

Gewöhnliche Differentialgleichungen
Wintersemester 2012/2013

Fritz Colonius
Institut für Mathematik
Universität Augsburg
fritz.colonius@math.uni-augsburg.de

5. Februar 2013

Inhaltsverzeichnis

1	Einführung	1
1.1	Anwendungsbeispiele	1
1.2	Einige einfache Resultate	5
1.3	Lösungsbegriff und Differentialgleichungen Höherer Ordnung	7
2	Existenz und Eindeutigkeit von Lösungen	9
2.1	Problemformulierung	9
2.2	Lipschitz-Stetigkeit und die Gronwall-Ungleichung	10
2.3	Eindeutigkeit von Lösungen	13
2.4	Existenz von Lösungen	15
3	Autonome Differentialgleichungen	25
3.1	Grundlegende Eigenschaften	25
3.2	Trajektorien	28
3.3	Phasenporträts und Richtungs-Vektorfeld	30
4	Skalare Nichtautonome Differentialgleichungen	33
4.1	Problemformulierung	33
4.2	Exakte Differentialgleichungen und Trennung der Variablen	34
4.3	Integrierende Faktoren	38
4.4	Weitere Lösungsmethoden	41
5	Ebene Autonome Systeme	45
5.1	Reduktion auf skalare Differentialgleichungen	45
5.2	Hamiltonsche Differentialgleichungen und erste Integrale	46
6	Lineare Differentialgleichungen	51
6.1	Nichtautonome Lineare Differentialgleichungen	51
6.2	Autonome Lineare Differentialgleichungen	58
6.3	Klassifikation in der Ebene	62
6.4	Ein Beispiel: Der Lineare Oszillator	65
7	Stabilität Autonomer Differentialgleichungen	71
7.1	Stabilität und Eigenwerte	71
7.2	Lyapunov-Funktionen	75
7.3	Lyapunov-Funktionen für das Pendel	81
7.4	Diskussion der Stabilität	83
7.5	Regelungsprobleme	85
8	Abhängigkeit von Parametern und Anfangswerten	89
8.1	Stetigkeit	89
8.2	Differenzierbarkeit	93
8.3	Verzweigung von Ruhelagen	97

Vorbemerkung

Diese Vorlesung gibt eine Einführung in die Theorie gewöhnlicher Differentialgleichungen. Sie richtet sich in erster Linie an die Studierenden des Bachelor-Studiengangs Mathematik und des Lehramt-Studiengangs Gymnasium. Als Vorkenntnisse genügen die Grundveranstaltungen Analysis I und II sowie Lineare Algebra I und II.

Von Zeit zu Zeit werden auch weiterführende Resultate erwähnt und es werden Querverbindungen zu anderen Gebieten der Mathematik aufgezeigt.

Von den zahlreichen einführenden Büchern in die Theorie von gewöhnlichen Differentialgleichungen sei hier insbesondere auf die Bücher [7] von L. Grüne und O. Junge, [15] von W. Walter, sowie [4] von B. Aulbach hingewiesen.

Die angegebenen Referenzen werden, auch ohne besonderen Hinweis im Text, intensiv benutzt.

Kapitel 1

Einführung

In erster Annäherung ist eine gewöhnliche Differentialgleichung eine Gleichung für eine Funktion einer Variablen, in der die Funktionswerte und die Ableitungen für jeden Wert des Arguments in Beziehung gesetzt werden.

Die Differential- und Integralrechnung ist historisch zusammen mit der Theorie gewöhnlicher Differentialgleichungen aus dem Versuch entstanden, Bewegungsgesetze von starren Körpern (Planeten) zu verstehen. Die Entstehung der Analysis ist also untrennbar mit der Entstehung der Theorie gewöhnlicher Differentialgleichungen verbunden. Heutzutage werden grundlegenden Vorlesungen zur Analysis und zur Theorie von Differentialgleichungen getrennt angeboten. Zugleich ist diese Theorie einer der ersten Berührungspunkte zwischen Mathematik und Physik; ihre Bedeutung reicht auch in viele andere Wissenschaften hinein.

1.1 Anwendungsbeispiele

In diesem Abschnitt sammeln wir einige Differentialgleichungen, die bei der Modellierung realer Phänomene auftreten.

Populationsmodelle. Es sei $x(t)$ die Größe einer Population zur Zeit t (zum Beispiel die Größe der Weltbevölkerung im Jahr t oder die Zahl der Atome einer radioaktiven Substanz). Sei

$$\dot{x}(t) = \frac{d}{dt}x(t)$$

die momentane Änderung zur Zeit t . Ist die Änderungsrate

$$\frac{\dot{x}(t)}{x(t)} = a$$

unabhängig von t , so erhält man

$$\dot{x}(t) = ax(t). \tag{1.1.1}$$

Dann gilt, nach elementaren Resultaten aus der Analysis,

$$x(t) = e^{at}x(0).$$

Für das Wachstum der Weltbevölkerung ist zum Beispiel für die Zeit 1700-1971 $a = 0.02$ eine gute Näherung. Kennt man hier also das Entwicklungsgesetz (1.1.1) so kann man die zeitliche Entwicklung beschreiben. Insbesondere erhält man für $a > 0$ ein exponentielles Wachstum, für $a < 0$ eine exponentielle Abnahme. Während diese Gleichung als Modell für die Entwicklung der Weltbevölkerung nur beschränkt tauglich ist, beschreibt es mit sehr großer Präzision den Zerfall radioaktiver Stoffe.

Ein verfeinertes Populationsmodell erhält man, wenn die Zuwachsrate a von der Populationsgröße abhängt, $a = a(x)$, also

$$\dot{x} = a(x)x$$

(der Kürze halber lassen wir hier das Argument t weg).

Zum Beispiel wurde von Verhulst (1837) das folgende Modell für die Entwicklung der Weltbevölkerung vorgeschlagen:

$$a(x) = a_0 - bx \text{ mit } a_0, b > 0,$$

also

$$\dot{x} = a_0x - bx^2.$$

Dann ist $x(t) \equiv \frac{a_0}{b}$ eine konstante Lösung, ebenso $x(t) \equiv 0$. Für $0 < x < \frac{a_0}{b}$ ist $a_0x - bx^2 > 0$, also ist $\dot{x} > 0$ und daher x hier monoton wachsend. Ferner gilt für die zweite Ableitung

$$\ddot{x} := \frac{d^2}{dt^2}x = a_0\dot{x} - 2bx\dot{x} = (a_0 - 2bx)(a_0 - bx)x.$$

Wegen $a_0 - bx > 0$ ist $\ddot{x} > 0$ für $0 < x < \frac{a_0}{2b}$ und $\ddot{x} < 0$ für $\frac{a_0}{2b} < x < \frac{a_0}{b}$. Die obige Diskussion erlaubt es, den Verlauf von Lösungen $x(t)$ in Abhängigkeit von der Zeit t zu skizzieren. Dabei kann $\frac{a_0}{b}$ als Sättigungswert interpretiert werden (zum Beispiel wegen Begrenztheit der Ressourcen).

Das folgende Modell für die zeitliche Entwicklung von zwei Spezies stammt (nach Beobachtungen von Fischpopulationen im Mittelmeer während des ersten Weltkriegs) von V. Volterra (1931).

Seien $x(t)$ und $y(t)$ die momentane Größe der Beute-Spezies bzw. der Räuber-Spezies und

$$\begin{aligned}\dot{x} &= (a - by)x \\ \dot{y} &= (-c + dx)y\end{aligned}$$

hierbei sind die Konstanten $a, b, c, d > 0$ durch Beobachtungen zu bestimmen. Die Zuwachsrate der Beute-Spezies ist

$$\frac{\dot{x}}{x} = a - by$$

wird also durch die Größe der Räuber-Population vermindert; analog wird die Zuwachsrate der Räuber-Population

$$\frac{\dot{y}}{y} = -c + dx$$

durch die Größe der Beute-Population vermehrt.

Ein weiteres verfeinertes Räuber-Beute-Modell ist zum Beispiel das Volterra-Verhulst Modell (mit Konstanten $a_1, a_2, b_1, b_2, d_1, d_2 > 0$)

$$\begin{aligned}\dot{x} &= a_1x - b_1xy - b_2x^2 \\ \dot{y} &= -a_2y + d_1xy - d_2y^2.\end{aligned}$$

Die Konkurrenz zweier Spezies um gleiche (Futter-)Ressourcen kann man folgendermaßen modellieren:

$$\begin{aligned}\dot{x} &= a_1x - b_1xy - b_2x^2 \\ \dot{y} &= a_2y - d_1xy - d_2y^2.\end{aligned}$$

Ein Problem dieser Populationsmodelle ist, dass die Populationsgröße eine natürliche Zahl ist; Modellierung durch eine Differentialgleichung macht also nur Sinn für große Populationen. Alternative Modelle benutzen Differenzgleichungen, bei denen der Differentialquotient durch einen Differenzenquotienten ersetzt wird.

Klassische (und wesentlich ältere) Beispiele von gewöhnlichen Differentialgleichungen stammen aus der *Mechanik*. Newtons Gesetz besagt, dass die Beschleunigung proportional zur ausgeübten Kraft ist, also

$$\frac{d^2}{dt^2}x =: \ddot{x} = F.$$

Ist die Kraft $F = f(x)$ abhängig von x , so erhält man eine Differentialgleichung der Form

$$\ddot{x} = f(x).$$

Ein Beispiel ist das sogenannte Zwei-Körper-Problem, bei dem die Bewegung zweier unterschiedlich großer Körper (zum Beispiel Erde und Sonne) unter dem Einfluss der Schwerkraft, die umgekehrt proportional zum Quadrat der Entfernung ist, gesucht wird.

Da sich die Bewegung im \mathbb{R}^3 abspielt, haben x und \dot{x} also im Allgemeinen jeweils drei Komponenten (zwei reichen hier aber aus). Aus diesen Bewegungsgleichungen lassen sich die Keplerschen Gesetze herleiten, die von Kepler (1609, 1619) nur aus Beobachtungen hergeleitet wurden.

- (i) Jeder Planetenorbit ist eine Ellipse, mit der Sonne in einem Brennpunkt.¹
- (ii) Eine Linie von der Sonne zu einem Planeten überdeckt in gleichen Zeiten gleiche Gebiete.
- (iii) Die Periodenlängen T (Umlaufzeiten) der Orbits verhalten sich zueinander wie $a^{3/2}$, wobei a die große Ellipsen-Halbachse ist.²

Sind also zum Beispiel T_E und T_V die Umlaufzeiten von Erde und Venus und a_E und a_V die entsprechenden Halbachsen, so gilt $\frac{T_E^2}{T_V^2} = \frac{a_E^3}{a_V^3}$.

Isaac Newton (1684) gelang es erstmals, eine dynamische Erklärung für diese Gesetze (insbesondere für das dritte Gesetz) herzuleiten. Man kann in der Tat die Keplerschen Gesetze aus der Differentialgleichung für das Zwei-Körperproblem herleiten, bei dem die momentane Kraft F proportional zu dem inversen Quadrat des Abstands r der beiden Massenpunkte ist,

$$F = \frac{gMm}{r^2};$$

hierbei ist g die Gravitationskonstante und M und m sind die beiden Massen; siehe hierzu z.B. Acheson [1, Chapter 6].

Anstelle der hierzu gehörenden, etwas komplizierten Differentialgleichungen diskutieren wir ein einfacheres mechanisches System.

Das Pendel. Wir betrachten ein Pendel, das aus einer masselosen starren Stange der Länge l und einem daran befestigten Massepunkt besteht. Das Pendel soll sich unter dem Einfluss der senkrecht nach unten wirkenden Schwerkraft befinden. Wir bezeichnen die Winkelposition des Pendels mit $\varphi \in [0, 2\pi)$, so dass $\dot{\varphi}(t)$ die Winkelgeschwindigkeit und $\ddot{\varphi}(t)$ die Winkelbeschleunigung zum Zeitpunkt t bezeichnet. Die Lage, Geschwindigkeit und Beschleunigung des Punktes M des sich auf dem Kreis mit Radius l bewegenden Punktes erhält man einfach durch Multiplikation von $\varphi(t)$, $\dot{\varphi}(t)$ und $\ddot{\varphi}(t)$ mit l . Zur Herleitung einer Differentialgleichung für φ benutzen wir wieder das Newtonsche Kraftgesetz:

$$\text{Kraft} = \text{Masse mal Beschleunigung}$$

zu jedem Zeitpunkt t . Hier ist Masse mal Beschleunigung gegeben durch

$$ml\ddot{\varphi}(t).$$

Die zur Zeit t wirkende Kraft ist die tangential Komponente der (nach unten gerichteten) Schwerkraft $-mg$, also

$$-mg \sin \varphi(t).$$

Darüberhinaus berücksichtigen wir eine Reibungskraft, die der momentanen Geschwindigkeit proportional und entgegengesetzt ist, also

$$-k\dot{\varphi}(t)$$

¹Ellipsen sind beschrieben durch $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a \geq b > 0$, Exzentrizität $e := \sqrt{1 - \frac{b^2}{a^2}}$, Fokus $(\pm ae, 0)$.

²Mit dem dritten keplerschen Gesetz konnten die relative Abstände der Planeten im Sonnensystem berechnet werden. Erst der Venustransit 1761 und 1769 ermöglichte es, die Entfernung der Erde zur Sonne und damit die Größe des Sonnensystems zu berechnen; eine spannende Darstellung liefert: Andrea Wulf, Die Jagd auf die Venus und die Vermessung des Sonnensystems, C. Bertelsmann 2012.

mit einer Proportionalitätskonstanten k (dies ist eine stark vereinfachende Annahme! Die Modellierung von Reibungskräften ist im Allgemeinen schwierig). Wir erhalten nach Newton also

$$-mg \sin \varphi(t) - k\dot{\varphi}(t) = ml\ddot{\varphi}(t)$$

oder

$$\ddot{\varphi}(t) = -\frac{k}{lm}\dot{\varphi} - \frac{g}{l}\sin \varphi.$$

In dieser Gleichung entspricht $\varphi = 0$ dem senkrecht nach unten hängenden Roboterarm und $\varphi = \pi$ dem senkrecht nach oben stehenden Roboterarm.

Nimmt man zusätzlich an, dass ein Motor am Drehpunkt angebracht ist, der eine Kraft (oder ein Drehmoment) $u(t)$ auf den Roboterarm ausübt, so erhält man ein Regelungssystem der Form

$$\ddot{\varphi} = -\frac{k}{lm}\dot{\varphi} - \frac{g}{l}\sin \varphi + u(t).$$

Solche Systeme werden in der Regelungs- oder Kontrolltheorie studiert, wo man etwa daran interessiert ist, Steuerungen $u(\cdot)$ zu finden, für die der Roboterarm eine festgelegte Bewegung ausführt; siehe etwa Sontag [14] oder Knobloch/Kappel [12]. Wir bleiben bei der ungesteuerten Version (mit $u(t) \equiv 0$). Der Kürze wegen sei $l = m = g = 1$ und wir betrachten jetzt den ungedämpften Fall, in dem $k = 0$ gesetzt wird, also

$$\ddot{\varphi} = -\sin \varphi.$$

Der folgende Trick erlaubt es, nur mit Ableitungen erster Ordnung zu arbeiten: Wir setzen

$$x = (x_1, x_2)^T := (\varphi, \dot{\varphi})^T$$

und erhalten das System

$$\dot{x}_1 = x_2, \dot{x}_2 = -\sin x_1, \text{ also } \dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\sin x_1 \end{bmatrix}.$$

Die rechte Seite $f(x) = \begin{bmatrix} x_2 \\ -\sin x_1 \end{bmatrix}$ ordnet jedem Punkt $x \in \mathbb{R}^2$ einen Vektor zu, man spricht daher von einem Vektorfeld $f(x), x \in \mathbb{R}^2$. Daraus kann man schnell einige spezielle Lösungen ablesen: In den Punkten $x_k^* = (k\pi, 0)^T$ mit $k \in \mathbb{Z}$ gilt $f(x_k^*) = 0$. Die zeitliche Ableitung der Lösung ist also gleich Null und damit sind die Lösungen zeitlich konstant. Man bezeichnet solche Lösungen als **Ruhelagen** oder **Gleichgewichtspunkte**. Geometrisch lässt sich die Differentialgleichung $\dot{x} = f(x)$ als die Forderung interpretieren, dass jede Lösungskurve in jedem Punkt tangential zum Vektorfeld f verläuft. Man kann den Verlauf einiger Lösungskurven an Hand des Vektorfelds vermuten: es gibt geschlossene Kurven und wellenförmige Kurven.

Diese Diskussion zeigt alternative Möglichkeiten, die Lösungen von $\dot{x} = f(x)$ darzustellen: In ihrem zeitlichen Verlauf, wie für das erste Beispiel, oder als Kurven in ihrem zeitlichen Verlauf im sog. Zustandsraum, hier im \mathbb{R}^2 .

Wir studieren genauer die Position $\varphi = \pi$ für

$$\ddot{\varphi} = -k\dot{\varphi} - \sin \varphi.$$

Setzen wir $\psi := \varphi - \pi$, so ändern sich die Gleichungen zu

$$\ddot{\psi}(t) = -k\dot{\psi}(t) - \sin[\psi(t) + \pi].$$

Die Ruhelage in der senkrechten Position $\psi(t) \equiv 0$ mit Geschwindigkeit $\dot{\psi} \equiv 0$ ist offenbar eine Lösung dieser Differentialgleichung. Aus physikalischen Gründen ist klar, dass diese Lösung instabil ist. Für kleine Auslenkungen kehrt das Pendel nicht in die Ruhelage zurück. Ein wesentlicher Teil der Theorie von Differentialgleichungen beschäftigt sich damit, dieses Stabilitätsverhalten zu analysieren. Damit werden wir uns in der zweiten Semesterhälfte beschäftigen und beschränken uns

hier auf die folgenden Bemerkungen. Der Sinus-Term ist nichtlinear; wir können aber ein "lineares" System herleiten, indem wir für "kleine" Abweichungen $\psi \approx 0$ die Approximation erster Ordnung

$$\sin[\pi + \psi] \approx \sin \pi + \left(\frac{d}{dx} \Big|_{x=\pi} \sin x \right) \psi = -\psi$$

benutzen. Man erhält

$$\ddot{\psi}(t) + k\dot{\psi}(t) - \psi(t) = 0.$$

Schreiben wir wieder

$$x_1 = \psi \text{ und } x_2 = \dot{\psi},$$

so erhält man

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -k \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Wir hoffen natürlich, dass diese Gleichung das Verhalten des Ausgangssystems in der Nähe der (oberen) Ruhelage annähernd beschreibt. Dieses Vorgehen (Linearisierung des Systems) werden wir später mathematisch analysieren.

Alle physikalisch-technischen Anwendungen mit räumlich inhomogenen Vorgängen werden an Stelle von gewöhnlichen Differentialgleichungen durch partielle Differentialgleichungen beschrieben. Beispiele sind die Wärmeleitungsgleichung (x bezeichnet die räumliche Variable)

$$\frac{\partial u(t, x)}{\partial t} = \frac{\partial^2 u(t, x)}{\partial x^2}, \quad (t, x) \in \mathbb{R} \times \mathbb{R},$$

oder die Wellengleichung

$$\frac{\partial^2 u(t, x)}{\partial t^2} = \frac{\partial^2 u(t, x)}{\partial x^2}, \quad (t, x) \in \mathbb{R} \times \mathbb{R}.$$

Die Theorie (und Numerik) partieller Differentialgleichung ist ein eigenes, weit verästeltes Gebiet, dem auch eigene Lehrveranstaltungen gewidmet sind. Man kann die Beschäftigung mit gewöhnlichen Differentialgleichungen auch als Vorbereitung für das Studium partieller Differentialgleichungen sehen. In dieser Vorlesung werden partielle Differentialgleichungen nur ganz am Rand auftreten.

Historische Anmerkungen zur Geschichte der Theorie gewöhnlicher Differentialgleichungen findet man zum Beispiel im ersten Kapitel des Buches Hairer/Norsett/Wanner [9]. Eine Fülle von sehr ausführlich dargestellten Anwendungsbeispielen findet man in M. Braun [6]. Das Buch [11] von Hirsch/Smale/Devaney bietet eine Einführung in gewöhnliche Differentialgleichungen, die einerseits auf Linearer Algebra aufbaut und andererseits zur allgemeineren, modernen Theorie dynamischer Systeme hinführt.

1.2 Einige einfache Resultate

Wir betrachten die folgende Klasse von Differentialgleichungen.

Definition 1.2.1 Sei $A \in \mathbb{R}^{n \times n}$. Eine differenzierbare Funktion $x : \mathbb{R} \rightarrow \mathbb{R}^n$ heißt Lösung der (linearen autonomen) Differentialgleichung in \mathbb{R}^n

$$\dot{x} = Ax, \tag{1.2.1}$$

falls für alle $t \in \mathbb{R}$ gilt

$$\frac{d}{dt} x(t) = Ax(t).$$

Aus den Grundveranstaltungen ist bekannt, dass die Matrix-Exponentialfunktion

$$e^{At} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

eine Potenzreihe in $\mathbb{R}^{n \times n}$ (identifiziert mit \mathbb{R}^{n^2}) mit Konvergenzradius ∞ ist und dass gilt:

$$e^{A0} = I, \quad e^{A(t+s)} = e^{At} e^{As} \quad \text{für alle } t, s \in \mathbb{R}.$$

Ferner dürfen wir die Reihe gliedweise differenzieren. Es folgt

$$\frac{d}{dt} e^{At} = \sum_{k=0}^{\infty} k \frac{t^{k-1}}{k!} A^k = \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k = A \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = A e^{At} = e^{At} A. \quad (1.2.2)$$

Man erhält das folgende Resultat.

Proposition 1.2.2 Sei $A \in \mathbb{R}^{n \times n}$. Für jedes $x_0 \in \mathbb{R}^n$ ist

$$x(t) = e^{At} x_0 = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} t^k \right) x_0 \quad (1.2.3)$$

die eindeutige Lösung von (1.2.1) mit $x(0) = x_0$.

Beweis. Nach (1.2.2) gilt für x aus (1.2.3)

$$\frac{d}{dt} x(t) = \left(\frac{d}{dt} e^{At} \right) x_0 = A e^{At} x_0 = A x(t).$$

Zum Beweis der Eindeutigkeit sei $y(t), t \in \mathbb{R}$, eine beliebige Lösung mit $y(0) = x_0$. Aus der Produktregel ergibt sich

$$\frac{d}{dt} e^{-At} y(t) = -e^{-At} A y(t) + e^{-At} A y(t) = 0.$$

Demnach ist $e^{-At} y(t)$ konstant in t , und wegen $e^{-A0} y(0) = x_0$ folgt also

$$e^{-At} y(t) = x_0 \quad \text{für alle } t \in \mathbb{R}.$$

Multiplikation von links mit e^{At} liefert dann $y(t) = e^{At} x_0 = x(t)$ für alle $t \in \mathbb{R}$ und damit die Behauptung. ■

Die Aussage dieser Proposition wird auch so formuliert: das Anfangswertproblem

$$\dot{x} = Ax, \quad x(0) = x_0,$$

besitzt eine eindeutige Lösung auf \mathbb{R} .

Im Folgenden diskutieren wir einige weitere Beispiele von Differentialgleichungen. Lösungen von

$$\dot{x} = 2tx \quad (1.2.4)$$

sind gegeben durch

$$x(t) = c e^{t^2}, \quad t \in \mathbb{R},$$

für Konstanten $c \in \mathbb{R}$. In der Tat gilt

$$\frac{d}{dt} x(t) = c 2t e^{t^2} = 2t x(t), \quad t \in \mathbb{R}.$$

Gleichung (1.2.4) ist ein Spezialfall der folgenden skalaren linearen Differentialgleichung

$$\dot{x} = a(t)x \quad (1.2.5)$$

mit einer stetigen Funktion $a : \mathbb{R} \rightarrow \mathbb{R}$. Dann ist die eindeutige Lösung mit der Anfangsbedingung $x(t_0) = x_0 \in \mathbb{R}$ gegeben durch

$$x(t) = e^{\int_{t_0}^t a(s)ds} x_0, \quad t \in \mathbb{R},$$

Die Existenz folgt mit der Kettenregel und die Eindeutigkeit folgt mit dem gleichen Trick wie in Proposition 1.2.2: Ist $y(t), t \in \mathbb{R}$ eine Lösung mit $y(t_0) = x_0$ so zeigt man, dass

$$\frac{d}{dt} \left(e^{-\int_{t_0}^t a(s)ds} y(t) \right) = 0, \text{ also } y(t) = e^{\int_{t_0}^t a(s)ds} x_0, t \in \mathbb{R}.$$

Lösungen von

$$\dot{x} = x^2 \quad (1.2.6)$$

sind gegeben durch

$$x(t) = -\frac{1}{t+c}, \quad t \neq -c,$$

für Konstanten $c \neq 0$, denn

$$\frac{d}{dt} x(t) = -\frac{d}{dt} (t+c)^{-1} = (t+c)^{-2} = x(t)^2.$$

Lösungen von

$$\ddot{x} = -x \quad (1.2.7)$$

sind gegeben durch

$$x(t) = a \sin t + b \cos t$$

für Konstanten $a, b \in \mathbb{R}$.

Bemerkung 1.2.3 Die Bezeichnung t für die unabhängige Variable erinnert an *tempus* oder *time*. Dies ist häufig, aber durchaus nicht immer, auch ihre Interpretation. Manchmal werden anstelle von t, x die Variablen x, y verwendet mit $\frac{d}{dx} y(x) = y'(x)$. Dann haben zum Beispiel die Gleichungen (1.2.6) und (1.2.7) die Form

$$y'(x) = y(x)^2 \text{ bzw. } y''(x) = -y(x).$$

1.3 Lösungsbegriff und Differentialgleichungen Höherer Ordnung

In diesem Abschnitt spezifizieren wir den Lösungsbegriff für Differentialgleichungen, und zeigen, wie man Differentialgleichungen höherer Ordnung auf solche erster Ordnung zurückführen kann. Wenn wir im Folgenden von einem Intervall sprechen, schließen wir immer den trivialen Fall einpunktiger Intervalle aus.

Definition 1.3.1 Es sei $F : D \rightarrow \mathbb{R}^n$, wobei $D \subset \mathbb{R} \times \mathbb{R}^{mn}$ offen ist. Eine m -mal differenzierbare Funktion $x : J \rightarrow \mathbb{R}^n$, $J \subset \mathbb{R}$ ein Intervall, heißt Lösung der Differentialgleichung

$$x^{(m)} = F(t, x, \dot{x}, \ddot{x}, \dots, x^{(m-1)}), \quad (1.3.1)$$

falls für alle $t \in J$ gilt

$$x^{(m)}(t) = \frac{d^m}{dt^m} x(t) = F(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, \frac{d^{m-1}x}{dt^{m-1}}(t)).$$

Dann heißt m die Ordnung der Differentialgleichung und n ihre Dimension.

Dabei ist natürlich implizit gefordert, dass für alle $t \in I$ gilt

$$\left(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, \frac{d^{m-1}x}{dt^{m-1}}(t) \right) \in D.$$

Ist $t \in J$ ein Randpunkt des Lösungsintervalls J , so fordern wir nur, dass die entsprechenden einseitigen Ableitungen existieren und die Differentialgleichung erfüllen. Typischerweise sind Differentialgleichungen aus der Mechanik zweiter Ordnung.

Der folgende einfache Trick zeigt, dass man Differentialgleichungen höherer Ordnung auf solche erster Ordnung zurückführen kann.

Proposition 1.3.2 (i) *Es sei $x : J \rightarrow \mathbb{R}^n$ eine Lösung von (1.3.1). Dann ist*

$$(x, \dot{x}, \dots, x^{(m-1)})$$

Lösung der folgenden Differentialgleichung erster Ordnung

$$\begin{aligned} \dot{x}_1 &= x_2 & (1.3.2) \\ \dot{x}_2 &= x_3 \\ &\dots \\ \dot{x}_m &= F(t, x_1, x_2, \dots, x_{m-1}). \end{aligned}$$

(ii) *Ist umgekehrt $(x_1, \dots, x_m) : J \rightarrow \mathbb{R}^{mn}$ Lösung von (1.3.2), so ist $x = x_1$ Lösung von (1.3.1).*

Beweis. Dies ist offensichtlich aus den Definitionen. ■

Dies Resultat zeigt, dass man sich für alle allgemeinen Fragen (etwa nach der Existenz und Eindeutigkeit von Lösungen) auf Gleichungen erster Ordnung beschränken kann. Dies schließt allerdings nicht aus, dass Gleichungen zweiter Ordnung, wie sie in der Mechanik auftauchen, spezielle Eigenschaften haben, die sich nicht in der allgemeinen Theorie wiederfinden.

