^3)$. Dabei ist b eine glatte Funktion mit $b(0) = f_2(0) \neq 0$. Sei $\eta = |b(\mu)|\xi$ und $\beta = |b(\mu)|\mu$. Dann ist $\dot{\eta} = \beta + s\eta^2 + O(\eta^3)$, wo $s = \pm 1$ entspricht dem Vorzeichen von $b(0)$.

Als nächstes wird der Fall betrachtet, dass im Ursprung f, f' und f'' verschwinden, aber f''' verschwindet nicht. Es stellt sich heraus, dass um dafür zu sorgen, dass man ein Modell für alle Parameterwerte in der Nähe von Null bekommt, man den Fall mit zwei Parametern betrachten muss. In diesem Fall braucht man also eine Funktion $f(x, \lambda_1, \lambda_2)$. Um den generischen Fall zu bekommen verlangt man, dass im Ursprung $\frac{\partial^2 f}{\partial x \partial \lambda_1} \frac{\partial f}{\partial \lambda_2} - \frac{\partial^2 f}{\partial x \partial \lambda_2} \frac{\partial f}{\partial \lambda_1} \neq 0$. Diese Verzweigung heißt generische Spitze. Ein Modell für diese Verzweigung ist $\lambda_1 + \lambda_2 x - x^3$. Es enthält Punkte wo es eine Falte gibt. Es sind die Punkte außer dem Ursprung wo die zwei Gleichungen $\lambda_1 x + \lambda_2 x - x^3 = 0$ und $\lambda_2 - 3x^2 = 0$ gelten. Man kann x von diesen Gleichungen eliminieren mit dem Ergebnis $4\lambda_2^3 = 27\lambda_1^2$. Die Lösungsmenge ist eine Kurve, die eine Spitze im Ursprung hat und daher kommt der Name dieser Verzweigung. Wenn die Parameter außerhalb des Gebiets liegen, das von der Spitze begrenzt wird hat das System genau eine stationäre Lösung und diese ist stabil. Innerhalb der Spitze gibt es drei stationäre Lösungen, zwei stabil und eine instabil.

Ein System, das eine generische Spitze enthält wird jetzt mit einem Vorgehen vereinfacht das dem im Fall der Falte ähnlich ist. Um die Notation zu verkürzen sei $\lambda = (\lambda_1, \lambda_2)$. Eine Taylor-Entwicklung in x liefert

$$f(x, \lambda) = f_0(\lambda) + f_1(\lambda)x + f_2(\lambda)x^2 + f_3(\lambda)x^3 + O(x^4). \tag{186}$$

Es gelten die Bedingungen $f_1(0) = 0$ und $f_2(0) = 0$. Wie bei der Falte versuchen wir eine Vereinfachung zu erreichen durch eine Verschiebung $\xi = x + \delta$. Einsetzen in das dynamische System gibt

$$\begin{aligned}\dot{\xi} &= [f_0(\lambda) - f_1(\lambda)\delta + \omega(\lambda, \delta)\delta^2] + [f_1(\lambda) - 2f_2(\lambda)\delta + \phi(\lambda, \delta)\delta^2]\xi \\ &+[f_2(\lambda) - 3f_3(\lambda)\delta + \psi(\lambda, \delta)\delta^2]\xi^2 + [f_3(\lambda)\delta + \theta(\lambda, \delta)]\xi^3 + O(\xi^4)\end{aligned} \tag{187}$$

für glatte Funktionen ω, ϕ, ψ und θ. Weil $f_2(0) = 0$ ist es hier nicht möglich, wie im Fall der Falte den Satz über implizite Funktionen zu benutzen, um den linearen Term in ξ zu eliminieren. Wir können aber den quadratischen Term in ξ eliminieren. Dazu setzen wir $F(\lambda, \delta) = f_2(\lambda) - 3f_3(\lambda)\delta + \psi(\lambda, \delta)\delta^2$ und stellen fest, dass $F(0, 0) = 0$ und $\frac{\partial F}{\partial \delta}(0, 0) = -3f_3(0) \neq 0$. Betrachten wir neue Parameter μ_1 und μ_2 mit

$$\mu_1(\lambda) = f_0(\lambda) - f_1(\lambda)\delta(\lambda) + \delta^2(\lambda)\omega(\lambda, \delta(\lambda)), \tag{188}$$

$$\mu_2(\lambda) = f_1(\lambda) - 2f_2(\lambda)\delta(\lambda) + \delta^2(\lambda)\phi(\lambda, \delta(\lambda)). \tag{189}$$

