Skriptum zur Vorlesung

Gewöhnliche Differentialgleichungen 2

Christian Schmeiser[1]

1 Einleitung und Voraussetzungen

Der Name dieser Vorlesung deutet auf die Existenz eines ersten Teiles (GDG1) hin, auf den im Folgenden einige Male verwiesen wird. Dabei werden Resultate zitiert, die in Grundvorlesungen über gewöhnliche Differentialgleichungen üblicherweise präsentiert werden (so auch in der im SS 2005 von Roland Steinbauer gehaltenen Vorlesung *Gewöhnliche Differentialgleichungen 1*). Abgesehen von den Beweisen dieser Resultate und von einigen Grundlagen aus der reellen Analysis sollte diese Vorlesung einigermaßen selbstkonsistent sein.

Eine Gleichung der Form

$$g\left(t, x, \frac{dx}{dt}, \frac{d^2}{dt^2}, \dots, \frac{d^n x}{dt^n}\right) = 0$$

mit $g : \mathbb{R}^{n+2} \to \mathbb{R}$ nennt man eine *implizite gewöhnliche Differentialgleichung n-ter Ordnung* für eine reellwertige Funktion $x(t)$ der reellen Veränderlichen t. 'n-ter Ordnung' allerdings nur dann, wenn die Funktion g von der n-ten Ableitung $\frac{d^n x}{dt^n}$ auch 'wirklich abhängt'. Bevor wir noch nachzudenken beginnen, was das genau bedeutet, beschränken wir uns auf *explizite gewöhnliche Differentialgleichungen n-ter Ordnung* der Form

$$\frac{d^n x}{dt^n} = g\left(t, x, \frac{dx}{dt}, \frac{d^2 x}{dt^2}, \dots, \frac{d^{n-1} x}{dt^{n-1}}\right) .$$

Jede solche Gleichung kann mit Hilfe der Definition $y_i = \frac{d^{i-1}x}{dt^{i-1}}$, $i = 1, \dots, n$ als *explizites System erster Ordnung* der Form

$$\dot{y} = f(t, y) \tag{1}$$

[1]Institut für Mathematik, Universität Wien, Nordbergstraße 15, 1090 Wien, Austria. christian.schmeiser@univie.ac.at

(mit $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ und $\dot{y} := dy/dt$) für die vektorwertige Funktion $y(t) = (y_1(t), \ldots, y_n(t))$ geschrieben werden. Bei allgemeinen Überlegungen werden wir im Folgenden explizite Systeme erster Ordnung betrachten.

Sei $I \subset \mathbb{R}$ ein nichtleeres Intervall und $y : I \to \mathbb{R}^n$ eine in I differenzierbare Funktion. Dann nennen wir y eine *Lösung* der Differentialgleichung (1) auf dem Intervall I, wenn $\dot{y}(t) = f(t, y(t))$ für alle $t \in I$ gilt.

Gleichungen der Form (1) haben üblicherweise viele Lösungen. Unter gewissen Bedingungen an die 'rechte Seite' f kann durch eine *Anfangsbedingung*

$$y(t_0) = y_0 \tag{2}$$

eine eindeutige Lösung festgelegt werden. Ein Beispiel für ein derartiges Resultat ist der *Satz von Picard-Lindelöf* (Beweis siehe GDG1):

Satz 1 *Sei $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$. Sei $f(t, y)$ in einer Umgebung von (t_0, y_0) definiert (mit Werten in $\mathbb{R}^n$) und dort stetig als Funktion von t und lipschitzstetig als Funktion von y. Dann gibt es ein offenes Intervall $I \ni t_0$ und eine eindeutige auf dem Intervall I stetig differenzierbare Lösung y von (1), die die Anfangsbedingung (2) erfüllt.*

Hat die oben erwähnte Umgebung die Form $U = (t_0 - \Delta t, t_0 + \Delta t) \times K_r(y_0)$, dann kann das Existenzintervall in der Form $I = (t_0 - T, t_0 + T)$ gewählt werden, wobei T von Δt, r und $\sup_U |f|$ abhängt.