Offenbar gibt es auch allgemeinere Differentialgleichungen, bei denen man nicht nach der höchsten Ableitung auflösen kann, die also nicht die Form (1.3.1) haben. Dies führt insbesondere auf sogenannte differentiell-algebraische Gleichungen, die in den letzten Jahrzehnten auf Grund ihrer hohen Anwendungsrelevanz (zum Beispiel in der Schaltkreis-Simulation) verstärkt untersucht werden. Im Rahmen dieser Einführung werden wir ihre Theorie nur streifen. Hier bemerken wir nur, dass die Existenz von Lösungen sehr problematisch ist, wenn man nicht nach der höchsten Ableitung auflösen kann, wie das Beispiel

$$(\dot{x})^2 + 1 = 0$$

zeigt.

Kapitel 2

Existenz und Eindeutigkeit von Lösungen

In diesem Kapitel beginnen wir die systematische Theorie gewöhnlicher Differentialgleichungen, indem wir Sätze über die Existenz und Eindeutigkeit von Lösungen beweisen.

2.1 Problemformulierung

Wir betrachten das folgende **Anfangswert-Problem** für eine Differentialgleichung erster Ordnung. Es sei $f : D \rightarrow \mathbb{R}^n$ eine stetige Funktion auf einer offenen Teilmenge $D \subset \mathbb{R} \times \mathbb{R}^n$, und es sei $(t_0, x_0) \in D$ gegeben. Gesucht ist eine differenzierbare Funktion $x : J \rightarrow \mathbb{R}^n$, definiert auf einem Intervall J mit $t_0 \in J$, so dass

$$\dot{x}(t) = f(t, x(t)) \text{ für alle } t \in J \text{ und } x(t_0) = x_0. \quad (2.1.1)$$

Gesucht ist also eine Lösung, die auch die Anfangsbedingung $x(t_0) = x_0$ erfüllt. Offenbar ist jede Lösung dieser Differentialgleichung sogar stetig differenzierbar, weil die Komposition stetiger Funktionen stetig ist.

Wir erinnern daran, dass eine Menge offen ist, falls jeder ihrer Punkte eine Kugel mit positivem Radius enthält. Insbesondere ist $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, falls für jedes $(t_0, x_0) \in D$ ein $\varepsilon > 0$ existiert, so dass

$$\mathbf{B}(t_0, x_0; \varepsilon) := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n \mid |t - t_0| < \varepsilon \text{ und } \|x - x_0\| < \varepsilon\} \subset D;$$

dabei ist $\|\cdot\|$ irgendeine Norm auf \mathbb{R}^n .

Die folgenden Beispiele illustrieren die Schwierigkeiten bei der Frage nach der Existenz und Eindeutigkeit von Lösungen.

Beispiel 2.1.1

$$\dot{x} = f(t, x) \text{ mit } f(t, x) = \begin{cases} -1 & \text{für } t < 0 \\ 0 & \text{für } t = 0 \\ 1 & \text{für } t > 0 \end{cases}.$$

Die einzigen Lösungskandidaten für diese Differentialgleichung mit unstetiger rechter Seite f sind offenbar

$$x(t) = |t| + c$$

mit $c \in \mathbb{R}$. Diese Funktionen sind in $t = 0$ nicht differenzierbar, also gibt es für Anfangsbedingungen der Form $x(0) = x_0$ keine Lösungen, die auf einem Intervall J , das 0 im Innern enthält, also $0 \in \text{int} J$, definiert sind..

Beispiel 2.1.2 Für $\dot{x} = x^2$ sind, wie oben gezeigt, die Funktionen

$$x(t) = -\frac{1}{t+c}, \quad t \neq -c$$

Lösungen. Zum Beispiel erhält man für den Anfangswert $x(-2) = 1$ die Lösung

$$x(t) = -\frac{1}{t+1}, \quad t \in (-\infty, -1).$$

Diese Lösung ist also nur auf dem Teilintervall $J := (-\infty, -1)$ definiert und wird unbeschränkt für $t \nearrow -1$.

Beispiel 2.1.3 Mit der Signum-Funktion $\text{sign}x := \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -1 & \text{für } x < 0. \end{cases}$ betrachte

$$\dot{x} = \text{sign}x \cdot \sqrt{|x|}, \quad x(0) = 0. \quad (2.1.2)$$

Neben der trivialen Lösung $x(t) = 0, t \in \mathbb{R}$, gibt es unendlich viele Lösungen auf \mathbb{R} gegeben für $\alpha \geq 0$ durch

$$x_\alpha(t) = \begin{cases} \pm \frac{1}{4}(t-\alpha)^2 & \text{für } t \geq \alpha \\ 0 & \text{für } t \leq \alpha. \end{cases}$$

In den Beispielen 2.1.2 und 2.1.3 findet man also bei stetiger rechter Seite Lösungen von Anfangswertproblemen auf einem offenen Intervall, das die Anfangszeit enthält. Dabei muss die Lösung nicht eindeutig sein; hierfür braucht man also stärkere Voraussetzungen.

2.2 Lipschitz-Stetigkeit und die Gronwall-Ungleichung

Für die Eindeutigkeit der Lösungen von Anfangswertproblemen reicht, wie gesehen, Stetigkeit der rechten Seite nicht aus. Stattdessen ist der folgende Begriff relevant.

Definition 2.2.1 Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}^n$ stetig. Dann heißt f Lipschitz-stetig bezüglich x , falls für jedes $(t_0, x_0) \in D$ eine Umgebung U und eine Konstante $L > 0$ existieren mit

$$\|f(t, x_1) - f(t, x_2)\| \leq L \|x_1 - x_2\| \quad \text{für alle } (t, x_1), (t, x_2) \in U.$$

Dabei heißt $L = L(t_0, x_0)$ Lipschitz-Konstante.

Will man betonen, dass die Lipschitz-Konstante L von dem betrachteten Paar (t_0, x_0) abhängt, so spricht man auch von lokaler Lipschitz-Stetigkeit.

Beispiel 2.2.2 Die Funktion $f(x) = |x|$ auf \mathbb{R} ist Lipschitz-stetig. Die Funktion $f(x) = \sqrt{|x|}$ ist nicht Lipschitz-stetig in $x = 0$, weil für

$$\frac{|f(x) - f(0)|}{|x - 0|} = \left| x^{-1/2} \right| \rightarrow \infty \quad \text{für } x \rightarrow 0, \quad x \neq 0.$$

Wir werden zeigen, dass Lipschitz-Stetigkeit eine Abschwächung von stetiger Differenzierbarkeit ist. Zunächst erinnern wir an die Definition von stetiger Differenzierbarkeit.

Eine Abbildung $f = (f_i) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist stetig differenzierbar in $x_0 \in \mathbb{R}^n$, falls die partiellen Ableitungen $\frac{\partial f_i}{\partial x_j}(x)$ in einer Umgebung von x_0 existieren und stetig in x sind.

Bemerkung 2.2.3 Etwas abstrakter kann man das folgendermaßen formulieren: Die Ableitung einer Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ in $x_0 \in \mathbb{R}^n$ ist (falls existent) die lineare Abbildung $f'(x_0) = Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit

$$\lim_{\|x-x_0\| \rightarrow 0} \frac{\|f(x) - f(x_0) - Df(x_0)(x-x_0)\|}{\|x-x_0\|} = 0.$$

Bezüglich der kanonischen Basen in \mathbb{R}^n und \mathbb{R}^m wird $Df(x_0)$ durch die Jacobi-Matrix $\left(\frac{\partial f_i}{\partial x_j}(x_0)\right)_{i,j}$ dargestellt. Ist f für alle x in einer Umgebung U von x_0 differenzierbar, so ist die stetige Differenzierbarkeit äquivalent dazu, dass die (nichtlineare) Abbildung $x \mapsto Df(x) : U \rightarrow \mathbb{R}^{m \times n}$ stetig in x_0 ist; hierbei wird $\mathbb{R}^{m \times n}$ mit einer Matrix-Norm betrachtet.

Die folgende Proposition zeigt, dass stetig differenzierbare Abbildungen Lipschitz-stetig sind.

Proposition 2.2.4 Für eine offene Menge $D \subset \mathbb{R} \times \mathbb{R}^n$ sei $f : D \rightarrow \mathbb{R}^n$ stetig und bezüglich x differenzierbar, und $D_x f(t, x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ hänge stetig von $(t, x) \in D$ ab (also alle partiellen Ableitungen $\frac{\partial}{\partial x_j} f_i(t, x)$ sind stetig in (t, x)). Dann ist f Lipschitz-stetig bezüglich x .

Beweis. Für $(t_0, x_0) \in D$ und $\varepsilon > 0$ gibt es wegen der Stetigkeit eine Umgebung $I \times V \subset D$ mit

$$\|D_x f(t, x) - D_x f(t_0, x_0)\| \leq \varepsilon \text{ für alle } (t, x) \in I \times V.$$

Also folgt für alle $(t, x) \in J \times V$

$$\begin{aligned} \|D_x f(t, x)\| &\leq \|D_x f(t_0, x_0)\| + \|D_x f(t, x) - D_x f(t_0, x_0)\| \\ &\leq \|D_x f(t_0, x_0)\| + \varepsilon =: L. \end{aligned}$$

Es gibt $\delta > 0$, so dass $\mathbf{B}(x_0; \delta) = \{x \in \mathbb{R}^n, \|x - x_0\| < \delta\} \subset W$. Dann folgt nach dem Mittelwertsatz für alle $(t, x_1), (t, x_2) \in J \times \mathbf{B}(x_0; \delta)$

$$\|f(t, x_1) - f(t, x_2)\| \leq \sup_{0 \leq s \leq 1} \|D_x f(t, x_1 + s(x_2 - x_1))\| \|x_2 - x_1\| \leq L \|x_2 - x_1\|,$$

weil $x_1 + s(x_2 - x_1) \in \mathbf{B}(x_0, \delta) \subset W$ für alle $s \in [0, 1]$. ■

Beispiel 2.2.5 Die Funktion $f(x) = x^2$ ist Lipschitz-stetig auf \mathbb{R} , weil

$$|x_1^2 - x_2^2| \leq |x_1 + x_2| |x_1 - x_2|.$$

Es gibt jedoch keine gemeinsame Lipschitz-Konstante L für alle Punkte $x_0 \in \mathbb{R}$.

Die folgende Proposition zeigt, dass bei Kompaktheit lokale Lipschitz-Stetigkeit die globale Lipschitz-Stetigkeit impliziert. Wir erinnern daran, dass eine Teilmenge von \mathbb{R}^n kompakt ist, wenn sie abgeschlossen und beschränkt ist. Äquivalent dazu ist, dass jede Überdeckung durch offene Mengen eine endliche Teilüberdeckung enthält (Satz von Heine-Borel).

Proposition 2.2.6 Es sei K eine kompakte Teilmenge einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n$. Ist $f : D \rightarrow \mathbb{R}^n$ stetig und Lipschitz-stetig bezüglich x , so gibt es eine offene Menge W mit $K \subset W \subset D$, so dass $f : W \rightarrow \mathbb{R}^n$ global (oder gleichmäßig) Lipschitz-stetig ist, das heißt, es gibt eine Konstante $L > 0$, so dass für alle $(t, x_1), (t, x_2) \in W$ gilt

$$\|f(t, x_1) - f(t, x_2)\| \leq L \|x_2 - x_1\|.$$

Beweis. Zu jedem $(t, x) \in K$ existieren nach Voraussetzung eine Lipschitz-Konstante $L = L(t, x) > 0$ und eine offene Umgebung $U_{t,x} \subset D$ mit

$$\|f(t_1, x_1) - f(t_1, x_2)\| \leq L(t, x) \|x_1 - x_2\| \text{ für alle } (t_1, x_1), (t_1, x_2) \in U_{t,x}.$$

Wir können annehmen, dass $U_{t,x} := \mathbf{B}(t, x; \delta_{t,x})$ für $\delta_{t,x} > 0$. Offensichtlich gilt $K \subset \bigcup_{(t,x) \in K} \mathbf{B}(t, x; \delta_{t,x}/2)$, dies ist also eine offene Überdeckung von K . Weil K kompakt ist, existiert eine endliche Teilüberdeckung, das heißt, es gibt es endlich viele $(t_i, x_i) \in K, i = 1, \dots, N$, mit

$$K \subset \bigcup_{i=1}^N \mathbf{B}(t_i, x_i; \delta_i/2) =: W \text{ mit } \delta_i := \delta_{t_i, x_i}.$$

Daher folgt, dass

$$K \subset W \subset \overline{W} \subset \bigcup_{i=1}^N \mathbf{B}(t_i, x_i; \delta_i) \subset D.$$

Weil \overline{W} kompakt ist, ist die stetige Funktion f auf \overline{W} beschränkt, es gibt daher $c > 0$ mit

$$\|f(t, x)\| \leq c \text{ für alle } (t, x) \in \overline{W}.$$

Setze

$$\delta := \min\{\delta_i/2 \mid i = 1, \dots, N\} \text{ und } L := \max\left\{\frac{2c}{\delta}, L(t_i, x_i) \mid i = 1, \dots, N\right\}.$$

Zum Nachweis der Lipschitz-Eigenschaft betrachte nun $(t, x), (t, y) \in W$.

1. Fall: $\|x - y\| < \delta$: Es gibt i mit $(t, x) \in \mathbf{B}(t_i, x_i; \delta_i/2)$, und dann ist $(t, y) \in \mathbf{B}(t_i, x_i; \delta_i)$, weil $\|y - x_i\| \leq \|y - x\| + \|x - x_i\| \leq \delta + \delta_i/2 \leq \delta_i$. Daher

$$\|f(t, x) - f(t, y)\| \leq L(t_i, x_i) \|x - y\| \leq L \|x - y\|.$$

2. Fall: $\|x - y\| \geq \delta$: Dann ist

$$\|f(t, x) - f(t, y)\| \leq 2c = \frac{2c}{\delta} \delta \leq L \|x - y\|,$$

Beide Fälle zusammen zeigen, dass L eine globale Lipschitz-Konstante ist. ■

Als nächstes stellen wir ein grundlegendes Hilfsmittel für die Analyse von Differentialgleichungen bereit, das Gronwall-Lemma (dieses Lemma existiert in vielen Varianten, wir benötigen nur die einfachste Form).

Lemma 2.2.7 (Gronwall) *Es sei $I = [t_0, t_1] \subset \mathbb{R}$ ein Intervall und $u : I \rightarrow [0, \infty)$ sei eine stetige Funktion, die für Konstanten $\alpha \geq 0$ und $\beta > 0$*

$$u(t) \leq \alpha + \beta \int_{t_0}^t u(s) ds \text{ für alle } t \in I \tag{2.2.1}$$

erfüllt. Dann gilt für alle $t \in I$

$$u(t) \leq \alpha e^{\beta(t-t_0)}. \tag{2.2.2}$$

Beweis. Wir definieren auf I die Hilfsfunktion

$$v(t) := \int_{t_0}^t u(s) ds.$$

Nach dem Hauptsatz der Differential- und Integralrechnung erhält man aus (2.2.1)

$$\dot{v}(t) = u(t) \leq \alpha + \beta v(t).$$

Definiere nun die weitere Hilfsfunktion

$$w(t) := v(t)e^{-\beta t}, t \in I.$$

Durch Differenzieren ergibt dies mit der Produktregel

$$\dot{w}(t) = \left(\frac{d}{dt} v(t) \right) e^{-\beta t} + v(t) \left(\frac{d}{dt} e^{-\beta t} \right) \leq (\alpha + \beta v(t)) e^{-\beta t} - \beta v(t) e^{-\beta t} = \alpha e^{-\beta t}.$$

Integration von t_0 bis t liefert wegen $w(t_0) = 0$

$$w(t) = w(t) - w(t_0) \leq \frac{\alpha}{\beta} (e^{-\beta t_0} - e^{-\beta t}).$$

Damit ergibt sich

$$v(t) \leq \frac{\alpha}{\beta} e^{\beta(t-t_0)} - \frac{\alpha}{\beta}.$$

Für u erhält man daraus und aus (2.2.1)

$$u(t) \leq \alpha + \beta v(t) \leq \alpha + \alpha e^{\beta(t-t_0)} - \alpha = \alpha e^{\beta(t-t_0)}.$$

■

2.3 Eindeutigkeit von Lösungen

In diesem Abschnitt beweisen wir, dass für Lipschitz-stetige rechte Seiten das Anfangswertproblem (2.1.1) höchstens eine Lösung auf einem Intervall um t_0 besitzt. Ferner beweisen wir Abschätzungen für approximative Lösungen. Dies werden wir für Existenzaussagen benötigen.

Wir beginnen mit einer Aussage über Eindeutigkeit von Lösungen.

Satz 2.3.1 (Eindeutigkeitssatz) *Für eine offene Menge $D \subset \mathbb{R} \times \mathbb{R}^n$ sei $f : D \rightarrow \mathbb{R}^n$ stetig und Lipschitz-stetig bezüglich x . Es seien $x_1 : I_1 \rightarrow \mathbb{R}^n$ und $x_2 : I_2 \rightarrow \mathbb{R}^n$ Lösungen von $\dot{x} = f(t, x)$ mit $x_1(t_0) = x_2(t_0)$ für ein $t_0 \in I_1 \cap I_2$. Dann gilt $x_1(t) = x_2(t)$ für alle $t \in I_1 \cap I_2$.*

Beweis. Im Allgemeinen ist $I_1 \cap I_2$ nicht kompakt, wir können aber schreiben

$$I_1 \cap I_2 = \bigcup_{k=0}^{\infty} J_k,$$

mit kompakten Intervallen $J_k \subset J_{k+1}$ für alle k . Daher reicht es, Folgendes zu zeigen: Es sei I ein kompaktes Intervall mit $t_0 \in I \subset I_1 \cap I_2$; dann folgt aus $x_1(t_0) = x_2(t_0)$, dass $x_1(t) = x_2(t)$ für alle $t \in I$.

Da x_1 und x_2 stetig sind, sind die Bildmengen $x_1(I)$ und $x_2(I)$ kompakt in \mathbb{R}^n , also ist auch

$$K := I \times (x_1(I) \cup x_2(I))$$

kompakt. Nach Proposition 2.2.6 ist die Einschränkung $f : K \rightarrow \mathbb{R}^n$ global Lipschitz-stetig mit einer Konstanten L . Es folgt für alle $t \in I$ mit $t \geq t_0$

$$\begin{aligned} \|x_1(t) - x_2(t)\| &= \left\| x_1(t_0) + \int_{t_0}^t f(s, x_1(s)) ds - x_2(t_0) - \int_{t_0}^t f(s, x_2(s)) ds \right\| \\ &\leq \int_{t_0}^t \|f(s, x_1(s)) - f(s, x_2(s))\| ds \\ &\leq L \int_{t_0}^t \|x_1(s) - x_2(s)\| ds. \end{aligned}$$

Aus Gronwalls Ungleichung, Lemma 2.2.7, folgt daher

$$\|x_1(t) - x_2(t)\| = 0 \text{ für alle } t \in I \text{ mit } t \geq t_0.$$

Für $t \leq t_0$ benutzen wir den folgenden Trick (Zeitumkehr), der auch allgemein von Interesse ist: Die Funktion

$$y_1(t) := x_1(-t), \quad t \in -I,$$

ist Lösung des Anfangswertproblems

$$\dot{y} = -f(-t, y), \quad y(-t_0) = x_1(t_0),$$

denn

$$\frac{d}{dt}y_1(t) = \frac{d}{dt}x_1(-t) = -f(-t, x_1(-t)) = -f(-t, y_1(t)), \quad t \in -I.$$

Entsprechend wird $y_2(t), t \in -I$ definiert. Die rechte Seite erfüllt die gleiche Lipschitz-Bedingung wie f , und $t \leq t_0$ ist äquivalent zu $-t \geq -t_0$. Damit folgt die Behauptung aus dem ersten Teil des Beweises (angewendet auf y_1 und y_2). ■

Bemerkung 2.3.2 Die rechte Seite in unserem Gegenbeispiel (2.1.2) für eindeutige Lösbarkeit ist nicht Lipschitz-stetig in $x = 0$. Lipschitz-Stetigkeit der rechten Seite ist aber nicht notwendig für eindeutige Lösbarkeit: Ein Beispiel findet sich in Hairer, Norsett und Wanner [10, Exercise 5 in Section I.8].

Bemerkung 2.3.3 Als Spezialfall ergibt sich aus Theorem 2.3.1 und Proposition 2.2.4, dass Anfangswertprobleme mit einer stetigen rechten Seite f , die stetig nach x differenzierbar ist, höchstens eine Lösung besitzen.

Unsere Strategie, um die Existenz von Lösungen nachzuweisen, beruht darauf, zunächst Näherungslösungen zu konstruieren und dann daraus geeignete konvergente Teilfolgen zu finden, die gegen eine Lösung konvergieren. Dafür zeigen wir jetzt, mit den gleichen Argumenten wie oben, eine Abschätzung für approximative Lösungen, die wir zunächst definieren.

Definition 2.3.4 Sei $\varepsilon > 0$ und $f : D \rightarrow \mathbb{R}^n$ stetig auf einer offenen Teilmenge $D \subset \mathbb{R} \times \mathbb{R}^n$. Es sei $I \subset \mathbb{R}$ ein beschränktes Intervall und $M \subset \mathbb{R}^n$ mit $I \times M \subset D$. Eine stetige Funktion $x : I \rightarrow M$ heißt ε -Näherungslösung von $\dot{x} = f(t, x)$, falls

- (i) $I = \bigcup_{k=1}^N I_k$ mit Intervallen I_k und die Einschränkung $x|_{\overline{I_k}}$ stetig differenzierbar ist;
(ii) für alle $k = 1, \dots, N$ und $t \in I_k$ gilt

$$\|\dot{x}(t) - f(t, x(t))\| \leq \varepsilon.$$

Hier ist also insbesondere stückweise stetige Differenzierbarkeit von x verlangt.

Proposition 2.3.5 Es sei $x : I \rightarrow M$ eine ε -Näherungslösung von $\dot{x} = f(t, x)$. Dann gilt für $t_0, t \in I$

$$\left\| x(t) - x(t_0) - \int_{t_0}^t f(s, x(s)) ds \right\| \leq \varepsilon |t - t_0|.$$

Beweis. Für $t > t_0$ kann man $[t_0, t]$ in Teilintervalle $[s_k, s_{k+1}]$ zerlegen,

$$[t_0, t] = \bigcup_{k=1}^N [s_k, s_{k+1}],$$

so dass x auf $[s_k, s_{k+1}]$ stetig differenzierbar ist. Dann ist

$$x(s_{k+1}) - x(s_k) = \int_{s_k}^{s_{k+1}} \dot{x}(s) ds,$$

also

$$\begin{aligned}
\left\| x(t) - x(t_0) - \int_{t_0}^t f(s, x(s)) ds \right\| &= \left\| \sum_{k=0}^{N-1} (x(s_{k+1}) - x(s_k)) - \int_{t_0}^t f(s, x(s)) ds \right\| \\
&= \left\| \sum_{k=0}^{N-1} \int_{s_k}^{s_{k+1}} \dot{x}(s) ds - \int_{t_0}^t f(s, x(s)) ds \right\| \\
&\leq \int_{t_0}^t \|\dot{x}(s) - f(s, x(s))\| ds \\
&\leq \varepsilon (t - t_0).
\end{aligned}$$

Analog argumentiert man für $t < t_0$. ■

2.4 Existenz von Lösungen

In diesem Abschnitt beweisen wir die Sätze von Peano und Picard-Lindelöf über die Existenz von Lösungen.

Zunächst konstruieren wir approximative Lösungen von (2.1.1) als Euler-Polygone.

Vorüberlegung: Eine erste Einschränkung für das Existenzintervall einer Lösung durch $(t_0, x_0) \in D \subset \mathbb{R} \times \mathbb{R}^n$ ergibt sich schon aus der Forderung, dass der Graph der Lösung in D bleiben muss. Es gibt $a, b > 0$ mit

$$R := [t_0 - a, t_0 + a] \times \overline{\mathbf{B}(x_0; b)} \subset D. \quad (2.4.1)$$

Weil R kompakt ist, nimmt f auf R sein Maximum an,

$$c := \max_{(t,x) \in R} \|f(t, x)\| < \infty.$$

Ist x eine Lösung in R von $\dot{x} = f(t, x)$ durch (t_0, x_0) , so folgt

$$\|x(t) - x(t_0)\| = \left\| \int_{t_0}^t f(s, x(s)) ds \right\| \leq c |t - t_0|.$$

Gilt daher außerdem $|t - t_0| \leq \frac{b}{c}$, so ist gesichert, dass der Graph der Lösung in R bleibt, weil dann

$$\|x(t) - x(t_0)\| \leq c \frac{b}{c} = b.$$

Lemma 2.4.1 *Mit den obigen Bezeichnungen sei $\alpha := \min(a, \frac{b}{c})$. Dann existiert für jedes $\varepsilon > 0$ eine ε -Näherungslösung $x : [t_0 - \alpha, t_0 + \alpha] \rightarrow \overline{\mathbf{B}(x_0; b)}$ mit $x(t_0) = x_0$ und*

$$\|x(t) - x(s)\| \leq c |t - s| \text{ für alle } t, s. \quad (2.4.2)$$

Beweis. Weil f auf der kompakten Menge R stetig ist, ist f gleichmäßig stetig, es existiert also für $\varepsilon > 0$ ein $\delta > 0$, so dass für alle $(t, x), (s, y) \in R$

$$|t - s| < \delta \text{ und } \|x - y\| < \delta \text{ impliziert } \|f(t, x) - f(s, y)\| < \varepsilon. \quad (2.4.3)$$

Im Folgenden schreiben wir die Argumente nur für $[t_0, t_0 + \alpha]$ auf; für $[t_0 - \alpha, t_0]$ argumentiert man analog. Zerlege $[t_0, t_0 + \alpha]$ in Teilintervalle

$$t_0 < t_1 < \dots < t_m = t_0 + \alpha,$$

so dass für die Feinheit der Zerlegung gilt

$$\max(t_i - t_{i-1}) < \min\left(\delta, \frac{\delta}{c}\right).$$

Ausgehend von $t = t_0$ definieren wir einen Polygonzug (eine stetige, stückweise lineare Funktion) durch

$$x(t) = x(t_i) + (t - t_i)f(t_i, x(t_i)) \text{ für } t \in (t_i, t_{i+1}] \text{ und alle } i.$$

Offenbar ist x stückweise stetig differenzierbar mit

$$\dot{x}(t) = f(t_i, x(t_i)) \text{ für } t \in [t_i, t_{i+1}).$$

Wir zeigen zunächst induktiv, dass $x(t_i) \in \overline{\mathbf{B}(x_0; b)}$ und $\|f(t_i, x(t_i))\| \leq c$ für alle i . In der Tat ist

$$\|x(t_i) - x_0\| = \left\| \sum_{j=0}^{i-1} [x(t_{j+1}) - x(t_j)] \right\| = \left\| \sum_{j=0}^{i-1} [t_{j+1} - t_j] f(t_j, x_j) \right\| \leq c \sum_{j=0}^{i-1} [t_{j+1} - t_j] \leq c \alpha,$$

also auch $(t_i, x(t_i)) \in R$ und damit $\|f(t_i, x(t_i))\| \leq c$. Ferner ist x eine ε -Näherungslösung: für $t \in [t_i, t_{i+1}]$ gilt

$$\|\dot{x}(t) - f(t, x(t))\| = \|f(t_i, x(t_i)) - f(t, x(t))\| < \varepsilon$$

nach (2.4.3), weil nach Definition der Zerlegung hier

$$|t_i - t| < \min\left(\delta, \frac{\delta}{c}\right) \text{ und } \|x(t_i) - x(t)\| = |t - t_i| \|f(t_i, x(t_i))\| < \frac{\delta}{c} \cdot c = \delta.$$

Es bleibt die Abschätzung (2.4.2) zu zeigen. Liegen t und s im selben Intervall $[t_i, t_{i+1}]$ so gilt

$$\|x(t) - x(s)\| = \|(t - t_i)f(t_i, x_i) - (s - t_i)f(t_i, x_i)\| \leq c|t - s|.$$

Andernfalls können wir annehmen, dass $t \in [t_i, t_{i+1}]$ und $s \in [t_j, t_{j+1}]$ mit $0 \leq i < j$. Dann folgt unter Verwendung des obigen Spezialfalls

$$\begin{aligned} \|x(t) - x(s)\| &\leq \|x(t) - x(t_{i+1})\| + \sum_{k=i+1}^{j-1} \|x(t_k) - x(t_{k+1})\| + \|x(t_j) - x(s)\| \\ &\leq c \left[(t_{i+1} - t) + \sum_{k=i+1}^{j-1} (t_{k+1} - t_k) + (s - t_j) \right] \\ &= c|t - s|. \end{aligned}$$

■

Zur Konstruktion einer Lösung betrachten wir für eine Folge $\varepsilon_k \rightarrow 0$ die in Lemma 2.4.1 konstruierten ε_k -Näherungslösungen und wollen zeigen, dass sie gegen eine Lösung konvergieren. Wir benötigen dafür ein Kriterium, das uns garantiert, dass eine Folge stetiger Funktionen gegen eine stetige Grenzfunktion konvergiert. Die beschränkte Folge stetiger Funktionen

$$g_k(t) := t^k, \quad t \in [0, 1],$$

konvergiert zwar für jedes $t \in [0, 1]$, die Grenzfunktion $g(t) := 0$ für $t \in [0, 1)$ und $g(1) := 1$, ist jedoch nicht stetig. Daher benötigt man stärkere Voraussetzungen.