Die Größe $\mu = (\mu_1, \mu_2)$ erfüllt $\mu(0) = 0$. Die neuen Parameter können eingeführt werden wenn die Jacobi-Determinante $\det(\partial\mu/\partial\lambda)$ ungleich Null ist. Diese Bedingung ist mit der zweiten Bedingung für die generische Spitze äquivalent. Dann kann der Satz über die Umkehrabbildung angewendet werden. Wir erhalten eine glatte Umkehrabbildung $\lambda = \lambda(\mu)$ mit $\alpha(0) = 0$. Nach der Transformation zu den neuen Parametern hat die Gleichung für $\dot{\xi}$ die Form

$$\dot{\xi} = \mu_1 + \mu_2\xi + c(\mu)\xi^3 + O(\xi^4) \tag{190}$$

wo $c(\mu) = f_3(\lambda(\mu)) + \delta(\lambda(\mu))\omega(\lambda(\mu), \delta(\lambda(\mu))$ eine glatte Funktion von μ ist und $c(0) = f_3(0) = \frac{1}{6}f'''(0,0) \neq 0$. Zum Schluss wird die lineare Skalierung $\eta = \sqrt{|c(\mu)|}\xi$ ausgeführt und neue Parameter $\beta_1 = \sqrt{|c(\mu)|}\mu_1$ und $\beta_2 = \mu_2$ definiert. Das Ergebnis ist

$$\dot{\eta} = \beta_1 + \beta_2\eta + s\eta^3 + O(\eta^4). \tag{191}$$

Es ist jetzt gezeigt worden, dass im Fall der Spitze eine approximative Normalform erreicht werden kann. Es ist auch möglich, wie im Fall der Falte, weiter zu gehen und eine exakte Normalform zu erreichen. Diese Aussage wird aber hier nicht bewiesen. Es gibt einen Spezialfall der Spitze, die Heugabel, bei dem das System die Symmetrie $x \mapsto -x$ besitzt. Diese Verzweigung ist in der Klasse der Systeme mit dieser Symmetrie generisch.

Es wurde schon erwähnt, dass Aussagen über Verzweigungstheorie in einer Dimension auf Probleme in höheren Dimensionen angewendet werden können. Jetzt soll diese Bemerkung ausgebaut werden. Dazu betrachten wir beispielhaft ein zweidimensionales dynamisches System $\dot{x} = f(x, \lambda)$ mit $f(0,0) = 0$. Nehmen wir an, dass der Rang der Ableitung $Df(0,0)$ Eins ist und dass der Eigenwert dieser Matrix, der nicht verschwindet negativ ist. Die Dimension der Zentrumsmannigfaltigkeit ist Eins. Betrachten wir jetzt das erweiterte System