Bemerkung 1 *Mit $K_r(y)$ bezeichnen wir die offene Kugel mit Mittelpunkt y und Radius r im $\mathbb{R}^n$.*

In dieser Vorlesung werden gewöhnliche Differentialgleichungen als Realisierungen *kontinuierlicher dynamischer Systeme* betrachtet. Deswegen werden wir die unabhängige Variable t als 'Zeit' bezeichnen. Von besonderem Interesse werden zeitinvariante dynamische Systeme sein, die durch *autonome* Differentialgleichungen beschrieben werden, deren rechte Seite nicht explizit von der Zeit abhängt:

$$\dot{y} = f(y)\,. \tag{3}$$

In diesem Fall kann durch eine Zeitverschiebung im *Anfangswertproblem* (2), (3) immer $t_0 = 0$ erreicht werden:

$$y(0) = y_0\,. \tag{4}$$

Ein Punkt im $\mathbb{R}^n$ (wie z.B. der Anfangswert y_0) wird als *Zustand* des dynamischen Systems bezeichnet. Die Menge aller möglichen Zustände (d.h.

der $\mathbb{R}^n$) heißt *Zustands-* oder *Phasenraum.* Sei $y(t)$, $t \in I$, eine Lösung der Differentialgleichung (3). Dann ist $\{y(t) : t \in I\}$ eine (durch die Zeit t parametrisierte) Kurve im Phasenraum, genannt *Trajektorie.* Beim Studium dynamischer Systeme sind nicht so sehr einzelne Trajektorien von Interesse, als vielmehr die Gesamtheit aller Trajektorien, das sogenannte *Phasenporträt.*

Wir wollen uns in dieser Vorlesung nicht mit Problemen herumschlagen, die durch Glattheitsmängel in der rechten Seite $f(y)$ verursacht werden und daher

$$f \in C^\infty(\mathbb{R}^n)^n \tag{5}$$

annehmen, d.h. dass jede Komponente der vektorwertigen Funktion f nach jeder Komponente des Argumentes $y \in \mathbb{R}^n$ beliebig oft partiell differenzierbar ist. Für die meisten der folgenden Resultate wäre eine endliche Differenzierbarkeitsordnung ausreichend, was aus den Beweisen zumeist leicht ersichtlich sein wird. Eine erste Konsequenz aus der Annahme (5) und aus dem Satz von Picard-Lindelöf ist, dass das Anfangswertproblem (3), (4) für alle $y_0 \in \mathbb{R}^n$ zumindest *lokal*, d.h. in einem beschränkten Zeitintervall (das klein genug ist), eindeutig lösbar ist.

Ist die Tatsache, dass Existenz nicht für alle Zeiten garantiert werden kann, eine Schwäche des Satzes von Picard-Lindelöf? Das Beispiel

$$\dot{y} = y^2\,, \qquad y(0) = 1\,,$$

zeigt, dass im allgemeinen kein besseres Resultat möglich ist. Die rechte Seite $f(y) = y^2$ erfüllt (5), die Lösung $y(t) = \frac{1}{1-t}$ existiert offensichtlich nur für $t < 1$. Das *maximale Existenzintervall*, d.h. das größtmögliche Intervall, in dem die Lösung existiert, ist $(-\infty, 1)$. Wir beobachten, dass es offen ist und dass $\lim_{t\to 1-} |y(t)| = \infty$ gilt. Das folgende Resultat zeigt, dass die Annahme (5) dafür ausreicht, dass nichts 'Schlimmeres' passieren kann.

Satz 2 *Sei (5) erfüllt und $y_0 \in \mathbb{R}^n$. Dann ist das maximale Existenzintervall I der eindeutigen Lösung des Problems (3), (4) offen, d.h. $I = (a, b)$ mit $-\infty \le a < 0 < b \le \infty$. In den Fällen $a > -\infty$ bzw. $b < \infty$ gilt*

$$\lim_{t\to a+} |y(t)| = \infty \qquad \textit{bzw.} \qquad \lim_{t\to b-} |y(t)| = \infty\,.$$

Bemerkung 2 *Mit $|\cdot|$ bezeichnen wir die euklidische Norm im $\mathbb{R}^n$, d.h. $|y|^2 = y \cdot y$.*

Beweis: Im Fall $I = \mathbb{R}$ ist nichts zu beweisen. Sei also zunächst $b < \infty$. Angenommen, $\lim_{t \to b-} |y(t)| = \infty$ gilt nicht. Dann gibt es eine Folge $t_n \to b-$, sodass die Folge $y(t_n)$ beschränkt ist und daher eine konvergente Teilfolge $y(t_{n_k}) \to \overline{y}$ besitzt (Satz von Bolzano-Weierstrass). Aus dem Satz von Picard-Lindelöf folgt, dass für eine Umgebung U des Punktes $(b, \overline{y})$ ein $T > 0$ existiert, sodass für alle $(\tilde{t}, \tilde{y})$ in U die Lösung des Anfangswertproblems (3), $y(\tilde{t}) = \tilde{y}$, im Intervall $(\tilde{t} - T, \tilde{t} + T)$ existiert. Da die Folge $(t_{n_k}, y(t_{n_k}))$ gegen $(b, \overline{y})$ konvergiert, existiert ein n_k, sodass $(t_{n_k}, y(t_{n_k})) \in U$ und $b - t_{n_k} < T$ gilt. Die Lösung kann daher bis zum Zeitpunkt $t_{n_k} + T > b$ fortgesetzt werden, was ein Widerspruch gegen die Tatsache ist, dass b das rechte Ende des Existenzintervalles ist. Die Offenheit des Existenzintervalles bei b ist eine triviale Konsequenz.