Dazu zitieren wir den folgenden Satz aus der Funktionalanalysis. Er besagt, dass jede Folge von gleichmäßig beschränkten und gleichgradig stetigen Funktionen von einem kompakten Intervall nach \mathbb{R}^n eine Teilfolge besitzt, die gleichmäßig gegen eine stetige Funktion konvergiert. (Dies ist ein Kriterium für Kompaktheit im Raum $C([a, b], \mathbb{R}^n)$ der stetigen Funktionen von $[a, b]$ nach \mathbb{R}^n versehen mit der Maximums-Norm.)

Satz 2.4.2 (Arzelá-Ascoli) Für $a < b$ in \mathbb{R} und $k \in \mathbb{N}$ sei $g_k : [a, b] \rightarrow \mathbb{R}^n$ mit

$$\sup_{k \in \mathbb{N}} \|g_k\|_\infty = \sup_{k \in \mathbb{N}} \max_{t \in [a, b]} \|g_k(t)\| < \infty,$$

und für alle $\varepsilon > 0$ gebe es $\delta > 0$, so dass für alle $k \in \mathbb{N}$ und $t, s \in [a, b]$

$$|t - s| < \delta \text{ impliziert } \|g_k(t) - g_k(s)\| < \varepsilon.$$

Dann existieren eine stetige Funktion $g : [a, b] \rightarrow \mathbb{R}^n$ und eine Teilfolge (g_{k_j}) mit

$$\max_{t \in [a, b]} \|g_{k_j}(t) - g(t)\| \rightarrow 0 \text{ für } j \rightarrow \infty.$$

Einen ausführlichen **Beweis** findet man zum Beispiel in Aulbach [4, Anhang B]. Er beruht darauf, dass man zunächst für jeden der abzählbar vielen Punkte in $[a, b] \cap Q$ die Konvergenz zeigt, dann eine Diagonalfolge wählt, und schließlich die gleichgradige Stetigkeit ausnutzt, um die Behauptung für alle $t \in [a, b]$ zu zeigen.

Dies liefert den folgenden grundlegenden Existenzsatz.

Satz 2.4.3 (Peano) Gegeben sei ein Anfangswertproblem der Form

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

mit $f : D \rightarrow \mathbb{R}^n$ stetig und $(t_0, x_0) \in D \subset \mathbb{R} \times \mathbb{R}^n$, offen. Dann gibt es eine (lokale) Lösung durch (t_0, x_0) , die für $t \in [t_0 - \alpha, t_0 + \alpha]$, $\alpha > 0$, existiert und in $\mathbf{B}(x_0; b)$ liegt. Hierbei sind $\alpha, b > 0$ folgendermaßen gegeben: Wähle $a, b > 0$, so dass

$$R := [t_0 - a, t_0 + a] \times \overline{\mathbf{B}(x_0; b)} \subset D \text{ und } \alpha := \min\left(a, \frac{b}{c}\right) \text{ mit } c := \max_{(t, x) \in R} \|f(t, x)\|. \quad (2.4.4)$$

Beweis. Vorbemerkung: Es reicht zu zeigen, dass eine Funktion $x(\cdot)$ existiert mit

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds \text{ für alle } t.$$

Wir konstruieren eine Lösung als Häufungspunkt einer Folge von approximativen Lösungen. Definiere den Doppelkegel

$$S = \{(t, x) \in R \mid \|x - x_0\| \leq c \|t - t_0\| \text{ und } |t - t_0| \leq \alpha\}.$$

Sei $\varepsilon_m = \frac{1}{m}$, $m \in \mathbb{N}$. Nach Lemma 2.4.1 existieren ε_m -Näherungslösungen $x_m : [t_0 - \alpha, t_0 + \alpha] \rightarrow \mathbb{R}^n$, $m \in \mathbb{N}$, deren Graph in S verläuft, sie sind also gleichmäßig beschränkt, und sie sind gleichgradig stetig. Nach Lemma 2.3.5 gilt für alle m

$$\left\| x_m(t) - x_m(t_0) - \int_{t_0}^t f(s, x_m(s)) ds \right\| \leq \frac{1}{m} |t - t_0|,$$

Wieder nach Lemma 2.4.1 sind die Funktionen $x_m : [t_0 - \alpha, t_0 + \alpha] \rightarrow \mathbb{R}^n$, $m \in \mathbb{N}$, gleichgradig stetig. Daher existiert nach dem Satz von Arzela-Ascoli eine Teilfolge (x_{m_k}) , die gleichmäßig gegen eine stetige Funktion $x : [t_0 - \alpha, t_0 + \alpha] \rightarrow \mathbb{R}^n$ konvergiert,

$$\max_t \|x_{m_k}(t) - x(t)\| \rightarrow 0 \text{ für } k \rightarrow \infty.$$

Dies impliziert die gleichmäßige Konvergenz

$$f(t, x_{m_k}(t)) \rightarrow f(t, x(t)), \quad t \in [t_0 - \alpha, t_0 + \alpha].$$

In der Tat: Sei $\varepsilon > 0$. Dann existiert wegen der Stetigkeit von f ein $\delta > 0$, so dass für alle $(t, x_1), (t, x_2) \in R$

$$\|x_1 - x_2\| < \delta \text{ impliziert } \|f(t, x_1) - f(t, x_2)\| < \varepsilon.$$

Es gibt $k_0 \in \mathbb{N}$, so dass für alle $k \geq k_0$ und alle $t \in [t_0 - \alpha, t_0 + \alpha]$

$$\|x_{m_k}(t) - x(t)\| < \delta \text{ und damit } \|f(t, x_{m_k}(t)) - f(t, x(t))\| < \varepsilon.$$

Natürlich ist für alle s auch $(s, x(s)) \in R$ und damit (Vertauschen der Grenzprozesse erlaubt!)

$$\int_{t_0}^t f(s, x_{m_k}(s)) ds \rightarrow \int_{t_0}^t f(s, x(s)) ds.$$

Wir erhalten für jedes t

$$\begin{array}{ccc} x_{m_k}(t) & - & x_0 & - & \int_{t_0}^t f(s, x_{m_k}(s)) ds & \rightarrow & 0 \\ \downarrow & & & & \downarrow & & \\ x(t) & - & x_0 & - & \int_{t_0}^t f(s, x(s)) ds & = & 0. \end{array}$$

Also ist die Grenzfunktion $x(\cdot)$ Lösung von $\dot{x} = f(t, x)$, $x(t_0) = x_0$. ■

Eine offensichtliche Folgerung aus dem Satz von Peano und der Eindeutigkeitsaussage in Theorem 2.3.1 ist, dass, falls f sogar Lipschitz-stetig bezüglich x ist, eindeutige (lokale) Lösungen existieren. Gleichmäßige Lipschitz-Stetigkeit erlaubt aber sogar einen konstruktiven Beweis dieser Aussage.

Satz 2.4.4 (Picard-Lindelöf) *Gegeben sei ein Anfangswertproblem der Form*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

wobei $f : [t_0 - a, t_0 + a] \times \overline{\mathbf{B}(x_0; b)} \rightarrow \mathbb{R}^n$, $a, b > 0$, eine stetige Funktion ist und für alle t, x, y die Lipschitz-Abschätzung

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{mit einem } L > 0,$$

erfüllt. Dann existiert genau eine Lösung des Anfangswertproblems auf dem Intervall $[t_0 - \alpha, t_0 + \alpha]$, wobei α gegeben ist durch

$$\alpha := \min \left(a, \frac{b}{c} \right) \quad \text{mit } c := \max_{(t, x) \in R} \|f(t, x)\|.$$

Beweis. Der Beweis beruht darauf, explizit eine Folge von Iterierten zu konstruieren, die dann gegen die gesuchte Lösung konvergiert. Diese sogenannten Picard-Iterierten, definiert auf $[t_0 - \alpha, t_0 + \alpha]$, sind

$$x^0(t) := x_0, \quad x^{k+1}(t) := x_0 + \int_{t_0}^t f(s, x^k(s)) ds, \quad k \geq 0, \quad (2.4.5)$$

Dann zeigt man mit Hilfe der Lipschitz-Abschätzung, dass diese Folge gleichmäßig gegen eine Funktion konvergiert, die die integrierte Version des Anfangswertproblems erfüllt.

Die Graphen der Funktionen x^k liegen in R . Dies ist klar für $k = 0$, und mit Induktion von k auf $k + 1$ gilt für $k \geq 0$

$$\|x^{k+1}(t) - x^0(t)\| = \left\| \int_{t_0}^t f(s, x^k(s)) ds \right\| \leq \left| \int_{t_0}^t \|f(s, x^k(s))\| ds \right| \leq |t - t_0| c \leq \alpha c \leq b.$$

Ebenfalls durch Induktion zeigen wir für alle t

$$\|x^{k+1}(t) - x^k(t)\| \leq c L^k \frac{|t - t_0|^{k+1}}{(k+1)!}. \quad (2.4.6)$$

Dies ist klar für $k = 0$; aus der Induktionsannahme für $k \geq 0$ folgt

$$\begin{aligned} \|x^{k+2}(t) - x^{k+1}(t)\| &\leq \left| \int_{t_0}^t \|f(s, x^{k+1}(s)) - f(s, x^k(s))\| ds \right| \leq \left| \int_{t_0}^t L \|x^{k+1}(s) - x^k(s)\| ds \right| \\ &\stackrel{I.A.}{\leq} \left| \int_{t_0}^t L c L^k \frac{|s - t_0|^{k+1}}{(k+1)!} ds \right| \\ &= c L^{k+1} \frac{|t - t_0|^{k+2}}{(k+2)!}. \end{aligned}$$

Die Funktionenfolge $(x^k(\cdot))_{k \in \mathbb{N}}$ konvergiert gleichmäßig auf $[t_0 - \alpha, t_0 + \alpha]$. In der Tat, wegen

$$x^k(t) - x_0 = x^k(t) - x^0(t) = \sum_{i=0}^{k-1} [x^{i+1}(t) - x^i(t)] \text{ für alle } t,$$

ist $(x^k(\cdot) - x_0)_{k \in \mathbb{N}}$ die Folge der Partialsummen von $\sum_{i=0}^{\infty} [x^{i+1}(t) - x^i(t)]$. Diese Reihe ist gleichmäßig konvergent, denn sie besitzt die konvergente Majorante

$$\sum_{i=0}^{\infty} cL^i \frac{\alpha^{i+1}}{(i+1)!} = \frac{c}{L} \sum_{i=0}^{\infty} \frac{L^{i+1} \alpha^{i+1}}{(i+1)!} = \frac{c}{L} (e^{\alpha L} - 1).$$

Die Grenzfunktion $x^\infty(\cdot)$ erfüllt die Gleichung

$$x^\infty(t) = x_0 + \int_{t_0}^t f(s, x^\infty(s)) ds, \quad t \in [t_0 - \alpha, t_0 + \alpha],$$

weil die Integranden in (2.4.5) gleichmäßig konvergieren und daher der Grenzübergang unter dem Integralzeichen erlaubt ist. Also ist x^∞ eine gesuchte Lösung. Die Eindeutigkeit folgt aus Theorem 2.3.1. ■

Bemerkung 2.4.5 Zusammen mit Proposition 2.2.6 zeigt der Satz von Picard-Lindelöf, dass für eine stetige und bezüglich x Lipschitz-stetige rechte Seite und Anfangsbedingung $x(t_0) = x_0$ eine eindeutige Lösung auf einem Intervall existiert, das t_0 im Innern enthält.

Wir wissen jetzt, dass bei Lipschitz-Stetigkeit lokale Lösungen von Anfangswertproblemen existieren (d.h. sie existieren in einem Intervall um die Anfangszeit) und dass sie eindeutig sind. Jetzt wenden wir uns der Frage zu, ob man die lokalen Lösungen zusammensetzen kann. Zunächst betrachten wir das folgende Beispiel.

Beispiel 2.4.6 Betrachte die Anfangswertprobleme

$$\dot{x} = 1 + x^2, \quad x(t_0) = x_0 \text{ mit } (t_0, x_0) \in \mathbb{R} \times \mathbb{R}.$$

Die rechte Seite f ist auf $D = \mathbb{R} \times \mathbb{R}$ definiert und Lipschitz-stetig. Um Lösungen zu konstruieren, ist es günstig, die Differentialgleichung in der Form

$$\frac{dx}{dt} = 1 + x^2$$

zu schreiben. Formales Rechnen (Trennung der Variablen und Integration) liefert

$$\int_{x_0}^x \frac{dx}{1+x^2} = \int_{t_0}^t dt.$$

Eine Stammfunktion von $\frac{1}{1+x^2}$ ist $\arctan x$, also

$$\arctan x - \arctan x_0 = t - t_0,$$

und Auflösen nach x ergibt

$$x = \tan(t - t_0 + \arctan x_0).$$

Nachrechnen liefert, dass

$$x(t) = \tan(t - t_0 + \arctan x_0), \quad t \in \left(t_0 - \arctan x_0 - \frac{\pi}{2}, t_0 - \arctan x_0 + \frac{\pi}{2} \right),$$

die Lösung des obigen Anfangswertproblems ist. Diese Lösung kann nicht über dies Intervall hinaus fortgesetzt werden kann. Es ist nützlich, sich dies anhand des Graphen des Tangens zu veranschaulichen (man beachte dabei, dass in der obigen Lösung der Graph um $-t_0 + \arctan x_0$ verschoben wird, um die Anfangsbedingung $x(t_0) = x_0$ zu erfüllen). Man sieht dabei, dass die Lösung zum Rand des Existenzintervalls hin unbeschränkt wird.

Jetzt formulieren und beweisen wir den folgenden globalen Existenz- und Eindeutigkeitsatz.

Satz 2.4.7 *Es sei $f : D \rightarrow \mathbb{R}^n$ mit einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n$ eine stetige und bezüglich x Lipschitz-stetige Funktion. Zu jedem $(t_0, x_0) \in D$ gibt es genau ein offenes Intervall $I_{\max} = (t^-, t^+) \subset \mathbb{R}$ mit $-\infty \leq t^- < t^+ \leq \infty$ und den folgenden Eigenschaften:*

(i) *Das Anfangswertproblem $\dot{x} = f(t, x)$, $x(t_0) = x_0$ besitzt genau eine Lösung $x_{\max}(\cdot)$ mit dem Existenzintervall (t^-, t^+) .*

(ii) *Ist $y : J \rightarrow \mathbb{R}^n$ eine weitere Lösung, so gilt $J \subset I_{\max}$ und $y(t) = x_{\max}(t)$ für alle $t \in J$. Das Intervall $I_{\max} = (t^-, t^+)$ heißt das maximale Existenzintervall für die Lösung des obigen Anfangswertproblems; wir schreiben auch $I_{\max}(t_0, x_0)$.*

Beweis. Nach dem lokalen Existenzsatz (Peano oder Picard-Lindelöf) existiert eine Lösung auf einem Intervall I_0 um t_0 . Wegen der lokalen Lipschitz-Stetigkeit bezüglich x liefert der Eindeutigkeitsatz, Satz 2.3.1, für jede Lösung $y : J_0 \rightarrow \mathbb{R}^n$, dass $x(t) = y(t)$ für $t \in I_0 \cap J_0$.

Sei I_{\max} die Vereinigung aller Intervalle, auf denen das obige Anfangswertproblem eine Lösung besitzt. Dies ist wieder ein Intervall und das Anfangswertproblem besitzt auch eine Lösung auf I_{\max} , die folgendermaßen definiert ist: Für $t \in I_{\max}$ gibt es ein Intervall J um t_0 mit $t \in J$, auf dem eine Lösung y definiert ist. Setze

$$x_{\max}(t) := y(t).$$

Diese Definition ist nach dem Eindeutigkeitsatz, Satz 2.3.1, unabhängig von der Wahl der Lösung y . Dies zeigt auch sofort die Behauptung (ii).

Es bleibt zu zeigen, dass das Existenzintervall offen ist. Ist $t^+ = \infty$ (oder $t^- = -\infty$), so ist hier nichts zu zeigen. Ist $t^\pm \in I_{\max}$, so kann man den lokalen Existenzsatz auf $(t^\pm, x(t^\pm))$ anwenden. Dies liefert die Existenz einer Lösung auf einem Intervall, das t^\pm im Innern enthält und daher eine Fortsetzung der Lösung x_{\max} zu einer Lösung des Anfangswertproblems auf einem Intervall, das t^\pm im Innern enthält, im Widerspruch zur Definition von t^\pm . ■

Wir wenden uns jetzt der Frage zu, wie sich die Lösung verhält, wenn sie sich dem Rand ihres Existenzintervalls nähert. Zunächst einige Beispiele.

Beispiel 2.4.8 *Betrachte*

$$\dot{x} = -\frac{x}{t-1}, \quad x(0) = 1.$$

Die rechte Seite $f(t, x) := -\frac{x}{t-1}$ ist definiert auf $\{(t, x) \in \mathbb{R} \times \mathbb{R} \mid t \neq 1\}$. Die Lösung ist

$$x(t) = -\frac{1}{t-1}$$

mit maximalem Existenzintervall $I_{\max} = (t^-, t^+) = (-\infty, 1)$. Die Lösung ist unbeschränkt für $t \rightarrow 1, t \in I_{\max}$.

Beispiel 2.4.9 *Betrachte*

$$\dot{x} = -\frac{1}{x}, \quad x(0) = 1.$$

Die rechte Seite $f(t, x) := -\frac{1}{x}$ ist definiert auf $\{(t, x) \in \mathbb{R} \times \mathbb{R} \mid x \neq 0\}$. Die Lösung ist

$$x(t) = \sqrt{-2t+1}$$

mit maximalem Existenzintervall $I_{\max} = (t^-, t^+) = (-\infty, \frac{1}{2})$. Die Lösung divergiert für $t \rightarrow t^+, t \in I_{\max}$ nicht, sondern konvergiert gegen den Punkt $(t^+, 0) \in \partial D$.

Diese Beispiele zeigen, dass sich der Graph $(t, x(t))$ für $t \nearrow t^+$ (d.h., $t \rightarrow t^+$ und $t < t^+$) dem Rand des Definitionsbereichs D von f nähern kann; dabei kann die Lösung unbeschränkt werden oder konvergieren. Der folgende Satz zeigt insbesondere, dass die Lösung unbeschränkt werden muss, wenn sie sich nicht dem Rand des Definitionsbereichs nähert.

Bezeichnung: Für $x \in \mathbb{R}^n$ und $A \subset \mathbb{R}^n$ ist

$$\text{dist}(x, A) := \inf\{\|x - a\|, a \in A\}, \text{ falls } A \neq \emptyset \text{ und } \text{dist}(x, \emptyset) := \infty.$$

Aus der Dreiecksungleichung folgt dann für $x, y \in \mathbb{R}^n$

$$\text{dist}(x, A) \leq \|x - y\| + \text{dist}(y, A).$$

Ferner ist $\partial A = \overline{A} \cap (\overline{\mathbb{R}^n} \setminus A)$ der Rand von A .

Satz 2.4.10 *Es sei $f : J \times D_0 \rightarrow \mathbb{R}^n$ mit einem offenen Intervall $J \subset \mathbb{R}$ und einer offenen Menge $D_0 \subset \mathbb{R}^n$ eine stetige und bezüglich x Lipschitz-stetige Funktion. Ferner sei $x : I_{\max} = (t^-, t^+) \rightarrow \mathbb{R}^n$ die maximale Lösung des Anfangswertproblems mit $t_0 \in J, x_0 \in D_0$,*

$$\dot{x} = f(t, x), \quad x(t_0) = x_0.$$

Ist $\{\|x(t)\|, t \in [t_0, t^+)\}$ beschränkt, so folgt $t^+ = \sup J$ oder der Rand ∂D_0 von D_0 ist nichtleer und

$$\lim_{t \nearrow t^+} \text{dist}(x(t), \partial D_0) = 0.$$

Analog für $t^- > \inf J$.

Beweis. Es sei $t^+ < \sup J$ (im Fall $t^- > \inf J$ argumentiert man analog). Ist die Behauptung falsch, so gibt es $\varepsilon > 0$ und eine Folge $t_i \nearrow t^+$ (d.h. $t_i \rightarrow t^+, t_i < t^+$) mit

$$\text{dist}(x(t_i), \partial D_0) \geq 2\varepsilon; \tag{2.4.7}$$

Es seien

$$a := \sup\{\|x(t)\|, t \in [t_0, t^+)\}, c := \max\{\|f(t, x)\| \mid t_0 \leq t \leq t^+ \text{ und } \|x\| \leq a, \text{dist}(x, \partial D_0) \geq \varepsilon\}$$

und es sei $0 < \delta < \frac{\varepsilon}{c}$.

Behauptung: Dann gilt für alle $i \in \mathbb{N}$ und $0 \leq s \leq \min\{\delta, t^+ - t_i\}$

$$\text{dist}(x(t_i + s), \partial D_0) > \varepsilon. \tag{2.4.8}$$

In der Tat: Gilt diese Behauptung nicht, so gibt es $k \in \mathbb{N}$ und β mit $0 < \beta \leq \min\{\delta, t^+ - t_k\}$ und

$$\text{dist}(x(t_k + s), \partial D_0) \geq \varepsilon \text{ für } 0 \leq s \leq \beta$$

und

$$\text{dist}(x(t_k + \beta), \partial D_0) = \varepsilon. \tag{2.4.9}$$

Daraus folgt

$$\|f(t_k + s, x(t_k + s))\| \leq c \text{ für } 0 \leq s \leq \beta$$

und damit

$$\|x(t_k + \beta) - x(t_k)\| \leq \int_{t_k}^{t_k + \beta} \|f(s, x(s))\| ds \leq \beta c \leq \delta c < \varepsilon.$$

Folglich

$$\text{dist}(x(t_k + \beta), \partial D_0) \geq \text{dist}(x(t_k), \partial D_0) - (\|x(t_k + \beta)\| - \|x(t_k)\|) > 2\varepsilon - \varepsilon = \varepsilon,$$

was der Definition von β widerspricht. Daher gilt die **Behauptung**.

Nach (2.4.8) gilt für all $i \in \mathbb{N}$ mit $t^+ - t_i \leq \delta$

$$\|x(t) - x(s)\| \leq \left| \int_s^t \|f(\tau, x(\tau))\| d\tau \right| \leq c|t - s| \text{ für alle } s, t \in [t_i, t^+). \tag{2.4.10}$$

Für eine beliebige Folge $t_k \nearrow t^+$ zeigt dies, dass $(x(t_k))$ eine Cauchyfolge in \mathbb{R}^n ist. Folglich existiert der Grenzwert

$$y = \lim_{k \rightarrow \infty} x(t_k),$$

und, da $\text{dist}(x(t), \partial D_0) \geq \varepsilon$ ist für t nahe t^+ , gilt $y \in D_0$. Ebenso folgt aus (2.4.10), dass

$$\lim_{k \rightarrow \infty} \int_{t_0}^{t_k} f(s, x(s)) ds$$

existiert. Der Grenzwert y hängt nicht von der betrachteten Folge ab: Denn für $s_k \nearrow t^+$ erhalten wir analog $x(s_k) \rightarrow z \in D_0$ und man erhält (wegen (2.4.10), weil $|t_k - s_k| \rightarrow 0$)

$$\|y - z\| = \lim_{k \rightarrow \infty} \|x(t_k) - x(s_k)\| \leq c \lim_{k \rightarrow \infty} |t_k - s_k| = 0.$$

Also gilt $y = \lim_{t \nearrow t^+} x(t)$. Die Lösung ist also in $t = t^+$ durch $\bar{x}(t^+) := y$ stetig ergänzbar. Nach Definition ist

$$y = \lim_{t \nearrow t^+} x(t) = \lim_{t \nearrow t^+} \left[x_0 + \int_{t_0}^t f(s, x(s)) ds \right] = x_0 + \int_{t_0}^{t^+} f(s, \bar{x}_{\max}(s)) ds.$$

Insbesondere ist $\bar{x}(\cdot)$ in $t = t^+$ linksseitig differenzierbar und erfüllt die Differentialgleichung. Daher ist $\bar{x}(\cdot)$ eine Fortsetzung der maximalen Lösung, im Widerspruch zur Definition von $x(\cdot)$.

Wir haben gezeigt, dass (2.4.7) nicht gelten kann. Es folgt für die maximale Lösung, dass

$$\lim_{t \nearrow t^+} \text{dist}(x(t), \partial D_0) = 0.$$

■

Wir führen jetzt eine Notation ein, die es erlaubt, die Abhängigkeit von den Anfangsbedingungen explizit zu machen. Wir wissen, dass Anfangswertprobleme für Differentialgleichungen mit rechter Seite, die stetig und Lipschitz-stetig bezüglich x sind, eindeutige Lösungen auf einem maximalen Existenzintervall $I_{\max}(t_0, x_0)$, das von den Anfangsbedingungen abhängt, besitzen.

Sei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : D \rightarrow \mathbb{R}^n$ stetig und Lipschitz-stetig bezüglich x . Für die Differentialgleichung

$$\dot{x} = f(t, x)$$

bezeichnen wir die auf

$$\Omega = \{(t, t_0, x_0) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n, (t_0, x_0) \in D \text{ und } t \in I_{\max}(t_0, x_0)\}$$

definierte Funktion mit

$$\lambda(t; t_0, x_0) := x_{\max}(t),$$

wobei $x_{\max}(\cdot)$ die maximale Lösung zur Anfangsbedingung $x(t_0) = x_0$ ist.

Manchmal (zum Beispiel in Aulbach [4]) wird dies auch als **allgemeine Lösung** dieser Differentialgleichung bezeichnet. In dieser Notation wird also die Abhängigkeit der Lösung von der Anfangszeit t_0 und dem Anfangspunkt x_0 betont. Der Einfachheit halber verwenden wir auch wieder unsere Notation

$$\dot{\lambda}(t; t_0, x_0) := \frac{d}{dt} \lambda(t; t_0, x_0).$$

Ein erstes Resultat zu den Eigenschaften dieser Funktion, also zu der Abhängigkeit von den Anfangsbedingungen, ist die folgende **Kozyklus-Eigenschaft**.

Proposition 2.4.11 *Unter den Voraussetzungen von Satz 2.4.7 sei $(t_0, x_0) \in D$. Dann gilt für jedes $t_1 \in I_{\max}(t_0, x_0)$*

- (i) $I_{\max}(t_1, \lambda(t_1; t_0, x_0)) = I_{\max}(t_0, x_0)$;
- (ii) $\lambda(t; t_1, \lambda(t_1; t_0, x_0)) = \lambda(t; t_0, x_0)$ für alle $t \in I_{\max}(t_0, x_0)$ und $\lambda(t_0; t_0, x_0) = x_0$.

Beweis. Wegen $t_1 \in I_{\max}(t_0, x_0)$ ist $(t_1, \lambda(t_1; t_0, x_0)) \in D$. Daher sind

$$x_1(t) := \lambda(t; t_0, x_0) \text{ und } x_2(t) := \lambda(t; t_1, \lambda(t_1; t_0, x_0))$$

nach Definition die maximalen Lösungen zu den Anfangswerten (t_0, x_0) bzw. $(t_1, \lambda(t_1; t_0, x_0))$, und für $t = t_1$ gilt

$$x_1(t_1) = \lambda(t_1; t_0, x_0) = \lambda(t_1; t_1, \lambda(t_1; t_0, x_0)) = x_2(t_1).$$

Wegen der Eindeutigkeit der Lösung dieses Anfangswertproblems sind daher $x_1(t)$ und $x_2(t)$ identisch, insbesondere sind ihre Existenzintervalle gleich. ■

Der globale Existenz- und Eindeutigkeitssatz, Satz 2.4.7, garantiert die Existenz von Lösungen auf einem maximalen (offenen) Existenzintervall. Theorem 2.4.10 zeigt, dass die Lösung unbeschränkt werden muss, wenn sie sich dem Rand des Existenzintervalls nähert (sofern sie sich nicht dem Rand des Definitionsbereichs von f nähert). Dies kann man ausnutzen, um die Existenz von Lösungen auf einem gegebenen Intervall zu erzwingen. Das Gronwall-Lemma wird einen einfachen Beweis des folgenden Existenzsatzes bei linear beschränkter Seite liefern.