$$\dot{x} = f(x, \lambda), \tag{192}$$

$$\dot{\lambda} = 0. \tag{193}$$

In diesem System ist der Parameter zu einer Unbekannten geworden. Das erweiterte System hat eine stationäre Lösung bei $(0,0,0)$ und die Zentrumsmannigfaltigkeit dort ist zweidimensional. Nennen wir sie M. Die Durchschnitte M_λ der λ-Koordinatenebenen mit M sind invariante Mannigfaltigkeiten für die Systeme im parameterabhängigen Formulierung. Die Mannigfaltigkeit M_0 ist eine Zentrumsmannigfaltigkeit des Ursprungs im System für $\lambda = 0$. Wenn die Systeme für $\lambda \neq 0$ stationäre Lösungen haben, die nahe genug beim Ursprung im (x, λ)-Raum sind, dann müssen sie auf M_λ liegen. Die Mannigfaltigkeit M_λ ist invariant, aber im allgemeinen keine Zentrumsmannigfaltigkeit für die stationären Lösungen, die sie enthält. Kommen wir jetzt zu unserem zweidimensionalen System zurück und nehmen wir an dass die Einschränkung des Systems auf M eine Verzweigung bei $\lambda = 0$ hat und zwar eine Falte. Wir wissen wie diese Situation für die Einschränkung charakterisiert werden kann. Nach dem Satz von Shoshitaishvili wissen wir dann auch wie das qualitative Verhalten des ursprünglichen Systems in der Nähe des Ursprungs aussieht. Für $\lambda < 0$ gibt es keine stationären Lösungen. Für $\lambda > 0$ gibt es zwei. Davon ist eine asymptotisch stabil während die andere ein Sattelpunkt ist. Für $\lambda = 0$ gibt es genau eine stationäre Lösung und sie ist ein Sattelknoten. Das heißt, sie hat eine Umgebung, die durch die stabile Mannigfaltigkeit in zwei Gebiete aufgeteilt wird, wobei die eine Seite wie ein Knoten aussieht (hyperbolische Senke) und die andere wie ein Sattel. Daher kommt der alternative Name für diese Verzweigung. Ob das System auf der Zentrumsmannigfaltigkeit die gewünschten Eigenschaften besitzt kann durch direkte Rechnungen mit dem vollen System geprüft werden. Wir geben hier diese Bedingungen an, ohne Beweis. Damit es eine Falte im Punkte $(0,0)$ gibt müssen folgende Bedingungen gelten. $f(0,0) = 0$ und der Rang von $Df(0,0)$ ist Eins. Die Bedingung $l(D^2 f(r,r)) \neq 0$ gilt, wo l und r Links- und Rechtseigenvektoren der Matrix $Df(0,0)$ mit Eigenwert Null sind. Schließlich gilt $l(\partial f/\partial\lambda) \neq 0$. Es gibt eine ähnliche Charakterisierung der Spitze in höheren Dimensionen, die allerdings ein wenig komplizierter aussieht, als man vielleicht vermuten könnte. Die Bedingungen sind $f(0,0) = 0$, der Rang von $Df(0,0)$ ist Eins, $l(D^2 f(r,r)) = 0$, $l(D^3 f(r,r,r)) - 3D^2 f(r,z) \neq 0$ für einen geeigneten Vektor z dessen Definition wir hier nicht angeben. Die Korrektur in der Bedingung für die dritte Ableitung hat damit zu tun, dass man auf der Zentrumsmannigfaltigkeit arbeiten muss und nicht nur auf dem Zentrumsteilraum. Die Bedingung die Ableitungen bezüglich der Parameter enthält muss auch geeignet verallgemeinert werden. Für weitere Einzelheiten verweisen wir auf das Buch von Kuznetsov [7].

Wir haben jetzt ein paar Aspekte des Falles betrachtet, im dem die Linearisierung des Systems am Verzweigungspunkt einen verschwindenden Eigenwert hat. Als nächstes betrachten wir den Fall, dass es ein Paar rein imaginäre Eigenwerte gibt, die nicht verschwinden. Dazu braucht man natürlich ein System der Dimension mindestens zwei. Das Modellsystem in diesem Fall ist

$$\dot{x}_1 = \lambda x_1 - x_2 - x_1(x_1^2 + x_2^2), \tag{194}$$

$$\dot{x}_2 = x_1 - \lambda x_2 - x_2(x_1^2 + x_2^2). \tag{195}$$

Das System hat eine stationäre Lösung im Punkt $(0,0)$ für alle λ und die Jacobi-Matrix dort ist

$$\begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} \tag{196}$$

mit Eigenwerten $\lambda \pm i$. In dieser Form ist die Struktur des Systems nicht sehr transparent. Man sieht sie besser in Polarkoordinaten. Es ist hilfreich für die Rechnungen eine komplexe Größe $z = x_1 + ix_2$ einzuführen. Dann ist $|z|^2 = x_1^2 + x_2^2$ und

$$\dot{z} = (\lambda + i)z - z|z|^2. \tag{197}$$

Mit $z = \rho e^{i\phi}$ bekommen wir

$$\dot{\rho} = \rho(\lambda - \rho^2), \tag{198}$$

$$\dot{\phi} = 1. \tag{199}$$

Diese Gleichungen sind entkoppelt. Die erste Gleichung ist natürlich nur für $\rho \geq 0$ relevant. Für $\lambda < 0$ ist die stationäre Lösung bei $\rho = 0$ asymptotisch stabil und hyperbolisch. Für $\lambda = 0$ ist sie immer noch stabil aber nicht mehr hyperbolisch. Für $\lambda > 0$ ist die stationäre Lösung bei $\rho = 0$ instabil und es entsteht eine neue stabile stationäre Lösung bei $\rho = \sqrt{\lambda}$. Die zweite Gleichung beschreibt eine Rotation mit konstanter Geschwindigkeit. Wenn wir diese Informationen kombinieren bekommen wir folgendes Bild der Dynamik des zweidimensionalen