Das linke Ende behandelt man analog. ∎

Mit Hilfe dieses Resultates kann man in manchen Fällen leicht *globale Existenz* von Lösungen zeigen, d.h. dass die Lösung für alle Zeiten existiert. Ein oft gebrauchtes Hilfsresultat dafür ist das *Gronwall-Lemma* (Beweis siehe GDG1):

Satz 3 *Sei* $z : [0, T] \to [0, \infty)$ *stetig und es gelte die Ungleichung*

$$z(t) \leq z_0 + \lambda \int_0^t z(s) ds \,, \qquad 0 \leq t \leq T \,.$$

Dann gilt $z(t) \leq e^{\lambda t} z_0$, $0 \leq t \leq T$.

Der folgende Satz ist ein typisches globales Existenzresultat.

Satz 4 *Seien die Annahmen von Satz 2 erfüllt und sei das Wachstum der rechten Seite höchstens linear, d.h. es existieren* $\lambda, \mu \geq 0$ *sodass* $|f(y)| \leq \lambda |y| + \mu$ *für alle* $y \in \mathbb{R}^n$. *Dann existiert die Lösung des Anfangswertproblems (3), (4) global.*

Beweis: Aus der Integralgleichungsdarstellung

$$y(t) = y_0 + \int_0^t f(y(s)) ds$$

des Anfangswertproblems folgt

$$|y(t)| \leq |y_0| + \int_0^t (\lambda |y(s)| + \mu) ds$$

Im Fall $\lambda = 0$ gibt das $|y(t)| \le |y_0| + t\mu$. Für $\lambda > 0$ verwenden wir das Gronwall-Lemma mit $z(t) = |y(t)| + \mu/\lambda$ und erhalten

$$|y(t)| \le e^{\lambda t}|y_0| + \frac{\mu}{\lambda}(e^{\lambda t} - 1)\,.$$

In beiden Fällen ist es ausgeschlossen, dass $|y(t)|$ zu einem endlichen Zeitpunkt unbeschränkt wächst, und die Lösung ist daher als Konsequenz aus Satz 2 global. ∎

Für Phasenporträts hat Satz 2 die Konsequenz, dass Schwierigkeiten aufgrund mangelnder globaler Existenz in beschränkten Teilbereichen des Phasenraumes nicht sichtbar werden. Für die lokale Struktur des Phasenporträts sind *stationäre Punkte* bedeutsam, d.h. Lösungen der Gleichung $f(y) = 0$. Jeder Punkt des Phasenraumes liegt auf genau einer Trajektorie. Für stationäre Punkte besteht diese nur aus einem Punkt. Umgebungen jedes anderen Punktes werden durch eine schlichten Schar glatter Trajektorien überdeckt.

Die nächsten 3 Abschnitte beschäftigen sich mit ein-, zwei- und dreidimensionalen dynamischen Systemen, d.h. mit den Fällen $n = 1$, $n = 2$ und $n = 3$. Es wird sich zeigen, dass mit wachsender Dimension auch mit wachsender Komplexität des dynamischen Verhaltens gerechnet werden muss. Andererseits wird der letzte Abschnitt über höherdimensionale Systeme zeigen, dass auch dort oft ein 'niedrigdimensionales Verhalten' vorliegt.

2 Eindimensionale dynamische Systeme – Stabilität

Eine Differentialgleichung der Form $\dot{y} = f(y)$ mit $f : \mathbb{R} \to \mathbb{R}$ kann im Prinzip explizit gelöst werden: Die Lösung des Anfangswertproblems mit der Anfangsbedingung $y(0) = y_0$ ist implizit gegeben durch