Satz 2.4.12 *Es sei $D = (a, b) \times \mathbb{R}^n$ mit $-\infty \leq a < b \leq \infty$ und $f : D \rightarrow \mathbb{R}^n$ sei stetig und bezüglich x Lipschitz-stetig. Ferner sei f linear beschränkt, d.h.,*

$$\|f(t, x)\| \leq \rho(t) \|x\| + \sigma(t) \text{ für alle } t \in (a, b) \text{ und } x \in \mathbb{R}^n$$

mit stetigen Funktionen $\rho, \sigma : (a, b) \rightarrow [0, \infty)$. Dann existiert für jedes $(t_0, x_0) \in D$ die maximale Lösung von

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

auf $I_{\max} = (a, b)$.

Beweis. Sei $(t_0, x_0) \in D$. Betrachte die maximale Lösung $\lambda(t; t_0, x_0)$, $t \in I_{\max} = (t^-, t^+) \subset (a, b) =: J$. Dann gilt für alle $t \in I_{\max}$ mit $t \geq t_0$

$$\begin{aligned} \|\lambda(t; t_0, x_0)\| &\leq \|x_0\| + \int_{t_0}^t \|f(s, \lambda(s; t_0, x_0))\| ds \\ &\leq \|x_0\| + \int_{t_0}^t [\rho(s) \|\lambda(s; t_0, x_0)\| + \sigma(s)] ds \\ &= \|x_0\| + \int_{t_0}^t \sigma(s) ds + \int_{t_0}^t \rho(s) \|\lambda(s; t_0, x_0)\| ds. \end{aligned}$$

Ist $t^+ < b$, so gilt für jedes $t \in [t_0, t^+] \subset [t_0, b)$

$$\|x_0\| + \int_{t_0}^t \sigma(s) ds \leq \|x_0\| + \int_{t_0}^{t^+} \sigma(s) ds =: \alpha \text{ und } \rho(t) \leq \max\{\rho(s) \mid s \in [0, t^+]\} =: \beta.$$

Daher folgt nach Gronwalls Lemma, Lemma 2.2.7, aus der Ungleichung

$$\|\lambda(t; t_0, x_0)\| \leq \alpha + \int_{t_0}^t \beta \|\lambda(s; t_0, x_0)\| ds, \quad t \in [0, t^+],$$

dass

$$\|\lambda(t; t_0, x_0)\| \leq \alpha e^{\beta(t-t_0)} \leq \alpha e^{\beta(t^+-t_0)}, \quad t \in [t_0, t^+].$$

Also bleibt die Lösung für $t \rightarrow t^+$ beschränkt, im Widerspruch zu Satz 2.4.10 (mit $J = (a, b)$ und $D_0 = \mathbb{R}^n$).

Analog argumentiert man für $t^- = a$. Insgesamt folgt $I_{\max} = (a, b)$. ■

Eine unmittelbare Konsequenz ist der folgende Existenz- und Eindeutigkeitssatz für **inhomogene lineare Differentialgleichungen**.

Satz 2.4.13 Gegeben seien $-\infty \leq a < b \leq \infty$ und eine lineare inhomogene Differentialgleichung

$$\dot{x} = A(t)x + g(t)$$

mit stetigen Funktionen $A : (a, b) \rightarrow \mathbb{R}^{n \times n}$ und $g : (a, b) \rightarrow \mathbb{R}^n$. Dann existiert für jedes Anfangswertepaar $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$ die maximale Lösung auf dem Intervall (a, b) .

Beweis. Die rechte Seite $f(t, x) = A(t)x + g(t)$ ist stetig und bezüglich x Lipschitz-stetig wegen

$$\|f(t, x_1) - f(t, x_2)\| = \|A(t)x_1 - A(t)x_2\| \leq \|A(t)\| \|x_1 - x_2\|.$$

Ferner ist sie linear beschränkt, weil

$$\|f(t, x)\| \leq \|A(t)\| \|x\| + \|g(t)\| \quad \text{für alle } t \in (a, b) \text{ und } x \in \mathbb{R}^n.$$

Also folgt die Behauptung aus Theorem 2.4.12. ■

Bemerkung 2.4.14 Sind die Funktionen A und g auf einem abgeschlossenen Intervall definiert, so kann man offenbar auf ähnliche Weise Existenz und Eindeutigkeit von Lösungen nachweisen.

Kapitel 3

Autonome Differentialgleichungen

Ein wichtige Klasse von Differentialgleichungen erhält man, wenn die rechte Seite f nicht explizit von der Zeit t abhängt. Differentialgleichungen der Form $\dot{x} = f(x)$ mit einer stetigen Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen, heißen autonom. Die Voraussetzungen des globalen Existenz- und Eindeutigkeitssatzes sowie viele Annahmen und Aussagen vereinfachen sich für autonome Differentialgleichungen.

3.1 Grundlegende Eigenschaften

Wir betrachten autonome Differentialgleichungen der Form

$$\dot{x} = f(x) \tag{3.1.1}$$

mit einer stetigen Funktion $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$ offen. Wir setzen im Weiteren immer voraus, dass f Lipschitz-stetig ist. Die wichtigste Eigenschaft, die autonome Differentialgleichungen auszeichnet, ist die Translationsinvarianz der Lösungen.

Proposition 3.1.1 *Gegeben seien das autonome System (3.1.1) und $x_0 \in D$. Für alle $t_0, t_1 \in \mathbb{R}$ haben dann die Anfangswertprobleme*

$$\dot{x} = f(x), \quad x(t_0) = x_0 \text{ bzw. } x(t_1) = x_0,$$

die folgenden Eigenschaften:

- (i) $I_{\max}(t_1, x_0) = I_{\max}(t_0, x_0) - t_0 + t_1$;
- (ii) $\lambda(t; t_1, x_0) = \lambda(t + t_0 - t_1; t_0, x_0)$ für alle $t \in I_{\max}(t_0, x_0)$; insbesondere ist

$$\lambda(t; t_1, x_0) = \lambda(t - t_1; 0, x_0) \text{ für alle } t \in I_{\max}(t_0, x_0).$$

Beweis. Nach der Kettenregel ist $t \mapsto \lambda(t + t_0 - t_1; t_0, x_0)$ eine Lösung von $\dot{x} = f(x)$, $x(t_1) = \lambda(t_0; t_0, x_0) = x_0$, ist wegen der Eindeutigkeit also gleich $\lambda(t; t_1, x_0)$. Dann erfüllen die Definitionsbereiche die Eigenschaft (i), weil es sich um maximale Lösungen handelt. Die letzte Gleichung folgt für $t_0 = 0$. ■

Die Proposition zeigt insbesondere, dass man den Fall allgemeiner Anfangszeit t_0 auf Anfangszeit $t_0 = 0$ reduzieren kann.

Beispiel 3.1.2 *Für die Gleichung $\dot{x} = 1 + x^2$ ist die allgemeine Lösung*

$$\lambda(t; t_0, x_0) = \tan(t - t_0 + \arctan x_0), \quad t \in (t_0 - \arctan x_0 - \frac{\pi}{2}, t_0 - \arctan x_0 + \frac{\pi}{2}).$$

Die Translationsinvarianz zeigt sich darin, dass man diese Lösungen aus

$$\lambda(t; 0, x_0) = \tan(t + \arctan x_0), \quad t \in (-\arctan x_0 - \frac{\pi}{2}, -\arctan x_0 + \frac{\pi}{2})$$

durch Translation um t_0 erhält.

Bemerkung 3.1.3 Sind umgekehrt alle maximalen Lösungen einer Differentialgleichung $\dot{x} = f(t, x)$ translationsinvariant, so ist $f(t_1, x_0) = f(t_0, x_0)$ für alle t_0, t_1 und alle x_0 , d.h., f ist unabhängig von t . In der Tat: Translationsinvarianz impliziert für alle x_0 und alle t_0, t_1

$$f(t, \lambda(t; t_1, x_0)) = \dot{\lambda}(t; t_1, x_0) = \dot{\lambda}(t + t_0 - t_1; t_0, x_0) = f(t + t_0 - t_1, \lambda(t + t_0 - t_1; t_0; x_0)),$$

also für $t = t_1$

$$f(t_1, x_0) = f(t_0, x_0).$$

Translationsinvarianz der Lösungen charakterisiert also autonome Differentialgleichungen.

Das folgende einfache Resultat beleuchtet die Beziehungen zwischen nicht-autonomen und autonomen Differentialgleichungen.

Proposition 3.1.4 Es sei $x(t), t \in I$, eine Lösung des Anfangswertproblems

$$\dot{x} = g(t, x), \quad x(t_0) = x_0, \quad (3.1.2)$$

mit einer stetigen Funktion $g: D \rightarrow \mathbb{R}^n$, wobei $D \subset \mathbb{R} \times \mathbb{R}^n$ offen ist. Dann ist

$$(x(t), t), \quad t \in I,$$

eine Lösung des folgenden Anfangswertproblems für eine autonome Differentialgleichung in \mathbb{R}^{n+1}

$$\begin{pmatrix} \dot{x} \\ \dot{x}_{n+1} \end{pmatrix} = f(x, x_{n+1}) = \begin{pmatrix} g(x_{n+1}, x) \\ 1 \end{pmatrix}, \quad \begin{pmatrix} x(t_0) \\ x_{n+1}(t_0) \end{pmatrix} = \begin{pmatrix} x_0 \\ t_0 \end{pmatrix}; \quad (3.1.3)$$

hierbei ist f auf D definiert. Umgekehrt sind für jede Lösung $(x, x_{n+1})^T$ dieses Anfangswertproblems die ersten n Komponenten eine Lösung von (3.1.2).

Beweis. Folgt sofort durch Einsetzen. ■

Diese Proposition zeigt, dass man sich für allgemeine Aussagen, wie Existenz und Eindeutigkeit von Lösungen, maximales Existenzintervall, u.a., auf autonome Differentialgleichungen beschränken kann. Wir hätten uns also bei der Diskussion der Existenz und Eindeutigkeit von Lösungen im letzten Kapitel auf autonome Systeme beschränken können (allerdings hätten wir dann auch Lipschitz-Stetigkeit bezüglich t fordern müssen). Für viele Probleme ist diese Autonomisierung aber nicht geeignet, weil die Lösungen der autonomisierten Gleichung (3.1.3) unbeschränkt sind, wenn ihr Existenzintervall unbeschränkt ist. Später werden wir gerade die Beschränktheit von Lösungen voraussetzen, um ihr Verhalten besser zu verstehen.

Bemerkung 3.1.5 In den Beweisen des letzten Kapitels haben wir mehrmals benutzt, dass für stetige f Lösungen des Anfangswertproblems

$$\dot{x} = f(t, x), \quad x(t_0) = x_0,$$

auch charakterisiert sind als Lösungen der Integralgleichung

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) \, ds. \quad (3.1.4)$$

Diese Integralgleichung macht offenbar auch Sinn, wenn f nicht stetig in t ist, also zum Beispiel nur stückweise konstant in t ist. Solche Probleme sind typisch für Anwendungen in der Kontrolltheorie (oder „Regelungstheorie“), wo stückweise stetige Steuerungen benutzt werden. Zum Beispiel kann man die Bewegung eines gesteuerten Wagens durch

$$\ddot{x}(t) = u(t), \quad u(t) \in \{-1, 1\},$$

beschreiben (dabei bedeutet $u(t) = 1$ Vollgas, $u(t) = -1$ Vollbremsung). Auch für eine stückweise stetige Funktion $u(\cdot)$ liefert die integrierte Form eine Lösung. Eine Reduktion auf eine autonome Gleichung wie in Proposition 3.1.4 ist hier nicht möglich.

Ein anderes Beispiel ist die Pendelgleichung

$$\ddot{x} + k\dot{x} + \sin x = u(t),$$

wobei wieder die Steuerung u zwischen den Werten $+1$ und -1 umschalten kann. Nach der Standardreduktion auf eine Differentialgleichung erster Ordnung in \mathbb{R}^2 führt dies, für gegebene Steuerung $t \mapsto u(t)$, zu einer Differentialgleichung, bei der die rechte Seite

$$f(x, y, t) = \begin{pmatrix} y \\ -ky - \sin x + u(t) \end{pmatrix}$$

nicht stetig von t abhängt. Diese nichtautonome Gleichung kann durch den obigen Trick nicht auf eine autonome Gleichung zurückgeführt werden. Allgemeiner betrachtet man Differentialgleichungen, bei denen die rechte Seite messbar von t und stetig von x abhängt. Dies führt (unter zusätzlichen Integrierbarkeitsbedingungen) zu sogenannten Carathéodory-Lösungen, die nur als Lösungen der zugehörigen Integralgleichung (3.1.4) erklärt sind.

Die Kozykel-Eigenschaft vereinfacht sich, wie in Proposition 3.1.1 gezeigt, auf Grund der Translations-Invarianz im autonomen Fall wesentlich. Es reicht, nur Lösungen zu betrachten, die zur Zeit $t_0 = 0$ in einem Punkt x_0 starten. Dies führt zu der grundlegenden Begriffsbildung des **Flusses** einer autonomen Differentialgleichung. Zunächst führen wir die folgende Notation für die Systeme der Form (3.1.1) ein.

$$J_{\max}(x_0) := I_{\max}(0, x_0), \quad \varphi(t; x_0) := \lambda(t; 0, x_0) \text{ für alle } t \in J_{\max}(x_0).$$

Dann gilt für jedes $t_0 \in \mathbb{R}$ und $x_0 \in D$

$$\lambda(t; t_0, x_0) = \varphi(t - t_0; x_0) \text{ für alle } t \in J_{\max}(x_0) + t_0.$$

Wir nennen die zu (3.1.1) gehörende Abbildung

$$\varphi : \{(t, x_0) \in \mathbb{R} \times \mathbb{R}^n, x_0 \in D \text{ und } t \in J_{\max}(x_0)\} \rightarrow \mathbb{R}^n$$

den zugehörigen **lokalen Fluss**. Ist φ auf $\mathbb{R} \times D$ definiert, so sprechen wir von dem zugehörigen globalen Fluss oder einfach dem zugehörigen Fluss. Der lokale Fluss hat die folgenden Eigenschaften.

Proposition 3.1.6 Für den lokalen Fluss $\varphi(t; x_0)$, $x_0 \in D$ und $t \in J_{\max}(x_0) = (t^-, t^+)$, gilt:

- (i) $-\infty \leq t^- < 0 < t^+ \leq \infty$;
- (ii) $J_{\max}(\varphi(t; x_0)) = J_{\max}(x_0) - t$ für alle $t \in J_{\max}(x_0)$;
- (iii) $\varphi(0; x_0) = x_0$;
- (iv) $\varphi(t; \varphi(s; x_0)) = \varphi(t + s; x_0)$ für alle s, t mit $s, t + s \in J_{\max}(x_0)$;
- (v) $\varphi(-t; \varphi(t; x_0)) = x_0$ für alle t mit $t \in J_{\max}(x_0)$.

Beweis. Dies sind direkte Folgerungen aus den Kozyklus-Eigenschaften zusammen mit der Translationsinvarianz von Lösungen. ■

Ein globaler Fluss erfüllt für alle $t, s \in \mathbb{R}$ und alle $x_0 \in D$

$$\varphi(t; \varphi(s; x_0)) = \varphi(t + s; x_0) \text{ für } t, s \in \mathbb{R} \text{ und } \varphi(0; x_0) = x_0. \quad (3.1.5)$$

3.2 Trajektorien

Bisher haben wir Lösungen immer durch ihre Graphen in $\mathbb{R} \times \mathbb{R}^n$ illustriert. Für autonome Gleichungen ist eine andere graphische Darstellung häufig geeigneter.

Beispiel 3.2.1 Betrachte die lineare Differentialgleichung in \mathbb{R}^2

$$\dot{x} = y, \quad \dot{y} = -x,$$

mit Anfangsbedingung

$$x(0) = x_0, \quad y(0) = y_0,$$

also

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Man sieht leicht (durch Einsetzen), dass die Lösung gegeben ist durch

$$\varphi(t; x_0, y_0) = \lambda(t; 0, x_0, y_0) = \begin{pmatrix} x_0 \cos t + y_0 \sin t \\ y_0 \cos t - x_0 \sin t \end{pmatrix}.$$

Die Lösungen kann man in $\mathbb{R} \times \mathbb{R}^2$ oder als durch t parametrisierte Kurven in \mathbb{R}^2 darstellen.

Dies motiviert die folgende allgemeine Definition.

Definition 3.2.2 Betrachte die autonome Differentialgleichung $\dot{x} = f(x)$. Die Trajektorie oder der Orbit durch $x_0 \in D$ ist

$$\mathcal{O}(x_0) = \{\varphi(t; x_0) \in \mathbb{R}^n, t \in J_{\max}(x_0)\}.$$

Ferner sind

$$\mathcal{O}^+(x_0) = \{\varphi(t; x_0) \in \mathbb{R}^n, t \in J_{\max}(x_0) \text{ und } t \geq 0\}$$

und

$$\mathcal{O}^-(x_0) = \{\varphi(t; x_0) \in \mathbb{R}^n, t \in J_{\max}(x_0) \text{ und } t \leq 0\}$$

die positive bzw. negative Halbtrajektorie (der positive bzw. negative Orbit) durch x_0 .

Die Trajektorie ist eine Kurve in \mathbb{R}^n , die durch t parametrisiert ist. Die Lösung (oder Lösungskurve) ist eine Kurve in $\mathbb{R} \times \mathbb{R}^n$; sie enthält mehr Information als die Trajektorie. Betrachtet man alle Lösungskurven in \mathbb{R}^n , so spricht man vom Phasenporträt der Differentialgleichung (eine genauere Diskussion erfolgt in Abschnitt 3.3).

Beispiel 3.2.3

$$\dot{x} = 1 + x^2 \text{ und } \dot{x} = 1.$$

Die Lösungen sind für $x_0 \in \mathbb{R}$ gegeben durch

$$\varphi(t; x_0) = \tan(t - \arctan x_0), t \in \left(-\frac{\pi}{2} - \arctan x_0, \frac{\pi}{2} - \arctan x_0\right) \text{ bzw. } \varphi(t; x_0) = t + x_0, t \in \mathbb{R}.$$

Im Gegensatz zu den Lösungen (d.h. ihren Graphen in $\mathbb{R} \times \mathbb{R}$) stimmen die Trajektorien in $\mathbb{R}^n, n = 1$, überein: Die einzige Trajektorie ist \mathbb{R} .

Beispiel 3.2.4 Betrachte

$$\dot{x} = x - x^2.$$

Der zugehörige Fluss ist

$$\varphi(t; x_0) = x_0 / (x_0 + (1 - x_0)e^{-t}),$$

wie man durch Einsetzen leicht nachrechnet. Dann gilt

$$\varphi(t; 0) = 0 \text{ und } \varphi(t; 1) = 1 \text{ für alle } t \in \mathbb{R}.$$

Es gibt fünf Trajektorien:

$$(-\infty, 0), \{0\}, (0, 1), \{1\}, (1, \infty).$$

Die folgende Proposition gibt eine Klassifikation der Trajektorien für autonome Differentialgleichungen.

Proposition 3.2.5 Sei $x_0 \in D$ und $t \in J_{\max}(x_0)$. Dann gilt für die Trajektorie $\mathcal{O}(x_0)$ genau einer der folgenden Fälle:

(i) $J_{\max}(x_0) = (-\infty, \infty)$ und $\varphi(t; x_0)$ ist konstant, $\mathcal{O}(x_0) = \{x_0\}$; dann heißt x_0 Gleichgewicht oder Ruhelage.

(ii) $J_{\max}(x_0) = (-\infty, \infty)$ und $\varphi(t; x_0)$ ist nicht konstant und periodisch, $\mathcal{O}(x_0) = \mathcal{O}^+(x_0) \neq \{x_0\}$.

(iii) $\varphi(\cdot; x_0)$ ist injektiv, also ist $\mathcal{O}(x_0)$ eine Doppelpunkt-freie Kurve.

Beweis. Wir zeigen zunächst, dass jede nicht-injektive Lösung $\varphi(t; x_0)$ periodisch ist, es also $T > 0$ gibt mit $\varphi(t; x_0) = \varphi(t + kT; x_0)$ für alle $t \in [0, T]$ und $k \in \mathbb{Z}$. Gelte

$$\varphi(t_1; x_0) = \varphi(t_0; x_0) \text{ für } t_0 \neq t_1 \text{ in } J_{\max}(x_0).$$

Dann ist wegen Translationsinvarianz

$$x(t) := \varphi(t + t_0 - t_1; x_0)$$

Lösung auf $J_{\max}(x_0) - t_0 + t_1$ und

$$x(t_1) = \varphi(t_0; x_0) = \varphi(t_1; x_0).$$

Also ist $x(t) = \varphi(t; x_0)$ für alle $t \in J_{\max}(x_0) = J_{\max}(x_0) - t_0 + t_1$. Wegen $t_1 \neq t_0$ folgt, dass $J_{\max} = (-\infty, \infty)$. Ferner folgt die Periodizität mit $T = |t_0 - t_1|$ aus

$$\varphi(t; x_0) = \varphi(t + k(t_0 - t_1); x_0) \text{ für alle } t \in \mathbb{R} \text{ und } k \in \mathbb{Z}.$$

Ist $\varphi(\cdot; x_0)$ nicht-injektiv, so ist $\varphi(t; x_0)$ periodisch, also entweder konstant, oder periodisch und nicht konstant. Damit sind die Behauptungen gezeigt. ■

Beispiel 3.2.6 Betrachte

$$\begin{aligned} \dot{x} &= y + x(1 - x^2 - y^2) \\ \dot{y} &= -x + y(1 - x^2 - y^2). \end{aligned}$$

Der zugehörige Fluss in \mathbb{R}^2 ist

$$\varphi(t; x_0, y_0) = \frac{1}{\sqrt{x_0^2 + y_0^2 + (1 - x_0^2 - y_0^2)e^{-2t}}} \begin{pmatrix} x_0 \cos t + y_0 \sin t \\ y_0 \cos t - x_0 \sin t \end{pmatrix}.$$

Der Ursprung ist ein Gleichgewicht, jede Lösung mit Startwert auf dem Einheitskreis, also mit $(x_0)^2 + (y_0)^2 = 1$, liefert eine periodische Lösung, deren Trajektorie mit dem Einheitskreis übereinstimmt. Die anderen Trajektorien sind doppeltpunktfrei.

Proposition 3.2.7 Ist $n = 1$, also die Differentialgleichung skalar, so gibt es keine nichtkonstanten periodischen Trajektorien.

Beweis. Sei $x : \mathbb{R} \rightarrow \mathbb{R}$ eine T -periodische Lösung. Dann existiert $t_0 \in \mathbb{R}$ mit $\dot{x}(t_0) = 0$, weil die stetige Funktion x auf der kompakten Menge $[-1, T + 1]$ ihr Maximum (wegen der Periodizität) im Innern annimmt. Daher ist

$$0 = \dot{x}(t_0) = f(x(t_0)).$$

Wegen der Eindeutigkeit der Lösung folgt $x(t) = x(t_0)$ für alle t . ■

3.3 Phasenporträts und Richtungs-Vektorfeld

Die folgende Proposition zeigt, dass jeder Punkt des Definitionsbereichs $D \subset \mathbb{R}^n$ auf genau einer Trajektorie liegt.

Proposition 3.3.1 *Die auf D definierte Relation*

$$x_0 \sim y_0 \text{ falls } t_0 \in J_{\max}(x_0) \text{ existiert mit } y_0 = \varphi(t_0; x_0),$$

*ist eine Äquivalenzrelation, und für jedes $x_0 \in D$ ist die Trajektorie die von x_0 erzeugte Äquivalenzklasse. Diese Zerlegung von D nennt man das **Phasenporträt** der Differentialgleichung.*

Beweis. Reflexivität folgt aus $0 \in J_{\max}(x_0)$ und $\varphi(0; x_0) = x_0$ für alle $x_0 \in D$. Symmetrie gilt, weil aus $y_0 = \varphi(t_0; x_0)$ folgt, dass $x_0 = \varphi(-t_0; y_0)$. Transitivität schließlich folgt, weil

$$\varphi(t_1; \varphi(t_0; x_0)) = \varphi(t_1 + t_0; x_0).$$

■

Diese Zerlegung des Definitionsbereichs D kann genutzt werden, um einen Überblick über das Verhalten des Systems zu gewinnen. Man erhält das sogenannte **Phasenporträt** der Differentialgleichung. Allerdings muss man dabei im Auge behalten, dass - im Gegensatz zu den Lösungen - die Trajektorien keine Information über die Zeit enthalten. Am augenfälligsten wird das an der Tatsache, dass Differentialgleichungen mit endlichen Entweichzeiten, also beschränkten maximalen Existenzintervallen, das gleiche Phasenporträt wie solche mit unbeschränkten Existenzintervallen haben können. Ein Beispiel dafür liefern die skalaren Differentialgleichungen (vgl. Beispiel 3.2.3)

$$\dot{x} = 1 + x^2 \text{ und } \dot{x} = 1.$$

Die einzige Äquivalenzklasse ist \mathbb{R} .

Am wichtigsten sind Phasenporträts natürlich für $n = 2$, also für Differentialgleichungen in der Ebene. Wir werden später insbesondere die Phasenporträts linearer Differentialgleichungen in der Ebene ausführlich diskutieren.

Eine alternative Information über Differentialgleichungen liefert die Abbildung $f : D \rightarrow \mathbb{R}^n$ für eine autonome Differentialgleichung $\dot{x} = f(x)$. Man kann sich diese Abbildung veranschaulichen, indem man sich an jeden Punkt $x \in D$ den Vektor $f(x)$ angeheftet denkt. Daher heißt diese Abbildung auch Vektorfeld. Richtung und Länge des Vektors geben Aufschluss über die Lösung:

Sei $x_0 \in D$ und $\varphi(t; y_0)$, $t \in I$, eine maximale Lösung mit $x_0 \in \mathcal{O}(y_0)$, also $x_0 = \varphi(t_0; y_0)$ für ein $t_0 \in I$. Dann ist, wie aus der Analysis bekannt,

$$\frac{d}{dt} \varphi(t_0; y_0)$$

der Tangentialvektor an die durch t parametrisierte Kurve $\varphi(t; y_0)$, $t \in I$, im Kurvenpunkt $\varphi(t_0; y_0)$. Auf Grund der Differentialgleichung ist aber

$$\frac{d}{dt} \varphi(t_0; y_0) = f(\varphi(t_0; y_0)) = f(x_0).$$

Also ist $f(x_0)$ dieser Tangentialvektor. Insgesamt ordnet daher das Vektorfeld f jedem Punkt $x \in D$ den Tangentialvektor an die Trajektorie zu.

Man sieht, dass Definitions- und Bildbereich von $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ völlig unterschiedliche Rollen spielen. Das wird in der Differentialgeometrie präzisiert, wo der Zustandsraum einer Differentialgleichung eine Mannigfaltigkeit M ist, und die rechte Seite jedem Element von M einen Tangentialvektor zuordnet.

Bemerkung 3.3.2 *Man nennt die Trajektorien auch die zu dem Vektorfeld f gehörenden Integralkurven. Der Vektor $f(x_0)$ ist jeweils tangential an diese Kurve in x_0 . Analoge Begriffsbildungen werden in der Differentialgeometrie auch für höherdimensionale Objekte, etwa Integral-Flächen M , eingeführt. Offenbar braucht man für die Erzeugung dieser zwei-dimensionalen Gebilde zwei Vektorfelder, die dann jeweils tangential an M sind.*

Das folgende Beispiel illustriert die Beziehungen zwischen linearen und gewissen nichtlinearen Differentialgleichungen.

Beispiel 3.3.3 Der lineare (oder harmonische) Oszillator $\ddot{x} = -x$ hat, nach Reduktion auf die Gleichung erster Ordnung in \mathbb{R}^2 ,

$$\dot{x} = y, \quad \dot{y} = -x,$$

das bekannte Phasenporträt in der (x, \dot{x}) -Ebene. Im 18. Jahrhundert studierte J.F. Riccati die Frage, ob man für die Steigung längs einer Lösung,

$$v(t) := \frac{y(t)}{x(t)},$$

eine Differentialgleichung herleiten kann. In der Tat liefert die Kettenregel

$$\frac{d}{dt}v(t) = \dot{v} = \frac{\dot{y}x - y\dot{x}}{x^2} = -\frac{x^2}{x^2} - \frac{y^2}{x^2} = -1 - v^2.$$

Also erfüllt die Steigung die Differentialgleichung

$$\dot{v} = -1 - v^2,$$

solange $x(t) \neq 0$. Für $x(t) \rightarrow 0$ gilt offenbar $v(t) \rightarrow \infty$. Dies liefert eine Erklärung, warum wir die Lösungen dieser Differentialgleichung, die uns schon in Beispiel 3.1.2 begegnet ist, nur auf beschränkten Zeitintervallen existieren. Der Fluss ist gegeben durch

$$\varphi(t; x_0) = \tan(-t + \arctan x_0), \quad t \in \left(-\arctan x_0 - \frac{\pi}{2}, -\arctan x_0 + \frac{\pi}{2}\right).$$

Auf Grund der Beziehung zur linearen Differentialgleichung ist also klar, warum die Lösungen dieser Riccati-Differentialgleichung „endliche Entweichzeit“ haben: die Steigung wird in endlicher Zeit unbeschränkt (wie man auch an Hand der Lösungsformel sofort sieht).