Systems. Für $\lambda < 0$ gibt es eine hyperbolische Senke im Ursprung. Lösungen außerhalb des Ursprungs nähern sich diesem Punkt spiralförmig für $t \to \infty$. Das System für $\lambda = 0$ ist mit dem System für $\lambda < 0$ topologisch äquivalent. Für $\lambda > 0$ gibt es eine stabile periodische Lösung, die den Ursprung umkreist. Die soeben beschriebene Verzweigung ist die Hopf-Verzweigung. Es gibt eine ähnliche Verzweigung bei der das Vorzeichen im nichtlinearen Term im Modell geändert wird. In dem Fall ist die periodische Lösung instabil. Diese zwei Fälle heißen überkritisch bzw. unterkritisch.

Wie in den anderen bisher betrachteten Verzweigungen wollen wir allgemeinere Systeme mit dem Modellsystem in Beziehung setzen. Wenn auf der rechten Seite des Modellsystems ein Term hinzugefügt wird, der $O(|x|^4)$ ist, dann ist das Ergebnis mit dem Modellsystem topologisch äquivalent. Wir betrachten jetzt generische Hopf-Verzweigungen. Wenn ein zweidimensionales System eine stationäre Lösung im Ursprung besitzt und die Eigenwerte dort rein imaginär sind aber nicht Null, dann folgt aus dem Satz über implizite Funktionen, dass es für λ klein aber ungleich Null genau eine stationäre Lösung in der Nähe des Ursprungs gibt. Nach einer λ-abhängigen Koordinatentransformation können wir annehmen, dass $f(0, \lambda) = 0$ für alle λ. Damit die Verzweigung generisch ist und das System mit dem Modell topologisch äquivalent ist reicht es, dass zwei Bedingugen erfüllt sind. Die erste besagt, dass der Realteil des Eigenwerts der Jacobi-Matrix im Ursprung sich mit positiver Geschwindigkeit durch Null bewegt für $\lambda = 0$. Die zweite Bedingung ist, dass der erste Ljapunow-Koeffizient nicht verschwindet. Es handelt sich dabei um eine Kombination der Ableitungen zweiter und dritter Ordnung von f nach x im Ursprung. Insbesondere hängt diese Größe nur vom System bei $\lambda = 0$ ab. Die überkritischen und unterkritischen Fälle werden durch das Vorzeichen des Ljapunowkoeffizienten unterschieden. Mit Hilfe von Zentrumsmannigfaltigkeiten kann man Hopf-Verzweigungen in höheren Dimensionen definieren.

Nachdem wir den van der Pol-Oszillator im Grenzfall $k \to \infty$ untersucht haben, wollen wir den Fall $k \to 0$ betrachten mit dem Ziel, eine Hopf-Verzweigung zu finden. Es stellt sich heraus dass man dazu am besten die Gleichungen reskaliert [13]. Ausgehend von dem System erster Ordnung für u und v definieren wir $x = k^{\frac{1}{2}} u$ und $y = k^{\frac{3}{2}} v$. Das so transformierte System ist

$$\dot{x} = y - x^3/3 + kx = 0, \tag{200}$$

$$\dot{y} = -x. \tag{201}$$

Die Eigenwerte bei $(0, 0)$ sind $\frac{1}{2}(k \pm \sqrt{k^2 - 4})$. Die erste Bedingung für eine generische Hopf-Verzweigung ist erfüllt. Um die zweite Bedingung zu überprüfen können wir zuerst feststellen, dass der lineare Teil des Systems bei $k = 0$ schon in der Standardform ist. Unter diesen Umständen ist es einfach, eine Formel für den Ljapunowkoeffizienten benutzen, die in [9] angegeben wird. Der Koeffizient ist $-\frac{3\pi}{2}$ und es handelt sich um eine überkritische generische Hopf-Verzweigung. Daraus folgt, dass es eine stabile stationäre Lösung in der Nähe des Ursprungs gibt für $k > 0$. Die Amplitude der Oszillation, gemessen in den transformierten Variablen ist in führender Ordnung proportional $\sqrt{k}$. In der ursprünglichen

Größe u ist die Amplitude in führender Ordnung unabhängig von k, so dass es eine generische Hopfverzweigung in den ursprünglichen Variablen gar nicht geben könnte.