$$\int_{y_0}^{y(t)} \frac{d\eta}{f(\eta)} = t\,.$$

Das qualitative Verhalten der Lösung für verschiedene Anfangswerte y_0 ist oft nicht leicht aus dieser Formel ersichtlich, insbesondere, wenn man keine explizite Stammfunktion von $1/f(y)$ kennt. Andererseits ist das qualitative Verhalten viel leichter direkt aus dem Verhalten von $f(y)$ ablesbar. Wegen der Stetigkeit von f lässt sich der Phasenraum (d.h. die y-Achse) darstellen als Vereinigung von stationären Punkten ($f(y) = 0$) einerseits und von offenen Intervallen andererseits, auf denen f das Vorzeichen nicht wechselt. Ist zum Beispiel f in dem Intervall (y_1, y_2) positiv mit $f(y_1) = f(y_2) = 0$, dann existiert für jedes $y_0 \in (y_1, y_2)$ die Lösung des Anfangswertproblems für alle Zeiten, und es gilt

$$\lim_{t\to-\infty} y(t) = y_1\,, \quad \text{und} \quad \lim_{t\to\infty} y(t) = y_2\,.$$

Ist f in (y_1, ∞) positiv mit $f(y_1) = 0$, dann gilt für $y_0 > y_1$ ebenso $\lim_{t\to-\infty} y(t) = y_1$. Andererseits gilt $\lim_{t\to T} y(t) = \infty$, wobei $T \le \infty$ die rechte Grenze des maximalen Existenzintervalles ist. Analoge Aussagen gelten auch für alle anderen möglichen Fälle.

Ist y_0 ein stationärer Punkt, dann gilt natürlich $y(t) = y_0$ für alle $t \in \mathbb{R}$. In den meisten Anwendungen muss man allerdings damit rechnen, dass die einfache Differentialgleichung $\dot{y} = f(y)$ die Realität nicht vollständig beschreibt und daher immer wieder kleine Störungen auftreten, die bewirken, dass man sich vom Zustand y_0 entfernt. Wenn das keinen großen Effekt hervorruft, bezeichnet man den stationären Zustand als stabil. Eine präzisere Definition lautet:

Definition 1 *Sei $\overline{y}$ ein stationärer Zustand des dynamischen Systems $\dot{y} = f(y)$, d.h. $f(\overline{y}) = 0$. Dann heißt $\overline{y}$* stabil*, wenn für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass für alle Anfangsdaten y_0 mit $|y_0 - \overline{y}| < \delta$ die Lösung des Anfangswertproblemes für alle $t > 0$ existiert und $|y(t) - \overline{y}| < \varepsilon$ für alle $t > 0$ gilt. Gilt darüber hinaus*

$$\lim_{t\to\infty} y(t) = \overline{y}\,,$$

dann heißt $\overline{y}$ asymptotisch stabil*. Ist $\overline{y}$ nicht stabil, dann heißt es* instabil*.*

Es ist leicht einzusehen, wie das Verhalten von f die Stabilität eines stationären Punktes bestimmt. Ist ein stationärer Punkt $\overline{y}$ isoliert und das Vorzeichen von f wechselt an $\overline{y}$ von 'Plus' zu 'Minus', dann ist $\overline{y}$ asymptotisch stabil. Isolierte stationäre Punkte mit anderem Verhalten von f in ihrer Umgebung sind instabil. Liegt ein stationärer Punkt im Inneren eines Intervalles, in dem f verschwindet, dann ist er stabil, aber nicht asymptotisch stabil.

Es ist eine naheliegende Idee, die Stabilität mit Hilfe der am stationären Punkt *linearisierten Gleichung* zu untersuchen. Taylorentwicklung der rechten Seite gibt $f(y) \approx f(\overline{y}) + f'(\overline{y})(y - \overline{y}) = f'(\overline{y})z$ mit dem 'Fehler' $z = y - \overline{y}$, und daher die Linearisierung $\dot{z} = f'(\overline{y})z$. Die Stabilität des stationären Punktes $z = 0$ kann man leicht aus dem Vorzeichen von $f'(\overline{y})$ ablesen und daraus auf die Stabilität von $\overline{y}$ schließen, allerdings nur, wenn $f'(\overline{y}) \neq 0$ gilt (einen solchen stationären Punkt nennen wir *hyperbolisch*):

Satz 5 *Sei $\overline{y}$ ein stationärer Zustand des dynamischen Systems $\dot{y} = f(y)$. Gilt $f'(\overline{y}) < 0$, dann ist $\overline{y}$ asymptotisch stabil. Im Fall $f'(\overline{y}) > 0$ ist $\overline{y}$ instabil. Gilt $f'(\overline{y}) = 0$, dann kann aus der Linearisierung keine Aussage über die Stabilität von $\overline{y}$ abgeleitet werden.*

Der Beweis sei dem Leser überlassen.