Die Steigung legt eine Gerade durch den Ursprung eindeutig fest. Geometrisch betrachtet, haben wir die Geraden durch den Ursprung durch ihre Steigung parametrisiert. Diese Parametrisierung bricht für Geraden mit unendlicher Steigung zusammen. Parametrisiert man Geraden durch den Ursprung durch den Winkel

$$\theta \in \mathbb{R} \text{ modulo } \pi,$$

so sieht man, dass es nur die Parametrisierung durch die Steigung ist, die zu endlicher Entweichzeit führt. Geometrisch ist der Fluss auf der Menge der Geraden durch den Ursprung für alle Zeiten wohldefiniert. Die Geraden durch den Ursprung sind die eindimensionalen Unterräume in \mathbb{R}^2 , sie bilden den **projektiven Raum**.

Dieses einfache Beispiel zeigt, dass manchmal der „natürliche“ Zustandsraum einer Differentialgleichung nicht der \mathbb{R}^n , sondern eine Mannigfaltigkeit ist. Riccati-Gleichungen spielen in vielen modernen Anwendungen der Mathematik wie Computersehen oder Kontrolltheorie eine große Rolle (die obige Konstruktion lässt sich auf andere zwei- oder höherdimensionale lineare autonome Differentialgleichungen verallgemeinern und führt dann auf Matrizen-Differentialgleichungen, die das Verhalten von Unterräumen beschreiben).

Kapitel 4

Skalare Nichtautonome Differentialgleichungen

In den vorhergehenden Kapiteln haben wir die grundlegenden Fragen nach Existenz und Eindeutigkeit der Lösungen von Anfangswertproblemen geklärt. Dabei haben wir keine Lösungsformeln verwendet, weil diese häufig nicht bekannt sind oder nicht existieren. Im Allgemeinen muss man dann für die Berechnung von Lösungen auf numerische Verfahren zurückgreifen (eine erste Idee dafür liefern die Euler-Polygone, die wir im Beweis des Satzes von Peano kennengelernt haben, und die Picard-Iteration). In diesem Kapitel werden wir für skalare, nicht-autonome Differentialgleichungen Techniken beschreiben, die in vielen Fällen die explizite Lösung ermöglichen. Es handelt sich um klassische Techniken, die eine lange Geschichte haben; zum Teil sind sie auch in Computerprogrammen (wie MAPLE oder MATHEMATICA) implementiert. Ist solch eine Lösung verfügbar, so ist dies allen numerischen Verfahren weit überlegen.

4.1 Problemformulierung

An Stelle der bisher betrachteten Differentialgleichungen der Form

$$\dot{x} = f(t, x)$$

betrachten wir in diesem Kapitel *implizite* skalare nichtautonome Differentialgleichungen der Form

$$\dot{x}g(t, x) + h(t, x) = 0 \tag{4.1.1}$$

in $\mathbb{R} \times \mathbb{R}$. Dabei sind im Weiteren immer $\tilde{D} \subset \mathbb{R} \times \mathbb{R}$ eine offene Menge und $g, h : \tilde{D} \rightarrow \mathbb{R}$ stetig und bezüglich x Lipschitz-stetig. Ist $g(t, x) \neq 0$, so ist (4.1.1) offenbar äquivalent zu der expliziten Differentialgleichung

$$\dot{x} = -\frac{h(t, x)}{g(t, x)}. \tag{4.1.2}$$

(Äquivalent heißt, dass die Lösungen der beiden Differentialgleichungen übereinstimmen, falls für sie $g \neq 0$ ist). Daher führen wir den eingeschränkten Definitionsbereich

$$D = \{(t, x) \in \tilde{D}, g(t, x) \neq 0\}$$

ein. Auf D ist auch $\frac{h(t, x)}{g(t, x)}$ Lipschitz-stetig (siehe zum Beispiel Aulbach [4, Satz 2.4.7]. Daher gelten auf D alle Aussagen, die wir über Existenz und Eindeutigkeit von Lösungen für explizite Differentialgleichungen bewiesen haben. Insbesondere existiert für eine Anfangsbedingung

$$x(\tau) = \xi, (\tau, \xi) \in D \subset \mathbb{R} \times \mathbb{R}$$

eine eindeutige maximale Lösung $\lambda(t; \tau, \xi)$ auf dem Existenzintervall $I_{\max}(\tau, \xi)$ (in diesem Kapitel verwenden wir standardmäßig diese Buchstaben für Anfangsbedingungen). Sei $k : D \rightarrow \mathbb{R}$ stetig mit $k(t, x) \neq 0$ für alle (t, x) . Dann ist (4.1.2) offenbar äquivalent zu

$$\dot{x} = -\frac{h(t, x)k(t, x)}{g(t, x)k(t, x)}.$$

Diese triviale Beobachtung wird sich später als sehr nützlich erweisen.

Der skalare Fall mag als nicht besonders interessant erscheinen. Aber weil wir nichtautonome Gleichungen betrachten, haben wir eigentlich ein zwei-dimensionales Problem. Ferner werden wir später sehen, dass diese ein-dimensionalen nichtautonomen Differentialgleichungen benutzt werden können, um zwei-dimensionale autonome Differentialgleichungen wie die Pendelgleichung zu studieren.

4.2 Exakte Differentialgleichungen und Trennung der Variablen

Wir beginnen mit der folgenden Definition.

Definition 4.2.1 Die Differentialgleichung (4.1.1) heißt *exakt* auf \tilde{D} , falls es eine Stammfunktion genannte stetig differenzierbare Funktion (C^1 -Funktion) $S : \tilde{D} \rightarrow \mathbb{R}$ gibt mit

$$\frac{\partial S}{\partial t}(t, x) = h(t, x) \text{ und } \frac{\partial S}{\partial x}(t, x) = g(t, x) \text{ für alle } (t, x) \in \tilde{D}. \quad (4.2.1)$$

Was nutzt die Kenntnis einer Stammfunktion?

Satz 4.2.2 Die Differentialgleichung (4.1.1) sei exakt mit der Stammfunktion S . Dann gilt auf dem eingeschränkten Definitionsbereich D :

(i) die Lösung erfüllt

$$S(t, \lambda(t; \tau, \xi)) = S(\tau, \xi) \text{ für alle } (\tau, \xi) \in D, t \in I_{\max}(\tau, \xi).$$

(ii) Für alle $(\tau, \xi) \in D$ ist die Niveaumenge

$$N(\tau, \xi) := \{(t, x) \in D, S(t, x) = S(\tau, \xi)\}$$

Vereinigung von Lösungskurven.

Beweis. (i) Nach Definition gilt für jedes (τ, ξ)

$$\dot{\lambda}(t; \tau, \xi)g(t, \lambda(t; \tau, \xi)) + h(t; \lambda(t; \tau, \xi)) \equiv 0.$$

Mit (4.2.1) und der Kettenregel folgt für alle t

$$\begin{aligned} \frac{d}{dt}S(t, \lambda(t; \tau, \xi)) &= \frac{\partial S}{\partial t}(t, \lambda(t; \tau, \xi)) + \frac{\partial S}{\partial x}(t, \lambda(t; \tau, \xi))\dot{\lambda}(t; \tau, \xi) \\ &= h(t, \lambda(t; \tau, \xi)) + g(t, \lambda(t; \tau, \xi))\dot{\lambda}(t; \tau, \xi) \\ &= 0. \end{aligned}$$

Also ist S konstant längs $(t, \lambda(t; \tau, \xi))$.

(ii) Folgt aus (i). ■

Der Satz zeigt, dass zu $(\tau, \xi) \in D$ die Lösungskurve $(t, \lambda(t; \tau, \xi)), t \in I_{\max}(\tau, \xi)$, Teilmenge der Niveaumenge $N(\tau, \xi)$ von S ist. Die Niveaumengen von S sind nach dem Theorem über Implizite Funktionen Kurven, falls die Ableitung von S nach x regulär ist, d.h. $\frac{\partial S}{\partial x}(t, x) = g(t, x) \neq 0$. In den Punkten aus $\tilde{D} \setminus D$ (dies sind die Punkte mit $g(t, x) = 0$ ist) haben wir allerdings keine

Information. In den anderen Punkten ist die Lösungskurve aber durch die Niveaumenge von S eindeutig gegeben. Etwas formaler ausgedrückt: Mit

$$\tilde{N}(\tau, \xi) := \{(t, x) \in \tilde{D}, S(t, x) = S(\tau, \xi)\}$$

gilt

$$N(\tau, \xi) := \tilde{N}(\tau, \xi) \cap D$$

und nur in den Punkten aus $\tilde{N}(\tau, \xi) \setminus N(\tau, \xi)$ kann die Niveaumenge kompliziert sein.

Beispiel 4.2.3 Betrachte

$$\dot{x}t + x + t^2 = 0.$$

Die Gleichung ist exakt, weil auf \mathbb{R}^2

$$S(t, x) = tx + \frac{t^3}{3}$$

eine C^1 -Stammfunktion ist,

$$\frac{\partial S}{\partial x}(t, x) = t \text{ und } \frac{\partial S}{\partial t}(t, x) = x + t^2.$$

Hier ist $\tilde{D} = \mathbb{R}^2$ und der eingeschränkte Definitionsbereich D ist

$$D = \{(t, x) \in \mathbb{R}^2 \mid t \neq 0\}.$$

Hier sind die Niveaumengen

$$\tilde{N}(\tau, \xi) = \left\{ (t, x) \in \tilde{D} = \mathbb{R}^2 \mid tx + \frac{t^3}{3} = \tau\xi + \frac{\tau^3}{3} \right\}$$

und daher (Auflösen nach x)

$$\tilde{N}(\tau, \xi) = \begin{cases} \left\{ (t, x) \in \mathbb{R}^2 \mid x = \frac{1}{t} \left(\tau\xi + \frac{\tau^3}{3} \right) - \frac{t^2}{3} \text{ und } t \neq 0 \right\}, & \text{falls } \tau\xi + \frac{\tau^3}{3} \neq 0 \\ \{\{0\} \times \mathbb{R}\} \cup \left\{ \left(t, -\frac{t^2}{3} \right) \mid t \in \mathbb{R} \right\}, & \text{falls } \tau\xi + \frac{\tau^3}{3} = 0 \end{cases}.$$

Im ersten Fall besteht $\tilde{N}(\tau, \xi)$ aus zwei Kurven, von denen nur eine die Lösungskurve durch (τ, ξ) enthält (man muss $t > 0$ und $t < 0$ unterscheiden; vgl. hierzu auch Aubach [4], Kapitel 4). Man erhält hier für $\tau \neq 0, \xi \in \mathbb{R}$ sogar explizit die allgemeine Lösung. Denn aus

$$N(t, \xi) = \left\{ (t, x) \in \mathbb{R}^2 \mid tx + \frac{t^3}{3} = \tau\xi + \frac{\tau^3}{3} \text{ und } t \neq 0 \right\}$$

folgt nach Satz 4.2.2 und Satz 2.4.10

$$\lambda(t; \tau, \xi) = \frac{1}{t} \left(\tau\xi + \frac{\tau^3}{3} \right) - \frac{t^2}{3}, \quad t \neq 0.$$

Im Allgemeinen kann man allerdings $S(t, x) = S(\tau, \xi)$ nicht nach x auflösen, und erhält deshalb keine Formel für die allgemeine Lösung. Im folgenden Beispiel erhält man dennoch Informationen über die Lösungen.

Beispiel 4.2.4 Betrachte

$$S(t, x) = txe^x$$

als Stammfunktion einer Differentialgleichung. Diese ist dann gegeben durch

$$\dot{x}te^x(1+x) + xe^x = 0,$$

weil $\frac{\partial S}{\partial x}(t, x) = te^x(1+x) = g(t, x)$ und $\frac{\partial S}{\partial t}(t, x) = xe^x = h(t, x)$. Hier ist

$$D = \{(t, x) \in \mathbb{R}^2, t(1+x) \neq 0\}.$$

Auflösung von $S(t, x) = S(\tau, \xi)$ nach x ist nicht möglich. Aber Auflösung nach t ist für $x \neq 0$ möglich. Man erhält für $\xi \neq 0$

$$N(\tau, \xi) = \{(t, x) \in D \mid txe^x = \tau\xi e^\xi \neq 0\} = \left\{ \left(\frac{\tau\xi}{x} e^{\xi-x}, x \right) \in \mathbb{R}^2 \mid x \neq 0, -1 \right\}$$

und für $\xi = 0$

$$N(\tau, 0) = \{(t, x) \in D \mid txe^x = 0\} = \{(t, 0) \in \mathbb{R}^2 \mid t \neq 0\}.$$

Bis auf die beiden horizontalen Lösungskurven (die beiden t -Halbachsen) lassen sich also alle Lösungskurven als Graphen von Funktionen mit t in Abhängigkeit von x explizit darstellen.

Es ergeben sich die folgenden Fragen:

1. Wie stellt man fest, ob eine Differentialgleichung exakt ist?
2. Wie berechnet man gegebenenfalls eine Stammfunktion?

Sind g und h sogar C^1 -Funktionen, so folgt aus der Exaktheit, dass $S(t, x)$ sogar eine C^2 -Funktion ist mit

$$\frac{\partial g}{\partial t}(t, x) = \frac{\partial^2 S}{\partial t \partial x}(t, x) = \frac{\partial^2 S}{\partial x \partial t}(t, x) = \frac{\partial h}{\partial x}(t, x).$$

Also muss die folgende Integrabilitätsbedingung

$$\frac{\partial g}{\partial t}(t, x) = \frac{\partial h}{\partial x}(t, x) \tag{4.2.2}$$

gelten. Das folgende Beispiel zeigt, dass die Bedingung (4.2.2) zwar notwendig, aber nicht hinreichend für die Exaktheit ist.

Beispiel 4.2.5 Betrachte auf $\tilde{D} = \mathbb{R}^2 \setminus \{(0, 0)\}$

$$\dot{x} \frac{t}{t^2 + x^2} - \frac{x}{t^2 + x^2} = 0. \tag{4.2.3}$$

Dann gilt mit $g(t, x) = \frac{t}{t^2 + x^2}$ und $h(t, x) = -\frac{x}{t^2 + x^2}$ die Bedingung (4.2.2), weil

$$\frac{\partial g}{\partial t}(t, x) = \frac{t^2 + x^2 - t(2t)}{(t^2 + x^2)^2} = \frac{x^2 - t^2}{(t^2 + x^2)^2} = -\frac{t^2 + x^2 - x(2x)}{(t^2 + x^2)^2} = \frac{\partial h}{\partial x}(t, x).$$

Die Differentialgleichung (4.2.3) ist aber nicht exakt: Sei $S : \tilde{D} \rightarrow \mathbb{R}$ eine Stammfunktion, also

$$\frac{\partial S}{\partial t}(t, x) = -\frac{x}{t^2 + x^2} \quad \text{und} \quad \frac{\partial S}{\partial x}(t, x) = \frac{t}{t^2 + x^2}.$$

Die Funktion $P(\sigma) := S(\cos \sigma, \sin \sigma)$ ist natürlich periodisch, aber für $\sigma > 0$

$$\frac{d}{d\sigma} P(\sigma) = -\frac{\partial S}{\partial t}(\cos \sigma, \sin \sigma) \sin \sigma + \frac{\partial S}{\partial x}(\cos \sigma, \sin \sigma) \cos \sigma = \frac{\sin^2 \sigma}{\cos^2 \sigma + \sin^2 \sigma} + \frac{\cos^2 \sigma}{\cos^2 \sigma + \sin^2 \sigma} = 1.$$

Widerspruch zur Periodizität!

Der Grund für diese Lücke zwischen Exaktheit und der Integrabilität liegt in dem "Loch", auf dem g und h nicht definiert sind. Andernfalls ist die Integrabilität hinreichend für die Exaktheit. Wir zeigen dies für Rechtecksgebiete.

Proposition 4.2.6 *Betrachte die Differentialgleichung*

$$\dot{x}g(t, x) + h(t, x) = 0$$

auf $\tilde{D} = (a, b) \times (c, d) \subset \mathbb{R}^2$ mit $-\infty \leq a < b \leq \infty$ und $-\infty \leq c < d \leq \infty$. Die Funktionen g und h seien hier stetig differenzierbar. Dann ist die Integrabilitätsbedingung (4.2.2) notwendig und hinreichend für die Exaktheit.

Beweis. Gelte (4.2.2). Wir zeigen, dass für $(\tau, \xi) \in \tilde{D}$

$$S(t, x) := \int_{\tau}^t h(s, x) ds + \int_{\xi}^x g(\tau, y) dy, \quad (t, x) \in \tilde{D}, \quad (4.2.4)$$

eine Stammfunktion ist. In der Tat: Es gilt $\frac{\partial S}{\partial t}(t, x) = h(t, x)$ und wegen (4.2.2) (Differentiation unter dem Integral ist hier erlaubt, weil h und g C^1 -Funktionen sind)

$$\frac{\partial S}{\partial x}(t, x) = \int_{\tau}^t \frac{\partial h}{\partial x}(s, x) ds + g(\tau, x) = \int_{\tau}^t \frac{\partial g}{\partial t}(s, x) ds + g(\tau, x) = g(t, x) - g(\tau, x) + g(\tau, x) = g(t, x).$$

■

Die Methode der Trennung der Variablen kann als Spezialfall der obigen Konstruktion interpretiert werden.

Bemerkung 4.2.7 *Betrachte*

$$\dot{x} = p(t)q(x) \quad (4.2.5)$$

mit $p : D_1 \rightarrow \mathbb{R}$ stetig und $q : D_2 \rightarrow \mathbb{R}$ Lipschitz-stetig und $D_1, D_2 \subset \mathbb{R}$, offen. Jede Nullstelle von q liefert eine konstante Lösung. Also beschränken wir uns auf die offene Menge

$$\tilde{D} = \{(t, x) \in D_1 \times D_2 \mid q(x) \neq 0\}.$$

Hier ist (4.2.5) äquivalent zu der impliziten Gleichung

$$\dot{x} \frac{1}{q(x)} - p(t) = 0.$$

Offenbar ist hier der eingeschränkte Definitionsbereich $D = \{(t, x) \in \tilde{D} \mid \frac{1}{q(x)} \neq 0\} = \tilde{D}$. Auf jedem Rechteck $(a, b) \times (c, d) \subset D$ ist mit $g(t, x) = \frac{1}{q(x)}$ und $h(t, x) = -p(t)$ die Integrabilitätsbedingung

$$\frac{\partial g}{\partial t}(t, x) = \frac{\partial h}{\partial x}(t, x)$$

hier trivialerweise erfüllt. Wir können also die Stammfunktion (auf einem Rechteck, das (τ, ξ) enthält) schreiben als

$$S(t, x) = \int_{\tau}^t h(s, x) ds + \int_{\xi}^x g(\tau, y) dy = - \int_{\tau}^t p(s) ds + \int_{\xi}^x \frac{1}{q(y)} dy.$$

Dann lässt sich $S(t, x) = S(\tau, \xi)$ in der Form

$$\int_{\xi}^x \frac{1}{q(y)} dy = \int_{\tau}^t p(s) ds$$

schreiben. Berechnen der Integrale und Auflösen nach x liefert dann die allgemeine Lösung.

4.3 Integrierende Faktoren

Wir beginnen mit einem Beispiel. Es zeigt, dass eine explizite Differentialgleichung zu verschiedenen impliziten Differentialgleichungen äquivalent sein kann, von denen einige exakt, andere nicht exakt sein können.

Beispiel 4.3.1 Die Gleichung

$$\dot{x} = tx$$

kann auf verschiedene Weise als implizite Gleichung der Form $\dot{x}g(t, x) - h(t, x) = 0$ geschrieben werden, etwa

$$\begin{aligned} \dot{x} - tx &= 0 \text{ auf } D_a := \mathbb{R}^2; \\ \dot{x} \frac{1}{x} - t &= 0 \text{ auf } D_b := \{(t, x) \in \mathbb{R}^2 \mid x \neq 0\}; \\ \dot{x} \frac{1}{t} - x &= 0 \text{ auf } D_c := \{(t, x) \in \mathbb{R}^2 \mid t \neq 0\}; \\ \dot{x} \frac{1}{tx} - 1 &= 0 \text{ auf } D_d := \{(t, x) \in \mathbb{R}^2 \mid tx \neq 0\}. \end{aligned}$$

Auf jedem Quadranten in \mathbb{R}^2 sind die Integrabilitätsbedingungen ($\frac{\partial g}{\partial t} = \frac{\partial h}{\partial x}$) in den vier Fällen jeweils:

$$0 \equiv t; \quad 0 \equiv 0; \quad -\frac{1}{t^2} \equiv -1; \quad -\frac{1}{t^2 x} \equiv 0.$$

Offenbar ist nur die zweite Form exakt.

Wir können also durch geschickte Umformulierung (Multiplikation mit einer geeigneten Funktion) manchmal die Exaktheit erreichen; dabei sollte der Definitionsbereich erhalten bleiben.

Definition 4.3.2 Gegeben sei

$$\dot{x}g(t, x) + h(t, x) = 0$$

mit $g, h : \tilde{D} \rightarrow \mathbb{R}$ stetig und Lipschitz-stetig bezüglich x , $\tilde{D} \subset \mathbb{R}^2$ offen. Dann heißt eine stetige und bezüglich x Lipschitz-stetige Funktion $m : \tilde{D} \rightarrow \mathbb{R}$ integrierender Faktor (oder Eulerscher Multiplikator), falls $m(t, x) \neq 0$ auf \tilde{D} und die Differentialgleichung

$$\dot{x}g(t, x)m(t, x) + h(t, x)m(t, x) = 0$$

auf \tilde{D} exakt ist.

Offenbar sind die beiden Differentialgleichungen äquivalent, haben also dieselben Lösungen. Als Anwendung betrachten wir die **Bernoullischen Differentialgleichungen**.

Beispiel 4.3.3

$$\dot{x} = a(t)x + b(t)x^\alpha$$

mit $a, b : D_1 \rightarrow \mathbb{R}$ stetig und $D_1 \subset \mathbb{R}$ offen. Sei

$$D = D_1 \times (0, \infty) \subset \mathbb{R}^2.$$

Dies können wir als Vereinigung von Rechtecken darstellen, und wir nehmen an, dass

$$a(t) \not\equiv 0, \quad b(t) \not\equiv 0, \quad \alpha \neq 1.$$

Die Gleichung

$$\dot{x} - a(t)x - b(t)x^\alpha = 0,$$

also

$$g(t, x) \equiv 1 \text{ und } h(t, x) = -a(t)x - b(t)x^\alpha,$$

ist nicht exakt wegen

$$\frac{\partial g}{\partial t}(t, x) = 0 \text{ und } \frac{\partial h}{\partial x}(t, x) = -a(t) - \alpha b(t)x^{\alpha-1}.$$

Wir erhalten einen integrierenden Faktor folgendermaßen: Für eine Stammfunktion $A : D_1 \rightarrow \mathbb{R}$ von a definiere

$$m(t, x) = (1 - \alpha)x^{-\alpha}e^{(\alpha-1)A(t)}.$$

Mit

$$\hat{g}(t, x) = (1 - \alpha)x^{-\alpha}e^{(\alpha-1)A(t)}, \quad \hat{h}(t, x) = [a(t)x + b(t)x^\alpha](\alpha - 1)x^{-\alpha}e^{(\alpha-1)A(t)}$$

ist

$$\dot{x}\hat{g}(t, x) + \hat{h}(t, x) = 0$$

exakt, denn auf jedem Rechteck ist die Integrabilitätsbedingung erfüllt:

$$\begin{aligned} \frac{\partial \hat{h}}{\partial x}(t, x) &= \frac{\partial}{\partial x} [a(t)(\alpha - 1)x^{-\alpha+1} - b(t)(\alpha - 1)] e^{(\alpha-1)A(t)} = (\alpha - 1)e^{(\alpha-1)A(t)} (1 - \alpha)a(t)x^{-\alpha} \\ &= \frac{\partial \hat{g}}{\partial t}(t, x). \end{aligned}$$

Nach der Formel (4.2.4) berechnen wir jetzt eine Stammfunktion. Für $(\tau, \xi) \in D$ gilt

$$\begin{aligned} 0 &= S(t, x) - S(\tau, \xi) \\ &= \int_{\tau}^t \hat{h}(s, x) ds + \int_{\xi}^x \hat{g}(\tau, y) dy \\ &= \int_{\tau}^t [a(s)x + b(s)x^\alpha](\alpha - 1)x^{-\alpha}e^{(\alpha-1)A(s)} ds + \int_{\xi}^x (1 - \alpha)y^{-\alpha}e^{(\alpha-1)A(\tau)} dy \\ &= x^{1-\alpha} \int_{\tau}^t (\alpha - 1)a(s)e^{(\alpha-1)A(s)} ds + (\alpha - 1) \int_{\tau}^t b(s)e^{(\alpha-1)A(s)} ds + (1 - \alpha)e^{(\alpha-1)A(\tau)} \int_{\xi}^x y^{-\alpha} dy \\ &= x^{1-\alpha} \int_{\tau}^t \frac{d}{ds} [e^{(\alpha-1)A(s)}] ds + (\alpha - 1) \int_{\tau}^t b(s)e^{(\alpha-1)A(s)} ds + e^{(\alpha-1)A(\tau)} [x^{1-\alpha} - \xi^{1-\alpha}] \\ &= x^{1-\alpha}e^{(\alpha-1)A(t)} - x^{1-\alpha}e^{(\alpha-1)A(\tau)} + (\alpha - 1) \int_{\tau}^t b(s)e^{(\alpha-1)A(s)} ds + e^{(\alpha-1)A(\tau)}x^{1-\alpha} - e^{(\alpha-1)A(\tau)}\xi^{1-\alpha} \\ &= x^{1-\alpha}e^{(\alpha-1)A(t)} + (\alpha - 1) \int_{\tau}^t b(s)e^{(\alpha-1)A(s)} ds - e^{(\alpha-1)A(\tau)}\xi^{1-\alpha}. \end{aligned}$$

Auflösen nach x ist hier möglich und liefert die allgemeine Lösung

$$\lambda(t; \tau, \xi) = e^{A(t)} \left[e^{(\alpha-1)A(\tau)}\xi^{1-\alpha} + (1 - \alpha) \int_{\tau}^t b(s)e^{(\alpha-1)A(s)} ds \right]^{\frac{1}{1-\alpha}}.$$

Zum Beispiel erhält man für

$$\dot{x} = x + \frac{\sin t}{x}$$

(also $\alpha = -1, a(t) = 1, b(t) = \sin t, t \in \mathbb{R}$) die Lösung

$$\lambda(t; \tau, \xi) = \sqrt{e^{2(t-\tau)} \left[\frac{2}{5} \cos \tau + \frac{4}{5} \sin \tau + \xi^2 \right] - \frac{2}{5} \cos t - \frac{4}{5} \sin t}$$

auf $D = \mathbb{R} \times (0, \infty)$. Diese Differentialgleichung kann man offenbar auch für $\xi < 0$ betrachten, man erhält

$$\lambda(t; \tau, -\xi) = \lambda(t; \tau, \xi) \text{ für } \xi > 0.$$

Bemerkung 4.3.4 *Es gibt keine allgemeine Methode, integrierende Faktoren zu bestimmen. Einen Anhaltspunkt liefert aber die folgende Überlegung. Auf einem Rechteck ist $m(t, x)$ genau dann ein integrierender Faktor, wenn*

$$\dot{x}g(t, x)m(t, x) + h(t, x)m(t, x) = 0$$

die Integrabilitätsbedingung erfüllt. Also

$$\frac{\partial}{\partial t} [g(t, x)m(t, x)] = \frac{\partial}{\partial x} [h(t, x)m(t, x)]$$

oder, mit der Produktregel

$$\frac{\partial g}{\partial t}(t, x)m(t, x) + g(t, x)\frac{\partial m}{\partial t}(t, x) = \frac{\partial h}{\partial x}(t, x)m(t, x) + h(t, x)\frac{\partial m}{\partial x}(t, x),$$

also

$$g(t, x)\frac{\partial m}{\partial t}(t, x) - h(t, x)\frac{\partial m}{\partial x}(t, x) = m(t, x) \left[\frac{\partial h}{\partial x}(t, x) - \frac{\partial g}{\partial t}(t, x) \right].$$

Diese partielle Differentialgleichung ist nur in Spezialfällen lösbar. Vgl. Aulbach [4], pp. 158.