13 Das Lorenz-System

In dieser Vorlesung haben wir uns in relativ ruhigen Gewässern aufgehalten. Es ist aber wichtig zu wissen dass es im Meer der dynamischen Systeme viele Stürme gibt. Das Lorenz-System ist ein System von drei gewöhnlichen Differentialgleichungen, das von drei Parametern abhängt. Es sieht auf den ersten Blick harmlos aus:

$$\dot{x} = \sigma(y - x), \tag{202}$$

$$\dot{y} = rx - y - xz, \tag{203}$$

$$\dot{z} = xy - bz. \tag{204}$$

Der Meteorologe Edward Lorenz hat dieses System erhalten als ein Modellsystem für konvektive Rollen in der Atmosphäre. Die Parameter σ bzw. r haben direkte physikalische Interpretationen als die Prandtl-Zahl bzw. Rayleigh-Zahl. Der Parameter b hat keinen besonderen Namen. Alle Parameter werden als positiv vorausgesetzt. Das System ist symmetrisch unter der Transformation $(x, y, z) \mapsto (-x, -y, z)$. Der Ursprung ist eine stationäre Lösung für alle Werte der Parameter. Alle stationären Lösungen erfüllen $x = y$, $x^2 = bz$ und $(r-1)x = b^{-1}x^3$. Daraus folgt, dass es keine stationären Lösungen außerhalb des Ursprungs gibt für $r < 1$. Für $r > 1$ gibt es zwei stationäre Lösungen mit $x = y = \pm\sqrt{b(r-1)}$ und $z = r - 1$.

Jetzt wird die Stabilität der stationären Lösungen untersucht. Wenn man im Ursprung linearisiert dann entkoppelt die Gleichung für z. Die linearisierte Größe z klingt exponentiell ab. Die Linearisierung in x und y liefert eine Matrix mit Determinante $\sigma(1 - r)$. Wenn $r > 1$ dann gibt es einen positiven und einen negativen Eigenwert und der Ursprung ist ein Sattelpunkt. Die stabile Mannigfaltigkeit ist zweidimensional und die instabile Mannigfaltigkeit eindimensional. Die Spur ist $-\sigma - 1$ und deshalb ist der Ursprung für $r < 1$ eine hyperbolische Senke und asymptotisch stabil. Wir können mehr sagen mit Hilfe einer Ljapunow-Funktion. Sei $V(x, y, z) = \frac{1}{\sigma}x^2 + y^2 + z^2$. Diese Funktion verschwindet nur im Ursprung, wo sie ein globales Minimum hat.

$$\begin{aligned}\frac{1}{2}\dot{V} &= (r+1)xy - x^2 - y^2 - bz^2 \\ &= -\left[x - \frac{r+1}{2}y\right]^2 - \left[1 - \left(\frac{r+1}{2}\right)^2\right]y^2 - bz^2.\end{aligned} \tag{205}$$

$\dot{V} \leq 0$ und deshalb ist V eine Ljapunow-Funktion wenn $r < 1$. Außerdem verschwindet $\dot{V}$ nur im Urspung. Deshalb ist der Ursprung in diesem Fall global asymptotisch stabil. Für die Jacobi-Matrix der zwei anderen stationären Lösungen ist die charakteristische Gleichung

$$\lambda^3 + (\sigma + b + 1)\lambda^2 + (r + \sigma)b\lambda + 2b\sigma(r - 1) = 0. \tag{206}$$

Für r ein wenig grösser als Eins sagt das Kriterium von Routh und Hurwitz, dass alle Eigenwerte negativen Realteile haben. Es kann nicht passieren, dass für $r > 1$ ein Eigenwert Null wird. Deshalb bleiben die Realteile der Eigenwerte negativ so lange diese die imaginäre Achse außerhalb des Ursprungs treffen. Wenn $\lambda = i\omega$ mit ω reell wird diese Gleichung zu den Bedingungen $\omega^2 = (r+\sigma)b$ und $(\sigma+b+1)\omega^2 = 2b\sigma(r-1)$. Wir können ω von diesen Gleichungen eliminieren, mit dem Ergebnis, dass $r = r_H = \sigma\left(\frac{\sigma+b+3}{\sigma-b-1}\right)$. Diese Formel liefert nur dann eine relevante Lösung, wenn das Ergebnis positiv ist. Bei $r = r_H$ gibt es eine Hopf-Verzweigung, wo die zwei stabilen stationären Lösungen instabil werden. Wenn diese Verzweigung überkritisch wäre, dann könnte sie einen Kandidaten liefern für die ω-Limesmenge von anderen Lösungen. In Wirklichkeit ist sie aber unterkritisch.