Verzweigungen

Eindimensionale Dynamik ist also recht langweilig. Lösungen konvergieren entweder gegen stationäre Punkte oder verabschieden sich Richtung $\pm\infty$. Nun wollen wir allerdings unsere Fragestellung verallgemeinern und Familien von dynamischen Systemen betrachten, indem wir die rechte Seite der Differentialgleichung von *Parametern* abhängig machen. Beginnen wir mit eindimensionalen Sytemen, die von einem skalaren Parameter $r \in \mathbb{R}$ abhängen, d.h. wir betrachten Differentialgleichungen der Form

$$\dot{y} = f(y, r) \qquad \text{mit } f \in C^\infty(\mathbb{R}^2)\,.$$

Die rechte Seite soll also nicht nur vom Zustand y, sondern auch vom Parameter r auf glatte Art und Weise abhängen. Unser Interesse gilt Situationen, in denen Änderungen des Parameterwertes zu *qualitativen* Änderungen in der Dynamik führen, d.h. zu Änderungen in der Anzahl, der Anordnung bzw. der Stabilität stationärer Punkte. Wir sagen, dass am Parameterwert $r = r_0$ eine *Verzweigung* auftritt, wenn r_0 Parameterbereiche mit qualitativ unterschiedlicher Dynamik trennt.

Zunächst beschreiben wir Situationen, in denen Verzweigungen ausgeschlossen sind.

Satz 6 *Sei $B \subset \mathbb{R}$ eine beschränkte offene Teilmenge des Phasenraumes und $r_0 \in \mathbb{R}$ ein Parameterwert. Seien alle in B liegenden stationären Punkte ($\overline{y}$) des dynamischen Systems $\dot{y} = f(y, r_0)$ hyperbolisch ($f(\overline{y}, r_0) = 0$, $\partial f/\partial y(\overline{y}, r_0) \neq 0$) und es gäbe keine stationären Punkte auf ∂B. Dann gibt es endlich viele davon, und es existiert ein $\varepsilon > 0$ sodass für $|r - r_0| < \varepsilon$ das dynamische System $\dot{y} = f(y, r)$ gleich viele stationäre Punkte hat, die auch alle hyperbolisch sind und die dieselben Stabilitätseigenschaften wie die entsprechenden stationären Punkte von $\dot{y} = f(y, r_0)$ besitzen.*

Beweis: Nach dem Hauptsatz über implizite Funktionen sind hyperbolische stationäre Punkte isoliert. Gäbe es unendlich viele in der beschränkten Menge B, dann müssten sie einen Häufungspunkt besitzen. Der wäre wegen der Stetigkeit von f dann auch ein stationärer Punkt, der allerdings nicht isoliert ist und daher nicht hyperbolisch sein könnte.

Aus dem Hauptsatz für implizite Funktionen folgt auch, dass für jeden stationären Punkt $\overline{y}$ und für r nahe genug bei r_0 eine Lösung $\tilde{y}$ der Gleichung $f(\tilde{y}, r) = 0$ existiert, für die $\partial f / \partial y(\tilde{y}, r) \neq 0$ gilt, d.h. dass $\tilde{y}$ ein hyperbolischer stationärer Punkt von $\dot{y} = f(y, r)$ ist. Der Hauptsatz über implizite Funktionen sagt weiters, dass der stationäre Punkt $\tilde{y}$ in einer Umgebung von $\overline{y}$ eindeutig ist. Teilt man also B in kleine Umgebungen der stationären Punkte von $\dot{y} = f(y, r_0)$ und den Rest, dann folgt aus der Vorzeichenbeständigkeit stetiger Funktionen, dass für r nahe bei r_0 im Rest keine stationären Punkte von $\dot{y} = f(y, r)$ existieren, und dass es in jeder der Umgebungen genau einen hyperbolischen stationären Punkt gibt, der dieselben Stabilitätseigenschaften wie der entsprechende stationäre Punkt von $\dot{y} = f(y, r_0)$ besitzt. ■

Die Dynamik von eindimensionalen Systemen, die nur hyperbolische stationäre Punkte besitzen, wird also durch kleine Störungen qualitativ nicht verändert. Man nennt diese Eigenschaft *Strukturstabilität.* Man sagt auch, die Eigenschaft eines dynamischen Systems, nur hyperbolische stationäre Punkte zu besitzen, ist *generisch.* Um Verzweigungen zu studieren, müssen wir also für den kritischen Parameterwert r_0 die Existenz mindestens eines nichthyperbolischen stationären Punktes $\overline{y}$ annehmen. Im Folgenden werden wir o.B.d.A. immer $\overline{y} = r_0 = 0$ setzen.