Bemerkung 4.3.5 *Ausdrücke der Form*

$$g(t, x)dx + h(t, x)dt = 0 \tag{4.3.1}$$

heißen in der Literatur auch Pfaffsche Differentialgleichungen. Es sind keine Differentialgleichungen im Sinn dieser Vorlesung; zu ihrer präzisen Definition benötigt man die Theorie der Differentialformen vom Grad eins (oder Pfaffschen Formen). Vgl. hierzu zum Beispiel H. Amann, J. Escher, *Analysis II, Grundstudium Mathematik, Birkhäuser 2006 [3], Kapitel VIII.3*. Wir beschränken uns hier auf die folgenden Überlegungen: Eine Lösung von (4.3.1) ist eine stetig differenzierbare Abbildung $\psi = (\psi_1, \psi_2)$ eines Intervalls I nach \mathbb{R}^2 mit $\psi'(\tau) \neq (0, 0)$ für alle $\tau \in I$, so dass

$$g(\psi(\tau))\psi'_1(\tau) + h(\psi(\tau))\psi'_2(\tau) = 0 \text{ für alle } \tau \in I.$$

Die Gleichung (4.3.1) können wir auch als abkürzende Schreibweise ansehen für die beiden Differentialgleichungen

$$g(t, x)\frac{dx}{dt} + h(t, x) = 0 \text{ und } g(t, x) + h(t, x)\frac{dt}{dx} = 0. \tag{4.3.2}$$

Die Lösungen dieser Differentialgleichungen sind dann, nach Reparametrisierung, Lösungen ψ von (4.3.1). Die Stammfunktionen der beiden Gleichungen in (4.3.2) stimmen überein, denn es muss jeweils gelten

$$\frac{\partial S}{\partial t}(t, x) = h(t, x) \text{ und } \frac{\partial S}{\partial x}(t, x) = g(t, x).$$

Die Niveaulinien der gemeinsamen Stammfunktion stimmen also überein und ebenfalls ihre Lösungen, eingeschränkt auf den Durchschnitt der eingeschränkten Lösungsbereiche. Wir können (4.3.1) also als verallgemeinerte Differentialgleichung auffassen, die die beiden Differentialgleichungen (4.3.2) umfasst. In der (t, x) -Ebene sind dabei in den Lösungen auch senkrechte Tangenten erlaubt.

Bemerkung 4.3.6 *Im Beispiel (4.2.3) impliziert Integrabilität nicht Exaktheit. Hier hat der betrachtete Definitionsbereich $\tilde{D} = \mathbb{R}^2 \setminus \{(0, 0)\}$ ein "Loch". Ist dies ausgeschlossen (also \tilde{D} einfach zusammenhängend), so gilt diese Implikation. Man kann diese Eigenschaften auch nutzen, um die Zahl der Löcher im Definitionsbereich zu zählen. Beziehungen dieser Art werden in der algebraischen Topologie (mit Hilfe der Theorie der Differentialformen) ausgenutzt, um die de Rham-Gruppen zu definieren, die Mannigfaltigkeiten klassifizieren (vgl. zum Beispiel W.M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry, Academic Press 1986, Section VI.7, [5]*).*

4.4 Weitere Lösungsmethoden

Man kann Transformationen betrachten, um Differentialgleichungen auf bekannte Differentialgleichungen oder allgemeiner auf solche einfacherer Form zurückzuführen. Abstrakt kann man dies folgendermaßen formulieren. Sei

$$\frac{dx}{dt} = f(t, x), \quad f : D \rightarrow \mathbb{R}^n, \quad (4.4.1)$$

mit $D \subset \mathbb{R} \times \mathbb{R}^n$ offen. Betrachte eine bijektive Abbildung

$$T : D \rightarrow \mathbb{R} \times \mathbb{R}^n, \quad (t, x) \mapsto (s, y),$$

die die Lösungen von (4.4.1) auf die Lösungen einer (bekannten) Differentialgleichung

$$\frac{dy}{ds} = g(s, y), \quad g : T(D) \rightarrow \mathbb{R}^n \quad (4.4.2)$$

abbildet. Dann kann man mit T^{-1} auf die Lösungen von (4.4.1) zurückschließen.

Beispiel 4.4.1 Betrachte

$$\dot{x} = \frac{x}{t} + \sqrt{1 - \frac{x^2}{t^2}}. \quad (4.4.3)$$

Für die rechte Seite $f(t, x) := \frac{x}{t} + \sqrt{1 - \frac{x^2}{t^2}}$, $f : D = \{(t, x) \in \mathbb{R}^2 \mid x^2 < t^2\} \rightarrow \mathbb{R}$, dieser Gleichung gilt

$$f(\sigma t, \sigma x) = f(t, x) \text{ für alle } (t, x) \in D, \sigma > 0.$$

(man sagt auch, sie ist homogen vom Grad null). Wir wollen für einen Punkt $(t_0, x_0) \in D$, $t_0 > 0$, die maximale Lösung berechnen und verwenden die Transformation

$$y = \frac{x}{t}, \text{ also } T(t, x) = (t, \frac{x}{t}).$$

Ist $x(\cdot)$ eine Lösung von (4.4.3), so definiere $y(t) := \frac{x(t)}{t}$ und verwende die Homogenität, um zu zeigen

$$\begin{aligned} \dot{y}(t) &= \frac{d}{dt} y(t) = \frac{d}{dt} \frac{x(t)}{t} = \frac{\dot{x}(t)t - x(t)}{t^2} = \frac{tf(t, x(t)) - ty(t)}{t^2} = \frac{f(t, ty(t)) - y(t)}{t} = \frac{f(1, y(t)) - y(t)}{t} \\ &= \frac{y(t) + \sqrt{1 - y^2} - y(t)}{t}. \end{aligned}$$

Also ist $y(t)$ eine Lösung von

$$\dot{y} = \frac{\sqrt{1 - y^2}}{t}. \quad (4.4.4)$$

Der Definitionsbereich der rechten Seite ist

$$D' = \left\{ (t, y) \in \mathbb{R}^2 \mid t \neq 0, y \in \left\{ \frac{x}{t} \in \mathbb{R}, x^2 < t^2 \right\} \right\} = \{(t, y) \in \mathbb{R}^2 \mid t \neq 0, |y| < 1\}.$$

Die Gleichung (4.4.4) kann man mit Trennung der Variablen lösen und erhält für $y_0 := \frac{x_0}{t_0}$

$$\arcsin y - \arcsin y_0 = \int_{y_0}^y \frac{dw}{\sqrt{1 - w^2}} = \int_{t_0}^t \frac{ds}{s} = \ln t - \ln t_0,$$

also

$$y(t) = \sin\left(\arcsin y_0 + \ln \frac{t}{t_0}\right)$$

mit dem maximalen Existenzintervall (es muss $|y(t)| < 1$ gelten, also $\arcsin y_0 + \ln \frac{t}{t_0} \in (-\frac{\pi}{2}, \frac{\pi}{2})$)

$$I = (t_0 e^{-\frac{\pi}{2} - \arcsin y_0}, t_0 e^{\frac{\pi}{2} - \arcsin y_0}).$$

Rücktransformation liefert, mit $x_0 = ty_0$,

$$x(t) = ty(t) = t \sin \left(\arcsin(tx_0) + \ln \frac{t}{t_0} \right), \quad t \in I,$$

als Lösung zum Anfangswert $x(t_0) = x_0$.

Aulbach [4], Kapitel 4.3, diskutiert eine Reihe von weiteren Beispielen nützlicher Transformationen. Wir werden auf Transformationen zurückkommen, wenn wir die lokale Struktur von Flüssen (also autonome Differentialgleichungen) studieren.

Eine weitere Lösungsmethode (auch als Methode von Frobenius bekannt) beruht auf Reihendarstellungen von Lösungen. Wir illustrieren dies wieder an Hand eines Beispiels, wobei wir auf eine rigorose Herleitung verzichten. Betrachte die Differentialgleichung

$$t^2 \ddot{x} + t\dot{x} + (t^2 - \frac{1}{4})x = 0 \quad (4.4.5)$$

Die Idee besteht darin, eine Potenzreihenentwicklung der Lösung in der Form

$$x(t) = t^c \sum_{n=0}^{\infty} a_n t^n \quad (4.4.6)$$

anzunehmen, mit Konstanten $c, a_0, a_1, \dots \in \mathbb{R}$. Gliedweises Differenzieren liefert

$$\dot{x}(t) = \frac{d}{dt} \sum_{n=0}^{\infty} a_n t^{n+c} = \sum_{n=0}^{\infty} a_n (n+c) t^{n+c-1}, \quad \ddot{x}(t) = \sum_{n=0}^{\infty} a_n (n+c)(n+c-1) t^{n+c-2}.$$

Einsetzen in (4.4.5) und Umsortieren ergibt

$$\begin{aligned} 0 &= t^2 \sum_{n=0}^{\infty} a_n (n+c)(n+c-1) t^{n+c-2} + t \sum_{n=0}^{\infty} a_n (n+c) t^{n+c-1} + (t^2 - \frac{1}{4}) \sum_{n=0}^{\infty} a_n t^{n+c} \\ &= \sum_{n=0}^{\infty} a_n \left[(n+c)(n+c-1) + (n+c) - \frac{1}{4} \right] t^{n+c} + \sum_{n=0}^{\infty} a_n t^{n+c+2} \\ &= \sum_{n=0}^{\infty} a_n \left[(n+c)^2 - \frac{1}{4} \right] t^{n+c} + \sum_{n=2}^{\infty} a_{n-2} t^{n+c}. \end{aligned}$$

Da dies für alle t gelten soll, müssen die Koeffizienten alle gleich null sein. Wir erhalten für $n = 0$

$$a_0 \left(c^2 - \frac{1}{4} \right) = 0, \quad (4.4.7)$$

für $n = 1$

$$a_1 \left[(c+1)^2 - \frac{1}{4} \right] = 0, \quad (4.4.8)$$

und für $n \geq 2$

$$a_n \left[(n+c)^2 - \frac{1}{4} \right] + a_{n-2} = 0. \quad (4.4.9)$$

Nehmen wir (o.B.d.A.) an, dass $a_0 \neq 0$, so impliziert (4.4.7), dass $c = \pm \frac{1}{2}$.

Die Wahl $c = \frac{1}{2}$ in (4.4.8) liefert $a_1 = 0$. Da es sich um eine Gleichung zweiter Ordnung handelt, wird die Lösung von zwei Parametern abhängen. Wählen wir $c = -\frac{1}{2}$, so ist a_1 (neben a_0) ein

Parameter, dies erscheint also vernünftig. Mit dieser Wahl von c ergibt sich aus (4.4.9) für $n \geq 2$ die Rekursion

$$0 = a_n \left[\left(n - \frac{1}{2}\right)^2 - \frac{1}{4} \right] + a_{n-2} = a_n n(n-1) + a_{n-2}, \text{ also } a_n = -\frac{a_{n-2}}{n(n-1)}.$$

Induktion liefert

$$a_{2n} = (-1)^n \frac{a_0}{(2n)!}, \quad a_{2n+1} = (-1)^n \frac{a_1}{(2n+1)!}.$$

Einsetzen in die Reihe (4.4.6) für $x(t)$ ergibt dann

$$\begin{aligned} x(t) &= t^c \sum_{n=0}^{\infty} a_n t^n = t^{-1/2} \left[a_0 \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{(2n)!} + a_1 \sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} \right] \\ &= a_0 \frac{\cos t}{\sqrt{t}} + a_1 \frac{\sin t}{\sqrt{t}}. \end{aligned}$$

Für alle $a_0, a_1 \in \mathbb{R}$ ist dies, wie man leicht nachrechnet, eine Lösung von (4.4.5).

Will man dieses Vorgehen rechtfertigen und insbesondere auch das maximale Existenzintervall von Lösungen bestimmen, braucht man weitere Eigenschaften von Potenzreihen. Ferner ist es naheliegend, auch komplexe Lösungen zuzulassen (einige Informationen dazu findet man auch bei W. Walter [15]).

Kapitel 5

Ebene Autonome Systeme

In diesem Kapitel diskutieren wir autonome Differentialgleichungen in \mathbb{R}^2 . Zunächst zeigen wir, wie sie auf skalare, nicht-autonome Differentialgleichungen reduziert werden können. Dann diskutieren wir Hamiltonsche Systeme und erste Integrale.

5.1 Reduktion auf skalare Differentialgleichungen

Betrachte die autonome Differentialgleichung in \mathbb{R}^2

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y) \quad (5.1.1)$$

mit $f, g : D \rightarrow \mathbb{R}$ Lipschitz-stetig und $D \subset \mathbb{R}^2$ offen. Die Idee zur Analyse besteht darin, die Abhängigkeit von y bezüglich x zu studieren, wobei x als eine verallgemeinerte Zeit aufgefasst wird (oder analog x als Funktion von y). Das wird in der folgenden Proposition präzisiert.

Proposition 5.1.1 *Betrachte für (5.1.1) einen Punkt $(x_0, y_0) \in D$.*

(i) *Gilt $f(x, y) \neq 0$ auf D , so ist die Trajektorie $\mathcal{O}(x_0, y_0)$ gleich der maximalen Lösungskurve für*

$$\frac{dy}{dx} = \frac{g(x, y)}{f(x, y)}, \quad y(x_0) = y_0. \quad (5.1.2)$$

(ii) *Gilt $g(x, y) \neq 0$ auf D , so ist die Trajektorie $\mathcal{O}(x_0, y_0)$ gleich der maximalen Lösungskurve für*

$$\frac{dx}{dy} = \frac{f(x, y)}{g(x, y)}, \quad x(y_0) = x_0. \quad (5.1.3)$$

Beweis. (i) Sei

$$(\varphi_1(t), \varphi_2(t)) := \varphi(t; x_0, y_0), \quad t \in J := J_{\max}(x_0, y_0),$$

Lösung von (5.1.1). Wegen $f(x, y) \neq 0$ auf D ist $\varphi_1(t), t \in J$, streng monoton, hat also die Umkehrfunktion

$$\varphi_1^{-1} : \varphi_1(J) \rightarrow J \text{ mit } \varphi_1^{-1}(x_0) = 0.$$

Mit $t = \varphi_1^{-1}(x)$ erhält man für die Trajektorie

$$\begin{aligned} \mathcal{O}(x_0, y_0) &= \{(\varphi_1(t), \varphi_2(t)) \in \mathbb{R}^2 \mid t \in J\} \\ &= \{(\varphi_1(\varphi_1^{-1}(x)), \varphi_2(\varphi_1^{-1}(x))) \in \mathbb{R}^2 \mid x \in \varphi_1(J)\} \\ &= \{(x, \varphi_2(\varphi_1^{-1}(x))) \in \mathbb{R}^2 \mid x \in \varphi_1(J)\}. \end{aligned}$$

Ferner ist $y(x) := \varphi_2(\varphi_1^{-1}(x))$ die maximale Lösung von (5.1.2), weil $y(x_0) := \varphi_2(\varphi_1^{-1}(x_0)) = \varphi_2(0) = y_0$, und

$$\frac{dy}{dx}(x) = \frac{\dot{\varphi}_2(\varphi_1^{-1}(x))}{\dot{\varphi}_1(\varphi_1^{-1}(x))} = \frac{g(\varphi_1(\varphi_1^{-1}(x)), \varphi_2(\varphi_1^{-1}(x)))}{f(\varphi_1(\varphi_1^{-1}(x)), \varphi_2(\varphi_1^{-1}(x)))} = \frac{g(x, y(x))}{f(x, y(x))}.$$

(ii) folgt analog. ■

Bemerkung 5.1.2 Die obige Diskussion zeigt, dass die beiden Differentialgleichungen

$$\frac{dy}{dx}f(x, y) - g(x, y) = 0 \text{ und } -\frac{dx}{dy}g(x, y) + f(x, y) = 0 \quad (5.1.4)$$

für die Analyse benötigt werden. Dies entspricht den "Pfaffschen Differentialgleichungen" aus Bemerkung 4.3.5. Die Lösungen einer Pfaffschen Differentialgleichung sind Kurven in der Ebene, die eine C^1 -Parametrisierung (wie die Trajektorien von (5.1.1) durch t) haben; dabei setzt man noch voraus, dass für diese Parametrisierung $\psi : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ gilt, dass $\psi'(t) \neq (0, 0)$ für alle t .

5.2 Hamiltonsche Differentialgleichungen und erste Integrale

Im vorangehenden Kapitel haben wir Exaktheit von Gleichungen der Form (5.1.2), (5.1.3) zu ihrer Lösung benutzt. Hier kann man dies auch direkt an dem zwei-dimensionalen autonomen System (5.1.1) ablesen. Dafür ist das folgende Konzept, das wir nicht nur für skalare Systeme formulieren, nützlich.

Definition 5.2.3 Es sei $D \subset \mathbb{R}^{2n}$ offen und $H : D \rightarrow \mathbb{R}$ eine C^1 -Funktion. Dann heißt eine Differentialgleichung der Form

$$\dot{x} = \frac{\partial H}{\partial y}(x, y), \dot{y} = -\frac{\partial H}{\partial x}(x, y) \quad (5.2.5)$$

eine Hamiltonsche Differentialgleichung.

Man beachte, dass in (5.2.5) die rechte Seite als Spaltenvektor betrachtet wird. Die Differentialgleichung (5.1.1) in der Ebene \mathbb{R}^2 ist also eine Hamiltonsche Gleichung, falls es eine C^1 -Funktion $H : D \rightarrow \mathbb{R}$ gibt mit

$$\frac{\partial H}{\partial x}(x, y) = -g(x, y) \text{ und } \frac{\partial H}{\partial y}(x, y) = f(x, y) \text{ auf } D. \quad (5.2.6)$$

Dann ist offenbar H eine Stammfunktion für beide Differentialgleichungen in (5.1.4).

Proposition 5.2.4 Für eine Hamiltonsche Differentialgleichung (5.2.5) mit Hamiltonfunktion H gilt

$$H(\varphi(t; x_0, y_0)) = H(x_0, y_0) \text{ für alle } (x_0, y_0) \in D \text{ und } t \in J_{\max}(x_0, y_0).$$

Die Niveaumengen $\{(x, y) \in \mathbb{R}^{2n} \mid H(x, y) = c\}$ mit $c \in \mathbb{R}$ sind also Vereinigung von Trajektorien.

Beweis. Wir schreiben für die x und y -Komponenten der Lösung von (5.2.5) auch $(\varphi_1(t), \varphi_2(t)) := \varphi(t; x_0, y_0)$. Mit der Kettenregel folgt für alle t

$$\begin{aligned} \frac{d}{dt}H(\varphi_1(t), \varphi_2(t)) &= \frac{\partial H}{\partial x}(\varphi_1(t), \varphi_2(t))^T \cdot \dot{\varphi}_1(t) + \frac{\partial H}{\partial y}(\varphi_1(t), \varphi_2(t))^T \cdot \dot{\varphi}_2(t) \\ &= \frac{\partial H}{\partial x}(\varphi_1(t), \varphi_2(t))^T \cdot \frac{\partial H}{\partial y}(\varphi_1(t), \varphi_2(t)) - \frac{\partial H}{\partial y}(\varphi_1(t), \varphi_2(t))^T \cdot \frac{\partial H}{\partial x}(\varphi_1(t), \varphi_2(t)) = 0. \end{aligned}$$

■

Im einleitenden Kapitel haben wir mechanische Systeme diskutiert, die durch Differentialgleichungen der Form

$$\ddot{x} = F(x)$$

beschrieben werden. Sie sind äquivalent zu dem System erster Ordnung

$$\dot{x} = y, \quad \dot{y} = F(x)$$

in der Ebene. Weil Reibungsverluste unberücksichtigt bleiben, ist die Gesamtenergie konstant. Sie ist die Summe aus der kinetischen Energie und der potentiellen Energie

$$\frac{1}{2}(y^2 - y_0^2) \text{ bzw. } - \int_{x_0}^x F(w)dw;$$

hierbei sind x_0 und y_0 Normierungsgrößen. Es ist dann

$$H(x, y) = \frac{1}{2}(y^2 - y_0^2) - \int_{x_0}^x F(w)dw \quad (5.2.7)$$

die zugehörige Hamilton-Funktion mit $H(x_0, y_0) = 0$.

Beispiel 5.2.5 Das einfachste Beispiel solch einer (nichtlinearen) Hamiltonsche Differentialgleichung ist das reibungslose Pendel

$$\dot{x} = y, \quad \dot{y} = -\sin x \quad (5.2.8)$$

mit Hamiltonfunktion

$$H(x, y) = \frac{1}{2}(y^2 - y_0^2) - \cos x + \cos x_0.$$

Die Niveaumengen sind

$$N(x_0, y_0) = \{(x, y) \in \mathbb{R}^2 \mid y^2 - 2 \cos x = y_0^2 - 2 \cos x_0\}$$

Die Trajektorien sind nach Proposition 5.2.4 Teilmengen von Niveaumengen. Für $c \in \mathbb{R}$ sind die Niveaumengen

$$N_c = \{(x, y) \in \mathbb{R}^2 \mid y^2 - 2 \cos x = c\} = \{(x, \pm \sqrt{2 \cos x + c}) \in \mathbb{R}^2 \mid 2 \cos x + c \geq 0\}.$$

Offenbar ist $N_c = \emptyset$ für $c < -2$. Im Intervall $x \in [-\pi, \pi]$ erhält man (für $c = -2$) $N_{-2} = \{(0, 0)\}$; für $c \in (-2, 2)$ ergeben sich geschlossene Kurven um $(0, 0)$. Dies sind periodische Lösungen von (5.2.8). Für $c = 2$ erhält man zwei Kurven, die die Punkte $(-\pi, 0)$ und $(\pi, 0)$ verbinden. Diese beiden Punkte sind Gleichgewichte der Differentialgleichung, sie sind verbunden durch Trajektorien, die für $t \rightarrow \pm\infty$ gegen die Gleichgewichte gehen.

Beispiel 5.2.6 Als weiteres Beispiel betrachten wir

$$\ddot{x} = x - x^2,$$

also

$$\dot{x} = y, \quad \dot{y} = F(x) = x - x^2. \quad (5.2.9)$$

Die zugehörigen skalaren nichtautonomen Differentialgleichungen sind

$$(x^2 - x) \frac{dx}{dy} + y = 0 \text{ und } x^2 - x + y \frac{dy}{dx} = 0.$$

Man erhält aus (5.2.7) (mit $x_0 = 0, y_0 = 0$)

$$H(x, y) = \frac{1}{3}x^3 - \frac{1}{2}x^2 + \frac{1}{2}y^2.$$

Die Niveaumengen

$$\frac{c}{2} = H(x, y) = \frac{1}{3}x^3 - \frac{1}{2}x^2 + \frac{1}{2}y^2.$$

beschreiben das Phasenporträt. Auflösen nach y liefert

$$y = \pm \sqrt{c + x^2 - \frac{2}{3}x^3}$$

für $c + x^2 - \frac{2}{3}x^3 \geq 0$. Das Gleichgewicht $(0, 0)$ liegt in der Niveaumenge mit $c = 0$ und das Gleichgewicht $(1, 0)$ liegt in der Niveaumenge mit $c = -\frac{1}{3}$. Diskussion der Graphen der Funktionen $x \mapsto c + x^2 - \frac{2}{3}x^3$ liefert das Phasenporträt.

Die Verwendung von integrierenden Faktoren können wir auch für die autonomen Systeme in der Ebene ausnutzen, wenn also kein Hamiltonsches System vorliegt. Man beachte, dass sich das Phasenporträt des autonomen Systems (5.1.1) in der Ebene

$$\dot{x} = f(x, y), \quad \dot{y} = g(x, y)$$

nicht ändert, wenn man stattdessen

$$\dot{x} = f(x, y)m(x, y), \quad \dot{y} = g(x, y)m(x, y) \quad (5.2.10)$$

mit $m(x, y) \neq 0$ auf D betrachtet. An Stelle der beiden Gleichungen

$$f(x, y) \frac{dy}{dx} - g(x, y) = 0 \quad \text{und} \quad g(x, y) \frac{dx}{dy} - f(x, y) = 0$$

hat man dann die mit $m(x, y)$ multiplizierten Gleichungen

$$f(x, y)m(x, y) \frac{dy}{dx} - g(x, y)m(x, y) = 0 \quad \text{und} \quad g(x, y)m(x, y) \frac{dx}{dy} - f(x, y)m(x, y) = 0.$$

Eine Hamilton-Funktion von (5.2.10) ist nicht notwendigerweise eine Hamilton-Funktion von (5.1.1). Daher führen wir den folgenden allgemeineren Begriff ein.

Definition 5.2.7 Eine C^1 -Funktion $F : D \rightarrow \mathbb{R}$ heißt **erstes Integral** von (5.1.1), falls

$$\frac{\partial F}{\partial x}(x, y) \cdot f(x, y) + \frac{\partial F}{\partial y}(x, y) \cdot g(x, y) = 0 \quad \text{auf } D.$$

Offenbar ist jede Hamilton-Funktion ein erstes Integral.

Proposition 5.2.8 (i) Sei $F : D \rightarrow \mathbb{R}$ erstes Integral für (5.1.1). Dann gilt für alle $(x_0, y_0) \in D$ und $t \in J_{\max}(x_0, y_0)$

$$F(\varphi(t; x_0, y_0)) = F(x_0, y_0).$$

(ii) Sei $m : D \rightarrow \mathbb{R}$ ein integrierender Faktor für

$$f(x, y) \frac{dy}{dx} - g(x, y) = 0$$

und $S : D \rightarrow \mathbb{R}$ eine zugehörige Stammfunktion, so ist $S(x, y)$ ein erstes Integral von (5.1.1).

Beweis. (i) Es gilt mit $(x(t), y(t)) = \varphi(t; x_0, y_0)$

$$\begin{aligned} \frac{d}{dt} F(\varphi(t; x_0, y_0)) &= \frac{\partial F}{\partial x}(\varphi_1(t), \varphi_2(t)) \cdot \dot{\varphi}_1(t) + \frac{\partial F}{\partial y}(\varphi_1(t), \varphi_2(t)) \cdot \dot{\varphi}_2(t) \\ &= \frac{\partial F}{\partial x}(\varphi_1(t), \varphi_2(t)) \cdot f(\varphi_1(t), \varphi_2(t)) + \frac{\partial F}{\partial y}(\varphi_1(t), \varphi_2(t)) \cdot g(\varphi_1(t), \varphi_2(t)) = 0. \end{aligned}$$

(ii) Nach Voraussetzung gilt auf D

$$\frac{\partial S}{\partial x}(x, y) = -g(x, y)m(x, y) \quad \text{und} \quad \frac{\partial S}{\partial y}(x, y) = f(x, y)m(x, y),$$

also

$$\frac{\partial S}{\partial x}(x, y) \cdot f(x, y) + \frac{\partial S}{\partial y}(x, y) \cdot g(x, y) = 0.$$

■

Nicht jede Differentialgleichung besitzt ein nichtkonstantes erstes Integral, wie das folgende Beispiel zeigt.

Beispiel 5.2.9 Die Gleichung

$$\dot{x} = -x, \quad \dot{y} = -y$$

hat den Fluss

$$\varphi(t; x_0, y_0) = (x_0 e^{-t}, y_0 e^{-t}).$$

Ist F ein erstes Integral, so folgt für alle $(x_0, y_0) \in \mathbb{R}^2$ und $t \in \mathbb{R}$

$$F(x_0, y_0) = F(\varphi(t; x_0, y_0)) = \lim_{t \rightarrow \infty} F(\varphi(t; x_0, y_0)) = F(0, 0).$$

Also ist F konstant.