Wie ist das Langzeitverhalten von Lösungen des Lorenz-Systems? Wir haben jetzt ein paar einfache Kandidaten für ω-Limesmengen ausgeschlossen. Könnte es sein, dass die Lösungen für $t \to \infty$ gegen unendlich streben? Es wird jetzt gezeigt, dass auch diese Alternative ausgeschlossen werden kann. Nehmen wir an, dass zu einem bestimmten Zeitpunkt eine Lösung die Ungleichung $2\sigma x^2 + 2y^2 + b^2z^2 \geq C_1$ erfüllt für eine Konstante $C_1 > 0$.

$$\begin{aligned}
&\frac{d}{dt}[x^2 + y^2 + (z - r - \sigma)^2] = 2[-\sigma x^2 - y^2 - bz^2 + b(r + \sigma)z] \\
&\leq -2\sigma x^2 - 2y^2 - bz^2 + b(r + \sigma)^2 \\
&\leq b(r + \sigma)^2 - C_1.
\end{aligned} \tag{207}$$

Es folgt, dass wenn $C_1 > b(r + \sigma)^2$ die Größe $x^2 + y^2 + (z - r - \sigma)^2$ mit einer positiven Rate abnimmt. Wenn für eine Konstante $C_2 > 0$ die Kugel K die durch die Ungleichung $x^2 + y^2 + (z - r - \sigma)^2 \leq C_2$ definiert wird das Gebiet E enthält, das durch die Ungleichung $2\sigma x^2 + 2y^2 + b^2z^2 \leq C_1$ definiert wird, dann muss eine Lösung, die außerhalb von K anfängt K in endlicher Zeit erreichen und danach in K bleiben. Damit ist insbesondere bewiesen, dass die Lösungen des Lorenz-Systems beschränkt sind. Die Divergenz des Vektorfeldes, das das Lorenz-System definiert ist $-\sigma - 1 - b < 0$. Die Divergenz ist eine Konstante. Wenn wir das Volumen v eines Gebiets wie K betrachten, dann sehen wir, mit dem Satz von Stokes, dass $\dot{v} = -(\sigma + 1 + b)v$. Das Volumen konvergiert exponentiell gegen Null für $t \to \infty$.

In seiner ursprünglichen Arbeit hat Lorenz die Parameterwerte $\sigma = 10$, $b = \frac{8}{3}$ und $r = 28$ genommen. Später haben viele Autoren den Parameterraum erkundet, wobei sie meistens r variiert haben und die anderen beiden Parameter unverändert gelassen haben. Für die gegebenen Werte der anderen Parameter findet die schon erwähnte Hopf-Verzweigung bei einem Wert von r statt, der etwa 24,74 ist. Der instabile Mannigfaltigkeit des Urprungs ist eindimensional und man kann fragen wo sie hinführt. Bei einem bestimmten Wert von r, etwa 13,926 kommt sie zum Ursprung zurück und es gibt eine homokline Lösung. Wenn man von diesem Wert aus r nach oben verändert entspringen die instabilen periodischen Lösungen die bei den instabilen stationären Lösungen enden der homoklinen Lösung. Die Verzweigungen, die wir bisher besprochen haben sind lokale Verzweigungen. In dem Fall passiert alles in der Nähe einer stationären Lösung. Dagegen ist die homokline Verzweigung im Lorenz-System eine globale Verzweigung. Das gerade beschriebene Bild basiert auf numerischen Simulationen aber es ist anscheinend zumindest bewiesen, dass eine homokline Lösung existiert. Es gibt eine Theorie, die homokline Verzweigungen klassifiziert. In bestimmten Fällen gibt es ein sehr kompliziertes Verhalten in der Nähe der homoklinen Lösung. Deshalb spricht man in diesem Fall oft von einer homoklinen Explosion.

Für die Parameterwerte von Lorenz zeigen Computersimulationen, dass Lösungen gegen eine Menge konvergieren, die man Lorenz-Attraktor nennt. Auf dem Attraktor selbst kann man die Lösungen nicht lokalisieren. Sie springen zwischen zwei Teile des Attraktors, die wie Scheiben aussehen, hin und her. Die Scheiben sind keine Mannigfaltigkeiten sondern sehen eher wie viele Schichten aus, die aufeinander liegen, ein wenig wie Blätterteig. Gewisse Teile dieser Aussagen konnten durch Warwick Tucker bewiesen werden. Es handelt sich um einen Computergestützten Beweis. Das heißt, das Argument benutzt den Computer aber ist trotzdem ein strenger Beweis. Die dabei verwendete Technik heißt Intervall-Arithmetik.