Die Falte

Nehmen wir also an, dass für $r = 0$ an der Stelle $y = 0$ ein nichthyperbolischer stationärer Punkt auftritt. Die Taylorentwicklung der rechten Seite bezüglich y und r hat dann die Gestalt

$$f(y, r) = a_{01} r + a_{20} y^2 + a_{11} r y + a_{02} r^2 + O(y^3 + r^3)\,. \tag{6}$$

Ein einfaches Beispiel ist die Familie von dynamischen Systemen

$$\dot{y} = r + y^2\,. \tag{7}$$

Die Verzweigung, die hier bei $r = 0$ auftritt, kann man so beschreiben: Für $r < 0$ gibt es zwei hyperbolische stationäre Punkte, und zwar den instabilen Punkt $y = \sqrt{-r}$ und den asymptotisch stabilen Punkt $y = -\sqrt{-r}$. Die beiden verschmelzen für $r = 0$, und für positive r gibt es keinen stationären Punkt.

Diese Verzweigung wird in der Literatur *Falte* oder auch *saddle-node*-Verzweigung genannt. Warum gleich 2 Namen für ein einfaches spezielles Beispiel? Nun, das Beispiel ist nicht so speziell, wie es den Anschein hat. Kehren wir zurück zu der allgemeinen Taylorentwicklung (6) und betrachten wir die generische Situation, dass die ersten beiden Koeffizienten a_{01} und a_{20} verschieden von Null sind. Ich behaupte, dass dann durch eine Neudefinition des Parameters und eine Transformation im Zustandsraum das dynamische System mit der rechten Seite (6) in der Form (7) geschrieben werden kann. Im Folgenden wird diese Behauptung zwar nicht vollständig bewiesen aber hoffentlich glaubwürdig gemacht.

Als ersten Schritt ersetzen wir y durch $\frac{y}{a_{20}}$ und r durch $\frac{r}{a_{01}a_{20}}$. Das bringt das dynamische System mit der rechten Seite (6) in die Form

$$\dot{y} = r + y^2 + a_{11}ry + a_{02}r^2 + O(y^3 + r^3)\,,$$

wobei die Koeffizienten umbenannt wurden. Der Anfang des Taylorpolynoms hat also schon die Form (7). Ein heuristisches Argument wäre nun, dass $a_{11}ry$, $a_{02}r^2$ und $O(r^3)$ klein sind im Vergleich zu r und dass $O(y^3)$ klein ist im Vergleich zu y^2 und alle diese Terme daher vernachlässigt werden können. Das stimmt insofern, als durch 'identitätsnahe' Transformationen von r und y alle außer den ersten beiden Termen auf der rechten Seite eliminiert werden können. Wir werden nur die Elimination der beiden quadratischen Terme demonstrieren. Wir setzen

$$r = R + bR^2\,, \qquad y = Y + cY^2\,. \tag{8}$$

Man muss ein bisschen rechnen, bis man das transformierte System ermittelt hat:

$$\dot{Y} = R + Y^2 + (a_{11} - 2c)RY + (a_{02} + b)R^2 + O(Y^3 + R^3)\,.$$

Durch die Wahl $b = -a_{02}$ und $c = a_{11}/2$ produziert man (7) bis auf einen Rest dritter Ordnung. Ersetzt man die quadratischen Polynome auf den rechten Seiten in (8) durch vollständige Taylorentwicklungen, dann lässt sich exakt die Form (7) erzeugen (siehe, z.B., [1]).

Das bedeutet, dass in jedem dynamischen System mit einer rechten Seite der Form (6) bei $r = 0$ eine Falte auftritt, wenn die beiden Koeffizienten a_{01} und a_{20} nicht verschwinden. Die Gleichung (7) nennt man eine *Normalform* der Falte. Wir werden in Zukunft auch Normalformen anderer Verzweigungen analysieren, ohne jedes Mal die Transformation auf Normalform zu diskutieren.

Die transkritische Verzweigung

Die Falte ist *die* generische Verzweigung in eindimensionalen dynamischen Systemen. Andere Arten von Verzweigungen können dann auftreten, wenn das System spezielle Eigenschaften hat, die auch bei Variation von Parametern erhalten bleiben. Eine typische Eigenschaft dieser Art ist es, dass es einen ausgezeichneten stationären Punkt gibt, der immer erhalten bleibt, also z.B. $\bar{y} = 0$. In diesem Fall müssen in der allgemeinen Taylorentwicklung (6) die Koeffizienten a_{01} und a_{02} verschwinden. Nimmt man, abgesehen davon, generisches Verhalten an, dann gilt $a_{20}, a_{11} \neq 0$. Eine entsprechende Normalform ist

$$\dot{y} = ry - y^2 \,.$$

Die dadurch gegebene *transkritische Verzweigung* hat die folgenden Eigenschaften: Sowohl für $r < 0$ als auch für $r > 0$ gibt es die beiden stationären Punkte $y = 0$ und $y = r$. Für $r < 0$ ist $y = 0$ asymptotisch stabil und $y = r$ instabil; für $r > 0$ ist es umgekehrt. Bei der Verzweigung findet also ein *Stabilitätsaustausch* statt.