Bemerkung 5.2.10 Die obigen Resultate kann man sogar benutzen, um für (5.1.1) die Durchlaufzeit von Trajektorienstücken zu berechnen. Also: In welcher Zeit T wird für zwei Punkte $(x_0, y_0), (x_1, y_1)$ auf einer Trajektorie von (5.1.1) dieses Trajektorienstück durchlaufen? Die Zeit T ist positiv oder negativ, je nach der Durchlaufrichtung. Wir nehmen dazu an, dass auf diesem Trajektorienstück $f(x, y) \neq 0$ gilt. Der Einfachheit halber sei $f(x, y) > 0$ und $x_0 < x_1$. Ferner sei $y(x)$ die Lösung von

$$\frac{dy}{dx} = \frac{g(x, y)}{f(x, y)}, \quad y(x_0) = y_0.$$

Nach Voraussetzung ist $y(x_1) = y_1$. Wir behaupten:

$$T = \int_{x_0}^{x_1} \frac{dx}{f(x, y(x))}.$$

Zum **Beweis** beachte, dass für die Lösung $(\varphi_1(t), \varphi_2(t)) = \varphi(t; x_0, y_0), t \in [0, T]$ gilt

$$\{(x, y(x)), x \in [x_0, x_1]\} = \{(\varphi_1(t), \varphi_2(t)), t \in [0, T]\},$$

insbesondere ist $\varphi_1(0) = x_0$ und $\varphi_1(T) = x_1$. Wegen $\dot{\varphi}_1 = f(\varphi_1, \varphi_2) > 0$, existiert die Umkehrfunktion $\varphi_1^{-1} : [x_0, x_1] \rightarrow [0, T]$, und

$$(\varphi_1(\varphi_1^{-1}(x)), \varphi_2(\varphi_1^{-1}(x))) = (x, y(x)). \tag{5.2.11}$$

Ferner ist nach dem Transformationssatz für Änderung der Integrationsvariablen und dem Inverse-Funktionen-Theorem

$$0 < T = \int_0^T 1 dt = \int_{\varphi_1(0)}^{\varphi_1(T)} 1 \cdot (\varphi_1^{-1})'(x) dx = \int_{x_0}^{x_1} \frac{1}{\dot{\varphi}_1(\varphi_1^{-1}(x))} dx.$$

Wegen der Differentialgleichung (5.1.1) und wegen (5.2.11) ist dies gleich

$$\int_{x_0}^{x_1} \frac{1}{f(\varphi_1(\varphi_1^{-1}(x)), \varphi_2(\varphi_1^{-1}(x)))} dx = \int_{x_0}^{x_1} \frac{dx}{f(x, y(x))},$$

wie behauptet.

Kapitel 6

Lineare Differentialgleichungen

In diesem Kapitel diskutieren wir zunächst lineare nichtautonome Differentialgleichungen. Dann wenden wir uns speziell den autonomen linearen Differentialgleichungen zu, wo wir wesentlich explizitere Resultate geben können.

6.1 Nichtautonome Lineare Differentialgleichungen

In diesem Abschnitt betrachten wir nichtautonome lineare Differentialgleichungen der Form

$$\dot{x} = A(t)x + g(t), \quad (6.1.1)$$

wobei $A(t)$ und $g(t)$ stetige Funktionen von t in einem Intervall I sind mit Werten in $\mathbb{R}^{n \times n}$ bzw. in \mathbb{R}^n . Wir werden sehen, dass die Lösungstheorie starke Analogien zu der Lösungstheorie von linearen Gleichungen im \mathbb{R}^n der Form

$$Ax = b$$

mit $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ aufweist. Insbesondere spielt die homogene Gleichung

$$\dot{x} = A(t)x \quad (6.1.2)$$

eine besondere Rolle. Wir werden in diesem Kapitel viele Konzepte und Resultate aus der Linearen Algebra wiederfinden.

Die lineare Unabhängigkeit wird auch für Lösungen von (6.1.2) eine große Rolle spielen. Da es sich um Funktionen handelt, müssen wir einen Vektorraum von Funktionen betrachten. Geeignet ist dafür zum Beispiel der unendlich-dimensionale Vektorraum

$$C^1(I; \mathbb{R}^n) = \{\varphi : I \rightarrow \mathbb{R}^n \text{ stetig differenzierbar}\};$$

hierbei sind die Vektorraum-Operationen jeweils punktweise definiert (zum Beispiel ist $\varphi_1 + \varphi_2$ die Funktion aus $C^1(I; \mathbb{R}^n)$, die definiert ist durch $(\varphi_1 + \varphi_2)(t) := \varphi_1(t) + \varphi_2(t), t \in I$).

Bemerkung 6.1.1 Der Vektorraum $C^1(I, \mathbb{R})$ ist unendlich-dimensional, weil es für alle $k \in \mathbb{N}$ Funktionen $\varphi_i \in C^1(I, \mathbb{R}), i = 1, \dots, k$, gibt, so dass $\{\varphi_1, \dots, \varphi_k\}$ linear unabhängig sind: Wähle k paarweise verschiedene Punkte $t_i \in I$, und definiere φ_i als eine C^1 -Funktion mit $\varphi_i(t_i) = 1$ und $\varphi_i(t_j) = 0$ für $i \neq j$ (zum Beispiel $\varphi_i(t) := \prod_{j \neq i} \frac{t - t_j}{t_i - t_j}, t \in I$). Dann folgt aus

$$\alpha_1 \varphi_1 + \dots + \alpha_k \varphi_k = 0 \text{ mit } \alpha_i \in \mathbb{R},$$

dass für jedes i und $t = t_i$

$$0 = \alpha_1 \varphi_1(t_i) + \dots + \alpha_k \varphi_k(t_i) = \alpha_i \varphi_i(t_i) = \alpha_i.$$

Für jede Inhomogenität g im Vektorraum $C(I, \mathbb{R}^n)$ der stetigen Funktionen von I nach \mathbb{R}^n bezeichnen wir mit $\lambda_g(t; t_0, x_0)$ die Lösung von (6.1.1) zur Anfangsbedingung $x(t_0) = x_0$; sie existiert nach Theorem 2.4.13 auf dem Intervall I . Insbesondere ist $\lambda_0(t; t_0, x_0)$ Lösung der homogenen Gleichung (6.1.2). Ferner sei

$$\mathcal{L}(g) = \{\varphi \in C^1(I, \mathbb{R}^n) \mid \dot{\varphi}(t) = A(t)\varphi(t) + g(t), t \in I\}$$

die Menge der Lösungen von (6.1.1).

Es gilt die folgende als *Superpositionsprinzip* bekannte einfache Eigenschaft. Sie zeigt, dass $\mathcal{L}(g)$ ein Untervektorraum von $C^1(I, \mathbb{R}^n)$ ist.

Proposition 6.1.2 Für $\varphi_1 \in \mathcal{L}(g_1), \varphi_2 \in \mathcal{L}(g_2)$ und $c_1, c_2 \in \mathbb{R}$ gilt

$$c_1\varphi_1 + c_2\varphi_2 \in \mathcal{L}(c_1g_1 + c_2g_2).$$

Beweis. Dies folgt aus

$$\begin{aligned} \frac{d}{dt}(c_1\varphi_1(t) + c_2\varphi_2(t)) &= c_1A(t)\varphi_1(t) + c_1g_1(t) + c_2A(t)\varphi_2(t) + c_2g_2(t) \\ &= A(t)[c_1\varphi_1(t) + c_2\varphi_2(t)] + c_1g_1(t) + c_2g_2(t). \end{aligned}$$

■

Die folgende Aussage ist grundlegend.

Satz 6.1.3 Für jedes $t_0 \in I$ ist

$$x_0 \mapsto \lambda_0(\cdot; t_0, x_0) : \mathbb{R}^n \rightarrow \mathcal{L}(0) \tag{6.1.3}$$

ein Vektorraum-Isomorphismus. Daher ist $\mathcal{L}(0)$ ein n -dimensionaler Unterraum von $C^1(I, \mathbb{R}^n)$ und für alle $g \in C(I, \mathbb{R}^n)$ gilt

$$\mathcal{L}(g) = \{\varphi_g^* + \varphi_0 \in C^1(I, \mathbb{R}^n) \mid \varphi_0 \in \mathcal{L}(0)\}, \tag{6.1.4}$$

wobei φ_g^* eine beliebige Lösung des inhomogenen Systems (6.1.1) ist.

Beweis. Für alle $c_1, c_2 \in \mathbb{R}, x_1, x_2 \in \mathbb{R}^n$ ist

$$\lambda_0(t; t_0, c_1x_1 + c_2x_2) = c_1\lambda_0(t; t_0, x_1) + c_2\lambda_0(t; t_0, x_2), t \in I,$$

weil dies die eindeutige Lösung von $\dot{x} = A(t)x$ mit $x(t_0) = c_1x_1 + c_2x_2$ ist. Also ist die Abbildung (6.1.3) linear; sie ist injektiv wegen $\lambda(t_0; t_0, x_0) = x_0$ und surjektiv, weil für alle $\varphi_0 \in \mathcal{L}(0)$ gilt

$$\varphi_0 = \lambda_0(\cdot, t_0, \varphi_0(t_0)).$$

Es folgt, dass (6.1.3) ein Isomorphismus ist. Die Gleichheit (6.1.4) folgt aus dem Superpositionsprinzip: Sei $\varphi_g^* \in \mathcal{L}(g)$. Dann ist $\varphi_g^* + \varphi_0 \in \mathcal{L}(g)$ für alle $\varphi_0 \in \mathcal{L}(0)$. Umgekehrt, ist $\varphi_g \in \mathcal{L}(g)$, so ist $\varphi_g^* - \varphi_g \in \mathcal{L}(0)$. ■

Eine Konsequenz dieses Satzes ist die folgende Aussage.

Proposition 6.1.4 Gegeben seien $m \leq n$ Lösungen $\varphi_1, \dots, \varphi_m$ von (6.1.2). Dann sind äquivalent:

- (i) Die Funktionen $\varphi_1, \dots, \varphi_m$ sind linear unabhängig in $C^1(I, \mathbb{R}^n)$.
- (ii) Für jedes $t_0 \in I$ sind $\varphi_1(t_0), \dots, \varphi_m(t_0)$ linear unabhängig in \mathbb{R}^n .
- (iii) Es gibt ein $t_0 \in I$, so dass $\varphi_1(t_0), \dots, \varphi_m(t_0)$ linear unabhängig in \mathbb{R}^n sind.

Beweis. Offensichtlich impliziert (ii) die Bedingung (iii) und (iii) impliziert (i). Gelte (i) und betrachte für ein $t_0 \in I$ eine Linearkombination

$$\alpha_1\varphi_1(t_0) + \dots + \alpha_m\varphi_m(t_0) = 0 \text{ mit } \alpha_i \in \mathbb{R}.$$

Offenbar ist

$$\varphi_i(t) = \lambda_0(t; t_0, \varphi_i(t_0)), \quad t \in I.$$

und daher folgt aus der Linearität in Satz 6.1.3 für alle $t \in I$

$$\begin{aligned} 0 &= \lambda_0(t; t_0, 0) = \lambda_0(t; t_0, \alpha_1 \varphi_1(t_0) + \dots + \alpha_m \varphi_m(t_0)) = \alpha_1 \lambda_0(t; t_0, \varphi_1(t_0)) + \dots + \alpha_m \lambda_0(t; t_0, \varphi_m(t_0)) \\ &= \alpha_1 \varphi_1(t) + \dots + \alpha_m \varphi_m(t). \end{aligned}$$

Aus (i) folgt damit $\alpha_i = 0$ für alle i und (ii) ist bewiesen. ■

Den Spezialfall *autonomer* homogener linearer Differentialgleichungen

$$\dot{x} = Ax$$

mit $A \in \mathbb{R}^{n \times n}$ haben wir bereits im Kapitel 1 diskutiert. Die Lösungen sind von der Form

$$\lambda_0(t; t_0, x_0) = e^{A(t-t_0)} x_0 = \sum_{i=0}^{\infty} A^i \frac{(t-t_0)^i}{i!} x_0,$$

insbesondere

$$\lambda_0(t; 0, x_0) = e^{At} x_0 = \sum_{i=0}^{\infty} A^i \frac{t^i}{i!} x_0.$$

Für skalare nichtautonome lineare Differentialgleichungen der Form

$$\dot{x} = a(t)x$$

gilt eine analoge Formel: Die allgemeine Lösung ist, wie man sofort nachrechnet, gegeben durch

$$\lambda(t; t_0, x_0) = e^{\int_{t_0}^t a(s) ds} x_0.$$

Es ist naheliegend zu vermuten, dass eine ähnliche Formel auch im nichtautonomen Fall gilt, also

$$\lambda(t; t_0, x_0) \stackrel{?}{=} e^{\int_{t_0}^t A(s) ds} x_0.$$

Das folgende Beispiel zeigt, dass dies **falsch** ist. (In diesem Beispiel ist, um die Rechnung zu vereinfachen, $A(t)$ nur stückweise stetig).

Beispiel 6.1.5 Sei

$$A(t) := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ für } t \in [0, 1] \text{ und } A(t) := \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \text{ für } t > 1.$$

Man beachte, dass diese beiden Matrizen nicht kommutieren. Betrachte die lineare nichtautonome Differentialgleichung in \mathbb{R}^2 gegeben durch $\dot{x} = A(t)x$, also

$$\dot{x}_1 = x_2, \dot{x}_2 = 0 \text{ für } t \in [0, 1], \text{ und } \dot{x}_1 = 0, \dot{x}_2 = -x_1 \text{ für } t > 1.$$

Sei $x(0) = (1, 0)^T$. Für $t \in [0, 1]$ ist $x_2(t) = 0$ und $x_1(t) = 1$. Für $t \geq 1$ ist

$$x_1(t) = x_1(1) = 1 \text{ und } x_2(t) = 1 - t,$$

also für $t = 2$

$$\begin{pmatrix} x_1(2) \\ x_2(2) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Andererseits ist

$$\int_0^1 A(s) ds = \int_0^1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} ds = \begin{pmatrix} 0 & \int_0^1 1 ds \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

und für $t > 1$

$$\int_0^t A(s)ds = \int_0^1 A(s)ds + \int_1^t A(s)ds = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \int_1^t (-1)ds & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1-t & 0 \end{pmatrix}.$$

Insbesondere erhält man für $t = 2$ und den obigen Anfangswert

$$e^{\int_0^2 A(s)ds} \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos 1 \\ -\sin 1 \end{pmatrix} \neq \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Dies Beispiel zeigt, dass - trotz vieler Ähnlichkeiten - nichtautonome nichtskalare lineare Differentialgleichungen wesentlich komplizierter sind als skalare oder autonome lineare Differentialgleichungen. Sie liefern sogar auch jetzt noch aktuelle Forschungsfragen.

Wir wenden uns jetzt der Darstellung von Lösungen zu.

Definition 6.1.6 Ein Fundamentalsystem von (6.1.2) besteht aus n linear unabhängigen Lösungen $\varphi_1, \dots, \varphi_n$, und eine $n \times n$ -Matrixfunktion $\Phi(t)$, deren Spalten ein Fundamentalsystem bilden, heißt Fundamentalmatrix.

Nach Satz 6.1.3 kann daher jede Lösung als Linearkombination der Lösungen eines Fundamentalsystems dargestellt werden. Nach Proposition 6.1.4 ist dann für alle $t \in I$ die Determinante $\det \Phi(t) \neq 0$, also $\Phi(t)$ eine reguläre Matrix. Offenbar ist jede Fundamentalmatrix $\Phi(t)$ eine Lösung der Matrix-Differentialgleichung in $\mathbb{R}^{n \times n}$

$$\dot{X}(t) = A(t)X(t). \quad (6.1.5)$$

Die Lösungen dieser Differentialgleichung mit Anfangsbedingungen der Form

$$X(t_0) = I_n, \quad I_n \text{ die } n \times n\text{-Einheitsmatrix,}$$

bezeichnen wir mit $\Phi(t, t_0)$, $t, t_0 \in I$, und nennen sie **Übergangsmatrix** oder Hauptfundamentalmatrix. Sie ist eine spezielle Fundamentalmatrix, und die allgemeine Lösung der homogenen Gleichung $\dot{x} = A(t)x$ erfüllt

$$\lambda_0(t; t_0, x_0) = \Phi(t, t_0)x_0, \quad (6.1.6)$$

weil auch für die rechte Seite gilt

$$\frac{d}{dt} [\Phi(t, t_0)x_0] = A(t)\Phi(t, t_0)x_0 \text{ und } \Phi(t_0, t_0)x_0 = x_0.$$

Die Übergangsmatrix hat die folgenden Eigenschaften.

Proposition 6.1.7 Die Übergangsmatrix $\Phi(t, t_0)$ von (6.1.2) erfüllt für alle $t_0, t_1, t_2 \in I$:

- (i) $\Phi(t_0, t_0) = I_n$;
- (ii) $\Phi(t_2, t_1)\Phi(t_1, t_0) = \Phi(t_2, t_0)$;
- (iii) $\Phi(t_0, t_1)^{-1} = \Phi(t_1, t_0)$.
- (iv) Ferner ist $\Phi(t, t_0)$ stetig differenzierbar in jedem Argument.

Beweis. (i) ist klar; (ii) folgt aus (6.1.6) und der Kozykel-Eigenschaft für $\dot{x} = A(t)x$ (siehe Proposition 2.4.11): für alle $x_0 \in \mathbb{R}^n$ ist

$$\Phi(t_2, t_1) [\Phi(t_1, t_0)x_0] = \lambda_0(t_2; t_1, \lambda_0(t_1; t_0, x_0)) = \lambda_0(t_2; t_0, x_0) = \Phi(t_2, t_0)x_0.$$

(iii) folgt aus (i) und (ii), weil

$$I_n = \Phi(t_0, t_0) = \Phi(t_0, t_1)\Phi(t_1, t_0) = \Phi(t_0, t_1)\Phi(t_0, t_1)^{-1}. \quad (6.1.7)$$

Stetige Differenzierbarkeit von $\Phi(t, t_0)$ im ersten Argument gilt nach Definition. Stetige Differenzierbarkeit im zweiten Argument sieht man folgendermaßen. Es sei $Y(t), t \in I$ die eindeutige (stetig differenzierbare) Matrixlösung von

$$\dot{Y}(t) = -Y(t)A(t) \text{ mit } Y(t_0) = I.$$

Dann gilt nach der Produktregel

$$\begin{aligned} \frac{d}{dt} [Y(t)\Phi(t, t_0)] &= \dot{Y}(t)\Phi(t, t_0) + Y(t)\frac{d}{dt}\Phi(t, t_0) \\ &= -Y(t)A(t)\Phi(t, t_0) + Y(t)A(t)\Phi(t, t_0) = 0, \end{aligned}$$

also

$$Y(t)\Phi(t, t_0) \equiv Y(t_0)\Phi(t_0, t_0) = I_n.$$

Es folgt, dass $\Phi(t_0, t) = \Phi(t, t_0)^{-1} = Y(t)$ stetig differenzierbar in t ist. ■

Wir notieren die folgenden Eigenschaften von Fundamentalmatrizen.

Proposition 6.1.8 (i) Sei $\Phi(t), t \in I$, eine Fundamentalmatrix von (6.1.2). Dann gilt für die allgemeine Lösung

$$\lambda_0(t; t_0, x_0) = \Phi(t)\Phi(t_0)^{-1}x_0.$$

(ii) Ist $\Psi(t)$ eine weitere Fundamentalmatrix, so gibt es eine reguläre Matrix C mit $\Phi(t) = \Psi(t)C$ für alle $t \in I$.

Beweis. (i) Nach Proposition 6.1.4 ist $\Phi(t)$ invertierbar und die rechte Seite ist eine Lösung der Differentialgleichung mit Anfangswert x_0 zur Zeit t_0 .

(ii) Nach (i) gilt

$$\Phi(t)\Phi(t_0)^{-1}x_0 = \lambda_0(t; t_0, x_0) = \Psi(t)\Psi(t_0)^{-1}x_0.$$

Wählt man für x_0 der Reihe nach die Einheitsvektoren in \mathbb{R}^n , so sieht man, dass die beiden Matrizen spaltenweise übereinstimmen, also

$$\Phi(t)\Phi(t_0)^{-1} = \Psi(t)\Psi(t_0)^{-1}.$$

Multiplikation von rechts mit $\Phi(t_0)$ liefert die Behauptung mit $C := \Psi(t_0)^{-1}\Phi(t_0)$. ■

Wir wenden uns jetzt der inhomogenen Gleichung (6.1.1)

$$\dot{x} = A(t)x + g(t)$$

zu. Das Hauptergebnis dieses Abschnitts ist die folgende Lösungsdarstellung, bekannt als Formel der *Variation-der-Konstanten* (oder *Variation-der-Parameter*).

Satz 6.1.9 Für die allgemeine Lösung der inhomogenen Gleichung (6.1.1) gilt

$$\lambda_g(t; t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^t \Phi(t, s)g(s)ds, \quad (6.1.8)$$

wobei $\Phi(t, s)$ die Übergangsmatrix ist.

Beweis. Nach (6.1.6) ist $\Phi(t, t_0)x_0$ die Lösung der homogenen Gleichung. Es bleibt nach Satz 6.1.3 also nur zu zeigen, dass

$$\varphi_g^*(t) := \int_{t_0}^t \Phi(t, s)g(s)ds$$

eine (partikuläre) Lösung der inhomogenen Gleichung mit $\varphi_g^*(t_0) = 0$ ist. Nun ist aber für jede Fundamentalmatrix Ψ

$$\Phi(t, s) = \Psi(t)\Psi(s)^{-1},$$

weil beide Seiten Lösung von (6.1.5) mit $\Phi(t, t) = \Psi(t)\Psi(t)^{-1} = I_n$ sind (zum Beispiel gilt $\Phi(t, s) = \Phi(t, t_0)\Phi(s, t_0)^{-1}$). Differentiation nach der Produktregel liefert

$$\begin{aligned}\frac{d}{dt}\varphi_g^*(t) &= \frac{d}{dt} \int_{t_0}^t \Phi(t, s)g(s)ds = \frac{d}{dt} \left[\Psi(t) \int_{t_0}^t \Psi(s)^{-1}g(s)ds \right] \\ &= A(t)\Psi(t) \int_{t_0}^t \Psi(s)^{-1}g(s)ds + \Psi(t)\Psi(t)^{-1}g(t) \\ &= A(t)\varphi_g^*(t) + g(t).\end{aligned}$$

■

Bemerkung 6.1.10 Die Bezeichnung *Variation-der-Konstanten* oder *Variation-der-Parameter* erklärt sich folgendermaßen: Für die skalare Gleichung $\dot{x} = a(t)x$ ist

$$\Phi(t, s) = e^{\int_s^t a(r)dr}. \quad (6.1.9)$$

Zur Lösung der inhomogenen Gleichung $\dot{x} = ax + g, x(t_0) = x_0$, machen wir den folgenden Ansatz, der die Lösung $u(t) = e^{\int_{t_0}^t a(s)ds}$ (mit $u(t_0) = 1$) der homogenen Gleichung variiert:

$$x(t) = c(t)u(t),$$

mit einer stetig differenzierbaren, zu bestimmenden Funktion $c(t)$, den „variierenden Konstanten“. Offenbar ist dann $c(t_0) = x(t_0) = x_0$. Durch Einsetzen in (6.1.1) ergibt sich

$$\dot{c}u + \dot{u}c = acu + g,$$

also wegen $\dot{u} = au$,

$$\dot{c}u + auc = acu + g$$

und daher

$$\dot{c}(t) = \frac{g(t)}{u(t)} = \frac{g(t)}{e^{\int_{t_0}^t a(r)dr}} = g(t)e^{-\int_{t_0}^t a(r)dr}.$$

Daraus ergibt sich

$$\begin{aligned}x(t) &= c(t)u(t) = c(t_0)u(t) + \int_{t_0}^t u(t)\dot{c}(s)ds = e^{\int_{t_0}^t a(s)ds} x_0 + \int_{t_0}^t g(s)e^{-\int_{t_0}^s a(r)dr} e^{\int_{t_0}^t a(r)dr} ds \\ &= e^{\int_{t_0}^t a(s)ds} x_0 + \int_{t_0}^t e^{\int_s^t a(r)dr} g(s)ds.\end{aligned}$$

Mit (6.1.9) folgt die *Variation-der-Konstanten-Formel*.

Bemerkung 6.1.11 An dieser Stelle ist eine Warnung angebracht: Die *Variation-der-Konstanten-Formel* (6.1.8) liefert eine handliche Darstellung der Lösung. Für allgemeine nichtautonome Differentialgleichungen ist es aber unmöglich, eine explizite Formel für die Übergangsmatrix $\Phi(t, s)$ anzugeben.

Wie gezeigt, ist die Lösung von nicht-autonomen linearen Differentialgleichungen im Allgemeinen nicht mit Hilfe der Exponentialfunktion darstellbar. Dies geht nur im skalaren Fall. Man kann sich aber auch im n -dimensionalen Fall eine skalare Größe verschaffen, indem man die Determinante einer Fundamentalmatrix betrachtet. Etwas allgemeiner definieren wir für eine $n \times n$ -Matrixfunktion $W(t)$, deren Spalten Lösungen von (6.1.2) sind, die **Wronski-Determinante** als

$$w(t) = \det W(t).$$

Die Matrix $W(t)$ heißt dann auch *Wronski-Matrix*. Wegen Proposition 6.1.4 ist die Determinante entweder identisch Null oder verschwindet für kein t .

Die Spur $\text{spur} A$ einer quadratischen Matrix A ist die Summe der Diagonalelemente. Wir erhalten die folgende **Liouville-Formel**.

Satz 6.1.12 Die Wronski-Determinante von (6.1.2) ist eine Lösung der skalaren homogenen linearen Differentialgleichung

$$\dot{w}(t) = [\text{spur}A(t)] w(t).$$

Daher gilt

$$w(t) = w(t_0) e^{\int_{t_0}^t [\text{spur}A(s)] ds} \text{ für alle } t, t_0 \in I.$$

Beweis. Sei $W(t) = (w_{ij}(t))_{i,j=1,\dots,n}$ und schreibe

$$W(t) = \begin{pmatrix} w_1(t) \\ \vdots \\ w_n(t) \end{pmatrix},$$

wobei $w_i(t) = (w_{i1}(t), \dots, w_{in}(t))$ die i te Zeile von $W(t)$ bezeichnet. Es gilt

$$\begin{pmatrix} \dot{w}_{11}(t) & \dots & \dot{w}_{1n}(t) \\ \vdots & & \vdots \\ \dot{w}_{n1}(t) & \dots & \dot{w}_{nn}(t) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{pmatrix} \begin{pmatrix} w_{11}(t) & \dots & w_{1n}(t) \\ \vdots & & \vdots \\ w_{n1}(t) & \dots & w_{nn}(t) \end{pmatrix}$$

(oder $\dot{w}_{ij} = \sum_{k=1}^n a_{ik} w_{kj}$). Zeilenweise geschrieben ist dies

$$\begin{aligned} \dot{w}_1(t) &= a_{11}(t)w_1(t) + \dots + a_{1n}(t)w_n(t) \\ &\vdots \\ \dot{w}_n(t) &= a_{n1}(t)w_1(t) + \dots + a_{nn}(t)w_n(t) \end{aligned} \tag{6.1.10}$$

(oder $\dot{w}_i = (\dot{w}_{i1}, \dots, \dot{w}_{in}) = (\sum_{k=1}^n a_{ik} w_{k1}, \dots, \sum_{k=1}^n a_{ik} w_{kn})$). Betrachte die Determinanten-Entwicklung

$$\det W(t) = \sum_{\sigma \in \Sigma_n} (\text{sgn}\sigma) w_{1,\sigma(1)}(t) \dots w_{n,\sigma(n)}(t),$$

wobei Σ_n die Permutationsgruppe ist und σ die Signatur einer Permutation σ , also $= 1$, falls die Anzahl der Transpositionen, als deren Hintereinanderausführung σ geschrieben werden kann, gerade ist und $= -1$ sonst. Differentiation liefert

$$\begin{aligned} \frac{d}{dt} \det W(t) &= \sum_{\sigma \in \Sigma_n} (\text{sgn}\sigma) \dot{w}_{1,\sigma(1)}(t) \dots w_{n,\sigma(n)}(t) + \dots + \sum_{\sigma \in \Sigma_n} (\text{sgn}\sigma) w_{1,\sigma(1)}(t) \dots \dot{w}_{n,\sigma(n)}(t) \\ &= \det \begin{pmatrix} \dot{w}_1(t) \\ w_2(t) \\ \vdots \\ w_n(t) \end{pmatrix} + \dots + \det \begin{pmatrix} w_1(t) \\ w_2(t) \\ \vdots \\ \dot{w}_n(t) \end{pmatrix} \\ &= a_{11}(t) \det W(t) + \dots + a_{nn}(t) \det W(t). \end{aligned}$$

Die letzte Gleichheit folgt, wenn man (6.1.10) einsetzt, Multilinearität ausnutzt, und beachtet, dass Determinanten mit identischen Zeilen verschwinden. Zum Beispiel erhält man für den ersten

Summanden

$$\begin{aligned} \det \begin{pmatrix} \dot{w}_1(t) \\ w_2(t) \\ \cdot \\ \cdot \\ w_n(t) \end{pmatrix} &= \det \begin{pmatrix} a_{11}(t)w_1(t) + \dots + a_{1n}(t)w_n(t) \\ w_2(t) \\ \cdot \\ \cdot \\ w_n(t) \end{pmatrix} \\ &= a_{11}(t) \det \begin{pmatrix} w_1(t) \\ w_2(t) \\ \cdot \\ \cdot \\ w_n(t) \end{pmatrix} + \dots + a_{1n}(t) \det \begin{pmatrix} w_n(t) \\ w_2(t) \\ \cdot \\ \cdot \\ w_n(t) \end{pmatrix} \\ &= a_{11}(t) \det W(t). \end{aligned}$$

Wir erhalten

$$\frac{d}{dt} \det W(t) = \sum_{i=1}^n a_{ii}(t) \det W(t) = \operatorname{spur} A(t) \det W(t).$$

■

6.2 Autonome Lineare Differentialgleichungen

In diesem Abschnitt betrachten wir lineare Differentialgleichungen mit konstanten Koeffizienten. Bekanntermaßen lassen sich für Matrizen Transformationen angeben, die sie in einfachere Form überführen. Die bekannteste ist die Jordan-Normalform, die man durch Koordinatentransformationen erhält. Wir werden dies ausnutzen, um auch Lösungen linearer autonomer Differentialgleichungen nach Koordinatentransformation einfach aufzuschreiben.