Der Lorenz-Attraktor wird manchmal als 'seltsamen Attraktor' beschrieben. In dem Zusammenhang gibt es mehrere Probleme. Das erste ist die Definition von Attraktor. Es gibt mehr als eine Definition in der Literatur und es nicht klar, dass eine davon die beste ist. Das gleiche gilt für die Definition von 'seltsam' in diesem Zusammenhang. Das dritte Problem ist zu zeigen, wenn man eine Definition festgelegt hat, dass das Lorenz-System (oder eine anderes System) eine Menge besitzt, die diese Definitionen erfüllt. Wir versuchen hier lediglich, die intuitiven Ideen zu skizzieren, die bei der Betrachtung dieser Fragen eine Rolle spielen. Eine mögliche Definition eines Attraktors ist wie folgt. Es handelt sich um eine Menge A, die invariant ist unter dem Fluss, die eine offene Umgebung U hat mit der Eigenschaft, dass jede Lösung die in U startet für $t \to \infty$ gegen A konvergiert und die minimal ist in dem Sinne, dass sie nicht die Vereinigung ist von zwei anderen Mengen, die die ersten zwei Bedingungen erfüllen. Der Attraktor heißt seltsam, wenn er sensible Abhängigkeit der Lösungen von den Anfangsdaten zeigt. Die letzte Bedingung bedeutet, dass für hinreichend viele Anfangsdaten auf dem Attraktor, der maximale Ljapunowexponent der Lösung positiv ist. Diese letzte Größe misst, wie schnell benachbarte Lösungen sich von der gegebenen Lösung entfernen. Wenn der Attraktor ein Punkt ist, dann ist dieser Exponent der größte Realteil eines Eigenwerts der Linearisierung.

14 Virusdynamik in der Praxis

Wie wurden die Modelle der Virusdynamik in der Praxis angewendet? Eine stationäre Lösung des Modells entspricht dem Zustand einer Person, die mit HIV infiziert ist. Nehmen wir an, dass diese Person mit einem Medikament behandelt wird, das verhindert, dass neue Zellen infiziert werden. Um den Zustand während der Behandlung zu modellieren setzen wir den Parameter β zu Null. Dann entkoppeln die Gleichungen für y und v von der Gleichung für x. Wir haben dann das System

$$\dot{y} = -ay, \tag{208}$$

$$\dot{v} = ky - uv. \tag{209}$$

Diese Gleichungen sind linear und können explizit gelöst werden, mit dem Ergebnis

$$y(t) = y^* e^{-at} \tag{210}$$

$$v(t) = \frac{v^*(ue^{-at} - ae^{-ut})}{u - a}. \tag{211}$$

Die Population der infizierten Zellen klingt exponentiell ab und nach einer gewissen Zeit macht die Zahl der Virusteilchen das gleiche. Machen wir die plausible Annahme, dass die freien Virusteilchen schneller eliminiert werden als die infizierten Zellen sterben, d. h. $u > a$. Dann ist die Zahl der Virusteilchen ungefähr proportional e^{-at}.

Ein Fall der etwas komplizierter ist bekommt man, wenn man ein Medikament betrachtet, das dafür sorgt, dass neu produzierte Virusteilchen defekt sind und keine neuen Zellen infizieren können. Bezeichnen wir die Population der defekten Virusteichen mit w während v die Population der funktionsfähigen Virusteilchen ist. Dann gelten die Gleichungen

$$\dot{y} = \beta xv - ay, \tag{212}$$

$$\dot{v} = -uv, \tag{213}$$

$$\dot{w} = ky - uw. \tag{214}$$

Diese Gleichungen sind nicht mehr von der Gleichung für x entkoppelt. Wenn wir aber ein Zeitintervall betrachten, in dem x sich wenig ändert, dann können wir x durch eine Konstante ersetzen. Dann lassen sich die so definierten Gleichungen explizit lösen, mit dem Ergebnis

$$y(t) = \frac{y^*(ue^{-at} - ae^{-ut})}{u - a}, \tag{215}$$

$$v(t) = v^* e^{-ut}, \tag{216}$$

$$w(t) = v^* \left[(e^{-at} - e^{-ut}) \frac{u}{u - a} - ate^{-ut} \right] \frac{u}{u - a}. \tag{217}$$

Unter der Annahme, dass $u > a$ ist die Gesamtzahl der Virusteilchen $v + w$ nach einer gewissen Zeit ungefähr proportional e^{-at}. Die Halbwertszeit für die infizierten Zellen ist für beide Arten von Medikamenten gleich. Die Reverse-Transkriptase-Hemmer sind von der ersten Art, die Protease-Hemmer von der zweiten Art. Beide Möglichkeiten wurden in den bahnbrechenden Experimenten zu diesem Thema probiert.