Die Heugabel

Eine weitere Möglichkeit für spezielle Eigenschaften eines dynamischen Systems, die auch bei Parameteränderungen erhalten bleiben, sind *Symmetrien.* Verlangen wir z.B., dass das dynamische System sich nicht verändert, wenn wir y durch $-y$ ersetzen, dann führt das auf die Forderung, dass die rechte Seite eine ungerade Funktion von y ist. Die Normalform der dieser Annahme entsprechenden *Heugabelverzweigung* ist

$$\dot{y} = ry - y^3 \,.$$

Sie hat die folgenden Eigenschaften: Für $r < 0$ gibt es einen stationären Punkt, und zwar $y = 0$, und dieser ist asymptotisch stabil. Für positive r verliert $y = 0$ seine Stabilität, die auf die beiden neuen stationären Punkte $y = \pm\sqrt{r}$ übergeht. Aufgrund der Symmetrie ist klar, dass zu jedem neuen stationären Punkt auch sein Spiegelbild ein stationärer Punkt mit denselben Eigenschaften sein muss.

Der 'Spruce Budworm'

Der spruce budworm ist ein nordamerikanischer Baumschädling, der immer wieder eine große Gefahr für Nadelwälder darstellt. Wir wollen die Entwicklung einer budworm-Population durch ein dynamisches System beschreiben.

Sei $N(\tau)$ ein Maß für die Größe der Population zum Zeitpunkt τ. Die Gleichung

$$\frac{dN}{d\tau} = RN\left(1 - \frac{N}{K}\right) - \frac{BN^2}{A^2 + N^2}$$

ist ein typisches Modell der *Populationsdynamik*. Der Faktor $R(1 - N/K)$ im ersten Term ist die Differenz zwischen Geburten- und Sterberate. Für sehr kleine Populationen ist diese Differenz durch die positive Konstante R gegeben. Für wachsende Populationen wird sie aufgrund von Nahrungsmangel und Konkurrenz kleiner, bis sie schließlich negativ wird, wenn die Population den kritischen Wert K überschreitet. Der zweite Term beschreibt Verluste durch natürliche Feinde. Im Fall des spruce budworms sind das Vögel. Die Vögel fressen die budworms mit einer maximalen Rate B. Die Abhängigkeit der Rate von der Größe der Population hat die folgende Interpretation: Ist die Population kleiner als der Schwellwert A, dann ist es für die Vögel zu mühsam, nach den budworms zu suchen. Sie ernähren sich dann hauptsächlich von Anderem und lassen die budworms in Ruhe. Über dem Schwellwert A werden die budworms als Vogelfutter attraktiv, und sie werden fast mit der maximalen Rate B gefressen.

Bevor das Modell als dynamisches System analysiert wird, möchten wir uns mit Fragen der *Dimensionsanalyse* beschäftigen. Will man konkret mit dem Modell rechnen (d.h. Zahlen einsetzen), muss man zunächst *Einheiten* fixieren. Diese können auf zwei *Grundeinheiten* zurückgeführt werden, eine Einheit für die Zeit, z.B. die Sekunde (abgekürzt durch 'sec'), und eine Einheit für die Populationsgröße, abgekürzt durch 'pop' (z.B. 1pop = 1000 budworms). In der folgenden Tabelle sind die Einheiten der Unbekannten N, der unabhängigen Variable τ und der vier Parameter R, K, A, und B mit Hilfe von pop und sec ausgedrückt.

Größe	Einheit
N	pop
τ	sec
R	sec^{-1}
K	pop
A	pop
B	$\mathrm{pop\,sec}^{-1}$

Diese Wahl der Grundeinheiten ist willkürlich und unter Umständen für das konkrete Modell nicht sehr sinnvoll. Eine andere Möglichkeit der Einheitenwahl ist die Verwendung von *intrinsischen Refernzgrößen* als Einheiten für die abhängigen und unabhängigen Variablen N und τ. Intrinsische