In einem Fall sah das Experiment so aus. Zwanzig HIV-Infizierte wurden mit einem Protease-Hemmer behandelt und die Konzentration der Virus-Teilchen wurde etwa alle vier Tage gemessen. (Vor der Behandlung war die Konzentration etwa konstant.) Wenn diese Konzentration tatsächlich exponentiell abklingen sollte, dann wäre $\log v$ eine lineare Funktion von t und man könnte den Exponenten von der Steigung dieser Kurve ablesen. In den Daten finden man eine solche Gerade und man stellt fest, dass die Halbwertzeit für die infizierten Zellen etwa zwei Tage beträgt. Unter den gegebenen Annahmen müsste dann die Halbwertzeit der Virusteilchen noch kleiner sein. Später hat man Experimente durchgeführt, wo man die Viruskonzentration kurz nach Anfang der Behandlung wesentlich häufiger misst (alle zwei Stunden). Auf diese Weise kann man auch den Parameter u messen. Damit hat man eine Abschätzung, wie schnell die Virusteilchen vor der Behandlung aus dem System entfernt werden und, entsprechend, wie schnell sie produziert werden. Man kommt auf eine Zahl der Größenordnung 10^9 pro Tag.

15 Quellen

Die meisten Themen, die in diesem Skript vorkommen werden an vielen Stellen in der Literatur behandelt. Hier werden die Quellen aufgelistet auf die wir uns am meisten gestützt haben bei der Vorbereitung dieser Vorlesung. Die Hauptquelle für die Abschnitte 2, 4, 7, 9 und 10 ist [4]. Der größte Unterschied hier ist, dass wir uns auf den Fall konzentrieren in dem es Eindeutigkeit gibt. Dadurch werden manche Beweise einfacher und, so scheint es dem Verfasser, transparenter. Die Hauptquelle für Teile der Abschnitte 5 und 11 ist [3]. Die Hauptquelle für den Abschnitt 6 ist [10]. Die Hautquellen für die Abschnitte 12 bzw. 13 bzw. 14 sind [7] bzw. [13] bzw. [8].

Ich möchte mich bei Gerrit Pfluger bedanken, der mich auf einen Fehler in einer früheren Fassung dieses Skripts aufmerksam gemacht hat.

References

[1] Carr, J. 1981 Applications of centre manifold theory. Springer, Berlin.

[2] De Leenheer, P. und Smith, H. 2003 Virus dynamics: a global analysis. SIAM J. Appl. Math. 63, 1313–1327.

[3] Hale, J. K. 2009 Ordinary Differential Equations. Dover, Mineola.

[4] Hartman, P. 1982 Ordinary Differential Equations. Birkhäuser, Basel.

[5] Korobeinikov, A. 2004 Global properties of basic virus dynamics models. Bull. Math. Biol. 66, 879–883.

[6] Korobeinikov, A. 2004 Lyapunov functions and global properties for SEIR and SEIS epidemic models. Math. Medicine and Biol. 21, 75–83.

[7] Kuznetsov, Y. A. 2010 Elements of Applied Bifurcation Theory. Springer, Berlin.

[8] Nowak, M. A. und May, R. A. 2000 Virus Dynamics. Oxford University Press, Oxford.

[9] Perko, L. 2001 Differential equations and dynamical systems. Springer, Berlin.

[10] Rudin, W. 1987 Real and complex analysis. McGraw-Hill, New York.

[11] Schnakenberg, J. 1979 Simple chemical reaction systems with limit cycle behaviour. J. Theor. Biol. 81, 389–400.

[12] Selkov, E. E. 1968 Self-oscillations in glycolysis. I A simple kinetic model. Eur. J. Biochem. 4, 79–86.

[13] Strogatz, S. H. 1994 Nonlinear dynamics and chaos. Perseus, Cambridge.