Referenzgrößen sind Enheiten, die sich aus den Parametern des Problems berechnen lassen. Für die Populationsgröße haben wir reiche Auswahl: K, A und B/R haben alle dieselbe Dimension wie N. Für die Zeit gibt es die Möglichkeiten $1/R$, K/B und A/B. Wir treffen unsere Wahl, indem wir uns auf einen der modellierten Effekte konzentrieren, und zwar auf den Einfluss der natürlichen Feinde. Daher verwenden wir die Parameter A und B bei der *Skalierung:*

$$t := \frac{\tau}{A/B}\,, \qquad y(t) := \frac{N(tA/B)}{A}\,.$$

Die neuen Variablen y und t sind dimensionslos. Die Gleichung für y lautet

$$\dot{y} = ry\left(1 - \frac{y}{k}\right) - \frac{y^2}{1+y^2}$$

mit den beiden *dimensionslosen Parametern* $r = RA/B$ und $k = K/A$. Ein wesentlicher Effekt der Verwendung intrinsischer Referenzgr̈ßen für die Skalierung ist die Reduktion der Anzahl der Parameter von vier auf zwei. Das ist eine große Erleichterung der Analyse aller möglichen qualitativen Eigenschaften des Modells.

Neben dem trivialen stationären Punkt $y = 0$ (der immer instabil ist, d.h. die budworms sterben nicht aus) gibt es noch andere stationäre Zustände, die durch die Gleichung

$$r\left(1 - \frac{y}{k}\right) = \frac{y}{1+y^2}$$

bestimmt werden. Kurvendiskussion und/oder graphische Darstellung zeigen, dass diese Gleichung je nach den Werten von r und k 1–3 positive Lösungen hat. Faltenverzweigungen treten immer dann auf, wenn die Gerade auf der linken Seite die Kurve auf der rechten Seite berührt. Aus dieser Bedingung ergibt sich ein Zusammenhang zwischen den Schnittpunkten und den Parameterwerten:

$$r = \frac{2y^3}{(1+y^2)^2}\,, \qquad k = \frac{2y^3}{y^2-1} \qquad \text{mit } y > 1\,.$$

Das kann man als Parameterdarstellung (mit Parameter y) einer Kurve in der r-k-Ebene interpretieren. Diese Kurve hat am Punkt $(r_0, k_0) = (3\sqrt{3}/8, 3\sqrt{3})$ (für den Parameterwert $y_0 = \sqrt{3}$) eine Spitze. Zwischen den beiden Ästen der Kurve gibt es drei stationäre Punkte, und außerhalb einen.

Dieser eine ist immer stabil. Wenn es drei stationäre Punkte gibt, dann sind zwei davon stabil mit einem instabilen dazwischen.

Nach diesen Resultaten ist folgendes Szenario möglich: Sei r fest und zwischen $1/2$ und $3\sqrt{3}/8$; k wachse langsam (z.B. dadurch, dass die Nadelbäume wachsen). Das gibt eine waagrechte Linie in der r-k-Ebene, die die Verzweigungskurve zweimal schneidet. Bevor sie das tut, gibt es ein stabiles Gleichgewicht mit kleinen Werten der budworm-Population. Beim ersten Überqueren der Verzweigungskurve entstehen ein großes stabiles und ein instabiles Gleichgewicht. Das kleine Gleichgewicht bleibt dabei stabil und die Population daher auf niedrigem Niveau. Beim zweiten Überqueren der Kurve passiert allerdings etwas Dramatisches: Das kleine Gleichgewicht verschmilzt mit dem instabilen, und beide verschwinden. Nun ist nur mehr das große Gleichgewicht übrig, und es kommt zu einem sprunghaften Anstieg der Population.

Dieses qualitative Verhalten kann man vollständig erklären, indem man eine kleine Umgebung des kritischen Punktes (r_0, k_0) im Parameterbereich analysiert. Eine Normalform für die dort auftretende *cusp-Verzweigung* (cusp = Spitze) ist gegeben durch

$$\dot{y} = r + ky + y^3 \,.$$

Die cusp-Verzeigung ist die generische Verzweigung, wenn man (wie bei der Heugabel) annimmt, dass auch die zweite Ableitung nach y am Verzweigungspunkt verschwindet, wenn aber (im Gegensatz zur Heugabel) keine Symmetrie gefordert wird. Für die cusp-Verzweigung sind zwei Parameter notwendig. Ein Fachausdruck dafür ist *Verzweigung der Kodimension 2* zum Unterschied der bisher behandelten Verzweigungen, die die Kodimension 1 hatten.

References

[1] J. Guckenheimer und P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, New York, 1983.

[2] S.H. Strogatz, *Nonlinear Dynamics and Chaos (with Applications to Physics, Biology, Chemistry, and Engineering)*, Addison-Wesley Publ. Comp., Reading, 1994.