Wir betrachten in diesem Abschnitt lineare autonome Differentialgleichungen der Form

$$\dot{x} = Ax$$

mit $A \in \mathbb{R}^{n \times n}$. Wie wir in Kapitel 1 gesehen haben, ist der zugehörige Fluss gegeben durch

$$\lambda(t; x_0) = e^{At} x_0$$

und die Übergangsmatrix ist $\Phi(t, t_0) = e^{A(t-t_0)}$.

Das folgende Lemma sammelt einige weitere Eigenschaften der Matrix-Exponentialfunktion.

Lemma 6.2.1 *Sei $A \in \mathbb{R}^{n \times n}$ oder $A \in \mathbb{C}^{n \times n}$ und $t, s \in \mathbb{R}$. Dann gelten*

(i) $e^{(T^{-1}AT)t} = T^{-1}e^{At}T$, falls T eine reguläre Matrix ist.

(ii) $e^0 = I$, $e^{At}e^{As} = e^{A(t+s)}$, $(e^{At})^{-1} = e^{-At}$.

(iii) $e^{At}e^{Bt} = e^{(A+B)t}$, falls $AB = BA$.

(iv) Besitzt A Blockdiagonalform mit $A = \operatorname{diag}[A_1, \dots, A_k]$, so auch e^{At} und

$$e^{At} = \operatorname{diag}[e^{A_1 t}, \dots, e^{A_k t}].$$

Beweis. Übungsaufgaben. ■

Die Variation-der-Parameter-Formel für

$$\dot{x} = Ax + g(t), \quad x(t_0) = x_0$$

hat hier die Form

$$\lambda_g(t; t_0, x_0) = e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-s)} g(s) ds.$$

Die entscheidende Frage ist jetzt: Wie berechnet man (analytisch) e^{At} ? Dies ist offenbar sehr einfach, wenn A eine Diagonalmatrix ist, also

$$A = \text{diag}[\lambda_1, \dots, \lambda_n].$$

Dann ist

$$e^{At} = \text{diag}[e^{\lambda_1 t}, \dots, e^{\lambda_n t}]. \quad (6.2.1)$$

Die Berechnung ist auch einfach, wenn A nilpotent ist, also $A^k = 0$ für ein $k \in \mathbb{N}$. Dann ist

$$e^{At} = \sum_{i=0}^{\infty} \frac{A^i}{i!} t^i = \sum_{i=0}^{k-1} \frac{A^i}{i!} t^i, \quad (6.2.2)$$

ist also einfach gegeben durch eine Linearkombination von Potenzen von A . Unsere Strategie, um die Matrix-Exponentialfunktion und damit die Lösung der Differentialgleichung darzustellen, besteht jetzt zunächst darin, die Matrix A in die Summe einer Diagonalmatrix und einer nilpotenten Matrix zu zerlegen. Die *Jordansche Normalform* im Körper der komplexen Zahlen zeigt, dass dies nach Konjugation mit einer invertierbaren Matrix möglich ist.

Erinnerung an Eigenwerte: Die Eigenwerte einer quadratischen Matrix über \mathbb{R} oder \mathbb{C} sind die Nullstellen des charakteristischen Polynoms

$$\det[A - \lambda I] = (-1)^n \lambda^n + a_{n-1} \lambda^{n-1} \dots + a_0$$

(man beachte, dass $a_0 = \det A$). Dieses Polynom n ten Grades hat n (nicht notwendig verschiedene) Nullstellen in \mathbb{C} . Ist $A \in \mathbb{R}^{n \times n}$, so sind die Koeffizienten reell und mit $\lambda \in \mathbb{C}$ ist auch $\bar{\lambda}$ Nullstelle.

Satz 6.2.2 Sei $A \in \mathbb{C}^{n \times n}$. Dann existieren eine invertierbare Matrix $T \in \mathbb{C}^{n \times n}$ sowie eine Diagonalmatrix S und eine nilpotente Matrix N , so dass

$$T^{-1}AT = S + N \text{ und } SN = NS. \quad (6.2.3)$$

Hierbei ist $S = \text{diag}[\lambda_1, \dots, \lambda_n]$, wobei λ_i die (nicht notwendigerweise verschiedenen) Eigenwerte von A sind, und $N = \text{blockdiag}[N_1, \dots, N_i]$, mit Blöcken der Form

$$N_i = \begin{pmatrix} 0 & 1 & & & 0 \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ 0 & & & & 0 \end{pmatrix}.$$

Aus (6.2.3) folgt mit den Rechenregeln für die Exponentialfunktion

$$e^{At} = e^{T(S+N)T^{-1}t} = T e^{(S+N)t} T^{-1} = T e^{St} e^{Nt} T^{-1}.$$

Wie in (6.2.1) und (6.2.2) lassen sich dann e^{St} und e^{Nt} berechnen. Man erhält die Lösung der ursprünglichen Differentialgleichung $\dot{x} = Ax$ durch eine Koordinatentransformation: Sei $y(t)$ Lösung von $\dot{y} = (S + N)y(t) = (T^{-1}AT)y(t)$. Dann erfüllt $x(t) := Ty(t)$

$$\dot{x}(t) = T\dot{y}(t) = TT^{-1}ATy(t) = ATT^{-1}x(t) = Ax(t).$$

Es ist leicht, Matrizen anzugeben, die nicht diagonalisierbar sind, z.B. $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.

Leider gibt es die Zerlegung (6.2.3) im Körper der reellen Zahlen im Allgemeinen nicht. Für

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \text{ ist } \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

die Jordan-Normalform. Für diese Matrix zum linearen Oszillator $\dot{x}_1 = x_2$, $\dot{x}_2 = -x_1$ sind ja $\pm i$ die Eigenwerte. Die Übergangsmatrix ist gegeben durch (vgl. Beispiel 3.2.1)

$$e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}t} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}. \quad (6.2.4)$$

Das kann man durch Betrachten der Reihendarstellung herleiten, oder direkt durch Einsetzen in die Matrix-Differentialgleichung $\dot{X} = AX$ nachprüfen.

Ein etwas komplizierteres Beispiel beschreibt die folgende Proposition.

Proposition 6.2.3 Für $a, b \in \mathbb{R}$ betrachte

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Dann ist

$$e^{At} = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}. \quad (6.2.5)$$

Beweis. Wegen

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

ist

$$e^{At} = e^{at} e^{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}bt}.$$

Mit (6.2.4) folgt (6.2.5). ■

Ist man an reellen Lösungen interessiert, so kann man an Stelle der komplexen die etwas kompliziertere *reelle Jordan-Normalform* verwenden: Es gibt eine reelle invertierbare Matrix T , so dass $T^{-1}AT$ die Form

$$\begin{pmatrix} R & 0 \\ 0 & K \end{pmatrix}$$

hat, wobei R die reellen Eigenwerte von A und K die komplex-konjugierten Eigenwertpaare von A als Eigenwerte besitzt und die Matrizen R und K die folgende Form haben: Die reellen Eigenwerte können wir genauso wie im komplexen Fall behandeln:

$$R = \text{diag}[J_1, \dots, J_m]$$

mit Jordanblöcken der Gestalt

$$J_j = \begin{pmatrix} \lambda_j & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \lambda_j \end{pmatrix} \quad (6.2.6)$$

wobei die λ_j die reellen Eigenwerte von A sind. Die komplexen Eigenwerte von A treten in konjugiert-komplexen Paaren auf, also $a_j \pm ib_j$, $a_j, b_j \in \mathbb{R}$, $b_j \neq 0$, $i = \sqrt{-1}$. Die Matrix K hat Block-Diagonalgestalt

$$K = \text{blockdiag}[K_1, \dots, K_r].$$

Hierbei sind die K_i von ähnlicher Gestalt wie die J_i . Allerdings stehen an Stelle reeller Zahlen 2×2 -Matrizen,

$$K_j = \begin{pmatrix} \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \dots & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \vdots & \vdots & \vdots & \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \dots & \dots & \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix} \end{pmatrix}. \tag{6.2.7}$$

Damit ist die *reelle Jordanform* beschrieben.

Nach Lemma 6.2.1 (iv) reicht es, die Lösungen jeweils explizit für einzelne Jordanblöcke anzugeben.

Proposition 6.2.4 *Betrachte $\dot{x} = Jx$ wobei*

$$J = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 1 \\ 0 & \dots & \dots & \lambda \end{pmatrix}$$

ein d -dimensionaler reeller Jordanblock der Form (6.2.6) ist. Dann ist

$$e^{Jt} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{d-1}}{(d-1)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}. \tag{6.2.8}$$

Beweis. Mit $S = \lambda I$ und

$$N = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 1 \\ 0 & \dots & \dots & 0 \end{pmatrix}$$

folgt die Behauptung aus

$$e^{Jt} = e^{(S+N)t} = e^{\lambda t} e^{Nt} = e^{\lambda t} \sum_{k=0}^{d-1} \frac{t^k N^k}{k!}.$$

■

Es bleiben die Jordanblöcke für komplexe Eigenwerte.

Proposition 6.2.5 *Betrachte $\dot{x} = Kx$, wobei*

$$K = \begin{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \dots & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \vdots & \vdots & \vdots & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \dots & \dots & \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \end{pmatrix}$$

ein $2d$ -dimensionaler komplexer Jordanblock der Form (6.2.7) ist. Dann ist

$$e^{Kt} = e^{at} \begin{pmatrix} G(t) & tG(t) & \frac{t^2}{2}G(t) & \dots & \frac{t^{d-1}}{(d-1)!}G(t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & 0 & G(t) \end{pmatrix} \text{ mit } G(t) := \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}. \quad (6.2.9)$$

Beweis. Kürze ab

$$0_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} a & b \\ -b & a \end{pmatrix},$$

und

$$N = \begin{pmatrix} 0_2 & I_2 & \cdot & \cdot & 0_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & I_2 & \cdot \\ 0_2 & \cdot & \cdot & \cdot & 0_2 \end{pmatrix}, \quad S = \begin{pmatrix} C & 0_2 & \cdot & \cdot & 0_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0_2 \\ 0_2 & \cdot & \cdot & \cdot & C \end{pmatrix}.$$

Dann ist $K = N + S$ und

$$NS = \begin{pmatrix} 0_2 & C & \cdot & \cdot & 0_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & C \\ 0_2 & \cdot & \cdot & \cdot & 0_2 \end{pmatrix} = SN.$$

Es folgt

$$e^{Kt} = e^{(N+S)t} = e^{Nt}e^{St} = \left(\sum_{j=0}^{d-1} \frac{t^j}{j!} N^j \right) \text{blockdiag}[e^{Ct}, \dots, e^{Ct}].$$

Mit $e^{Ct} = e^{at}G(t) = e^{at} \begin{pmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{pmatrix}$ liefert dies (6.2.9). ■

6.3 Klassifikation in der Ebene

In diesem Abschnitt werden die möglichen Phasenporträts für autonome homogene lineare Differentialgleichungen in der Ebene beschrieben.

Bis auf Koordinatentransformationen im \mathbb{R}^n sind die Gleichungen

$$\dot{x} = Ax$$

durch Matrizen A in reeller Jordan-Normalform beschrieben. Der zugehörige Fluss ist dann gegeben durch

$$\varphi(t; x_0) = e^{At}x_0.$$

In diesem Abschnitt beschränken wir uns auf Gleichungen für $n = 2$. Hier gibt es die folgenden verschiedenen reellen Jordan-Normalformen

$$\begin{bmatrix} \rho & 0 \\ 0 & \sigma \end{bmatrix}, \quad \begin{bmatrix} \rho & 1 \\ 0 & \rho \end{bmatrix}, \quad \text{und} \quad \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \text{ mit } b \neq 0.$$

Die zugehörigen Eigenwerte sind im ersten Fall $\rho, \sigma \in \mathbb{R}$, im zweiten Fall $\rho \in \mathbb{R}$, und im dritten Fall $a \pm ib \in \mathbb{C} \setminus \mathbb{R}$. Die Matrix A besitzt dann jeweils zwei linear unabhängige Eigenvektoren, nur einen bzw. keinen reellen Eigenvektor. Nach den Propositionen 6.2.4 und 6.2.5 gilt

$$e^{\begin{bmatrix} \rho & 0 \\ 0 & \sigma \end{bmatrix}t} = \begin{bmatrix} e^{\rho t} & 0 \\ 0 & e^{\sigma t} \end{bmatrix}, \quad e^{\begin{bmatrix} \rho & 1 \\ 0 & \rho \end{bmatrix}t} = \begin{bmatrix} e^{\rho t} & te^{\rho t} \\ 0 & e^{\rho t} \end{bmatrix}, \quad e^{\begin{bmatrix} a & b \\ -b & a \end{bmatrix}t} = e^{at} \begin{bmatrix} \cos bt & \sin bt \\ -\sin bt & \cos bt \end{bmatrix}.$$

Um die zugehörigen Phasenporträts zu beschreiben, müssen wir auch noch genauer unterscheiden, welche Vorzeichen die (Realteile der) Eigenwerte haben (Illustrationen der Phasenporträts findet man in Aulbach [4, Abschnitt 5.3]).

Fall I. Die Matrix A ist regulär,

$$\text{Fall I.1: } A = \begin{bmatrix} \rho & 0 \\ 0 & \sigma \end{bmatrix}, \quad \rho, \sigma \in \mathbb{R}, \quad \rho \neq \sigma.$$

Dann ist

$$\dot{x} = \rho x, \quad \dot{y} = \sigma y.$$

Der Fluss ist $\varphi(t; x_0, y_0) = (e^{\rho t} x_0, e^{\sigma t} y_0), t \in \mathbb{R}$. Die Trajektorie durch $(x_0, y_0) \in \mathbb{R}^2$ ist also gegeben durch

$$\mathcal{O}(x_0, y_0) = \{(e^{\rho t} x_0, e^{\sigma t} y_0) \in \mathbb{R}^2 \mid t \in \mathbb{R}\}.$$

Neben der Ruhelage $(0, 0)$ sind die vier Halbachsen Trajektorien, und in den vier offenen Quadranten des \mathbb{R}^2 ist

$$\mathcal{O}(x_0, y_0) = \left\{ \left(x, y_0 \left(\frac{x}{x_0} \right)^{\frac{\sigma}{\rho}} \right) \mid \frac{x}{x_0} > 0 \right\},$$

was sich durch Elimination des Parameters t ergibt:

$$x = e^{\rho t} x_0 \iff t = \frac{1}{\rho} \log \frac{x}{x_0};$$

dies wird in $e^{\sigma t} y_0$ eingesetzt. Es ergeben sich drei Fälle: Ist $\rho < \sigma < 0$, so nennt man 0 einen stabilen (zwei-tangentigen) Knoten. Ist $\rho < 0 < \sigma$, so heißt 0 Sattelpunkt, und ist $0 < \rho < \sigma$, so heißt 0 instabiler (zwei-tangentiger) Knoten.

$$\text{Fall I.2: } A = \begin{pmatrix} \rho & 0 \\ 0 & \rho \end{pmatrix}, \quad \rho \in \mathbb{R}, \quad \rho \neq 0.$$

Dann ist

$$\dot{x} = \rho x, \quad \dot{y} = \rho y, \quad \text{also } \varphi(t; x_0, y_0) = (e^{\rho t} x_0, e^{\rho t} y_0).$$

Ist $\rho < 0$ so nennt man $0 \in \mathbb{R}^2$ einen stabilen (viel-tangentigen) Knoten und, ist $\rho > 0$, einen instabilen (viel-tangentigen) Knoten.

$$\text{Fall I.3: } A = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}, \quad \rho \in \mathbb{R}, \quad \rho \neq 0.$$

Dann ist

$$\dot{x} = \rho x + y, \quad \dot{y} = \rho y, \quad \text{also } \varphi(t; x_0, y_0) = (e^{\rho t} x_0 + y_0 t e^{\rho t}, e^{\rho t} y_0).$$

Setzen wir $y_0 = 0$, so erhalten wir neben der Ruhelage $(0, 0)$ die beiden x -Halbachsen als Trajektorien. Für $y_0 \neq 0$ liefert Elimination von t

$$\mathcal{O}(x_0, y_0) = \left\{ \left(\frac{x_0}{y_0} y + \frac{y}{\rho} \log \frac{y}{y_0}, y \right) \in \mathbb{R}^2, \quad \frac{y}{y_0} > 0 \right\}.$$

Für $\rho < 0$ heißt $0 \in \mathbb{R}^2$ stabiler (ein-tangentiger) Knoten und für $\rho > 0$ instabiler (ein-tangentiger) Knoten. Für die Herleitung des Phasenporträts ist es wichtig zu bemerken, dass auf der Linie $y = -\rho x$ die Trajektorien eine senkrechte Tangente haben, weil hier $\dot{x} = 0$ ist. Ferner sind die Trajektorien tangential im Nullpunkt an die x -Achse, weil für die Steigung $\frac{y}{x}$ gilt

$$\frac{y}{\frac{x_0}{y_0} y + \frac{y}{\rho} \log \frac{y}{y_0}} = \frac{1}{\frac{x_0}{y_0} + \frac{1}{\rho} \log \frac{y}{y_0}} \rightarrow 0 \text{ für } y \rightarrow 0.$$

$$\text{Fall I.4: } A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad a, b \in \mathbb{R}, \quad b \neq 0.$$

Dann ist

$$\dot{x} = ax + by, \quad \dot{y} = -bx + ay,$$

also

$$\varphi(t; x_0, y_0) = e^{at} (x_0 \cos(bt) + y_0 \sin(bt), y_0 \cos(bt) - x_0 \sin(bt)).$$

Hier ist es hilfreich, Polarkoordinaten zu betrachten: Jeder Punkt $0 \neq (x, y) \in \mathbb{R}^2$ lässt sich eindeutig in Polarkoordinaten $x = r \cos \phi, y = r \sin \phi, r > 0, \phi \in [0, 2\pi)$ darstellen. Dann definiert jede Lösung von $\dot{r} = ar, \dot{\phi} = -b$ eine Lösung $(x(t), y(t))$ der obigen Differentialgleichung. In der Tat, man rechnet nach

$$\begin{aligned}\dot{x} &= \dot{r} \cos \phi - r \sin \phi \cdot \dot{\phi} = ar \cos \phi + br \sin \phi = ax + by, \\ \dot{y} &= \dot{r} \sin \phi + r \cos \phi \cdot \dot{\phi} = ar \sin \phi - br \cos \phi = ay - bx.\end{aligned}$$

Daher ist der Radius $r(t) = e^{at}r(0)$ und der Winkel $\phi(t) = -bt$. Für $a < 0$ heißt $0 \in \mathbb{R}^2$ stabiler Strudel. Ist $b < 0$, so drehen sich die Trajektorien im mathematisch positiven Sinn, ist $b > 0$, so im mathematisch negativen Sinn. Für $a = 0$ heißt 0 Zentrum oder Wirbel und analog für $b < 0$ und $b > 0$. Für $a > 0$ heißt 0 instabiler Strudel und analog für $b < 0$ und $b > 0$.

Fall II: Die Matrix A ist singulär.

Fall II.1: $B = \begin{pmatrix} \rho & 0 \\ 0 & 0 \end{pmatrix}$, $\rho \in \mathbb{R}, \rho \neq 0$. Das zugehörige System

$$\dot{x} = \rho x, \dot{y} = 0$$

besitzt den Fluss

$$\varphi(t; x_0, y_0) = (e^{\rho t} x_0, y_0).$$

Die y -Achse besteht aus Ruhelagen, die anderen Trajektorien verlaufen parallel zur x -Achse.

Für $\rho < 0$ laufen die Trajektorien für $t \rightarrow \infty$ gegen die y -Achse und für $\rho > 0$ laufen die Trajektorien für $t \rightarrow -\infty$ gegen die y -Achse.

Fall II.2: $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Jeder Punkt ist Ruhelage.

Fall II.3: $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Der Fluss des Systems $\dot{x} = y, \dot{y} = 0$, ist gegeben durch

$$\varphi(t; x_0, y_0) = (x_0 + y_0 t, y_0).$$

Die x -Achse besteht aus Ruhelagen. Die anderen Trajektorien verlaufen parallel zur x -Achse.

Damit ist die Klassifizierung der linearen autonomen Systeme in der Ebene abgeschlossen. Man beachte, dass alle Trajektorien für $t \rightarrow \infty$ gegen den Ursprung laufen, falls alle Eigenwerte Realteil kleiner als Null haben, und für $t \rightarrow -\infty$ gegen den Ursprung laufen, falls alle Eigenwerte Realteil größer als Null haben. Koordinatentransformationen haben offenbar nur einen unwesentlichen Einfluss auf das Limesverhalten dieser Systeme für $t \rightarrow \pm\infty$. Daher ist dieses asymptotische Verhalten durch die Eigenwerte und die Eigenraumstruktur festgelegt.

Anstatt eine Matrix $A \in \mathbb{R}^{2 \times 2}$ in Jordan-Normalform überzuführen, um das Verhalten der Lösungen zu beschreiben, ist eine andere Darstellung häufig nützlich. Für

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ist

$$\det[A - \lambda I] = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = \lambda^2 - \lambda(a + d) + ad - bc = \lambda^2 - \lambda \operatorname{spur} A + \det A.$$

Mit $s = \operatorname{spur} A$ und $d = \det A$ sind die Eigenwerte daher

$$\lambda_{1,2} = \frac{1}{2} \left(s \pm \sqrt{s^2 - 4d} \right).$$

Damit ist klar, dass die Punkte in der (s, d) -Ebene die Eigenwerte bestimmen. Zum Beispiel gehören zu den Punkten auf der Parabel

$$d = \frac{1}{4} s^2$$

die Matrizen mit doppeltem reellem Eigenwert; zu den (s, d) -Paaren oberhalb der Parabel gehören komplex-konjugierte Eigenwert-Paare, zu denen unterhalb zwei verschiedene reelle Eigenwerte. Genau die Punkte auf der s -Achse haben einen Eigenwert 0. Die Punkte auf der d -Achse mit $d > 0$ haben imaginäre Eigenwerte.

Schließlich noch eine Warnung/Erinnerung: Am charakteristischen Polynom kann man die Zahl und Größe der Jordanblöcke nicht ablesen. Zum Beispiel haben die Matrizen (mit zwei bzw. einem Jordanblock)

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

beide das charakteristische Polynom λ^2 . Stattdessen muss man $\ker [A - \lambda I]^j, j \geq 1$, ansehen.

6.4 Ein Beispiel: Der Lineare Oszillator

In diesem Abschnitt diskutieren wir ein besonders wichtiges Beispiel von linearen Differentialgleichungen, den linearen Oszillator (oder die lineare Schwingungsgleichung), die in zahlreichen Anwendungen in Physik und Ingenieurwissenschaften eine Rolle spielt. Es handelt sich um die folgende Klasse von linearen Differentialgleichungen zweiter Ordnung:

$$\ddot{x} + \alpha \dot{x} + \beta x = f(t) \tag{6.4.1}$$

mit $\alpha, \beta \in \mathbb{R}$ und einer stetigen Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$. Dabei wird häufig (in Anlehnung an die linearisierte Pendelgleichung) $\alpha \dot{x}$ als Reibungs- oder Dämpfungsterm und βx als Rückstellkraft bezeichnet. Ist $f \equiv 0$, so spricht man bei den Lösungen von freien Schwingungen, sonst von erzwungenen Schwingungen (wir werden gleich sehen, inwieweit diese Bezeichnungen korrekt sind). Wir betrachten nur zwei Spezialfälle und beginnen mit

$$\ddot{x} + 2a\dot{x} + x = 0 \quad \text{mit } a \geq 0. \tag{6.4.2}$$

Dies ist äquivalent zum System

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -x_1 - 2ax_2,$$

also der linearen homogenen Differentialgleichung in \mathbb{R}^2 mit $A = \begin{bmatrix} 0 & 1 \\ -1 & -2a \end{bmatrix}$. Das charakteristische Polynom

$$p(\lambda) = \det \begin{bmatrix} 0 - \lambda & 1 \\ -1 & -2a - \lambda \end{bmatrix} = \lambda^2 + 2a\lambda + 1$$

hat die Nullstellen

$$\lambda_{1,2} = -a \pm \sqrt{a^2 - 1}.$$

Dies sind also die zugehörigen Eigenwerte.

1. Fall: $a = 0$ (keine Dämpfung): Wir haben hier die bekannte Gleichung $\ddot{x} = -x$, die Eigenwerte sind rein imaginär, man erhält die beiden Lösungen

$$\varphi(t) = \cos t \quad \text{und} \quad \psi(t) = \sin t, \quad t \in \mathbb{R},$$

und ein Fundamentalsystem ist

$$\begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}.$$

Man beachte, dass in der zweiten Zeile die Ableitungen der Einträge in der ersten Zeile stehen.

2. Fall: $0 < a < 1$ (schwache Dämpfung): Die Eigenwerte sind

$$\lambda_{1,2} = -a \pm i\sqrt{1 - a^2}.$$

Man erhält zwei Lösungen der Form

$$\varphi(t) = e^{-at} \cos(\sqrt{1-a^2}t) \quad \text{und} \quad \psi(t) = e^{-at} \sin(\sqrt{1-a^2}t), \quad t \in \mathbb{R}.$$

In der Tat, mit der Abkürzung $\sigma := \sqrt{1-a^2}$ ist

$$\begin{aligned} \dot{\varphi}(t) &= e^{-at} [-a \cos \sigma t - \sigma \sin \sigma t], \\ \ddot{\varphi}(t) &= e^{-at} [a^2 \cos \sigma t + a\sigma \sin \sigma t + a\sigma \sin \sigma t - \sigma^2 \cos \sigma t]. \end{aligned}$$

Eingesetzt gilt also

$$\begin{aligned} &e^{at} [\ddot{\varphi}(t) + 2a\dot{\varphi}(t) + \varphi(t)] \\ &= a^2 \cos \sigma t + a\sigma \sin \sigma t + a\sigma \sin \sigma t - \sigma^2 \cos \sigma t + 2a [-a \cos \sigma t - \sigma \sin \sigma t] + \cos \sigma t \\ &= a^2 \cos \sigma t + a\sigma \sin \sigma t + a\sigma \sin \sigma t - (1-a^2) \cos \sigma t - 2a^2 \cos \sigma t - 2a\sigma \sin \sigma t + \cos \sigma t \\ &= 0. \end{aligned}$$

Analog für ψ . Das zugehörige Fundamentalsystem (das wir hier der besseren Unterscheidbarkeit von der Übergangsmatrix mit der Notation für eine Wronski-Matrix mit $W(t)$ bezeichnen) ist

$$W(t) = \begin{bmatrix} e^{-at} \cos \sigma t & e^{-at} \sin \sigma t \\ e^{-at} [-a \cos \sigma t - \sigma \sin \sigma t] & e^{-at} [-a \sin \sigma t + \sigma \cos \sigma t] \end{bmatrix}$$

und

$$W(0) = \begin{bmatrix} 1 & 0 \\ -a & \sigma \end{bmatrix} \quad \text{und} \quad W(0)^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{a}{\sigma} & \frac{1}{\sigma} \end{bmatrix}.$$

Die Übergangsmatrix $\Phi(t, 0)$ (also das Fundamentalsystem mit $\Phi(0, 0) = I_2$) ist dann gegeben durch (vgl. Proposition 6.1.8(ii))

$$\Phi(t, 0) = W(t)W(0)^{-1}$$

und die Einträge der ersten Zeile sind

$$e^{-at}(\cos \sigma t + \frac{a}{\sigma} \sin \sigma t) \quad \text{und} \quad \frac{1}{\sigma} e^{-at} \sin \sigma t.$$

3. Fall: $a = 1$ (aperiodischer Grenzfall): das charakteristische Polynom hat die doppelte Nullstelle $a = 1$ und man rechnet nach, dass

$$\varphi(t) = e^{-t} \quad \text{und} \quad \psi(t) = te^{-t}$$

Lösungen von

$$\ddot{x} + 2\dot{x} + x = 0$$

sind. Das ist sofort klar für φ und

$$\ddot{\psi} + 2\dot{\psi} + \psi = -e^{-t} - (e^{-t} + te^{-t}) + 2e^{-t} - 2te^{-t} + te^{-t} = 0.$$

Die zugehörige Wronski-Matrix ist

$$W(t) = \begin{bmatrix} \varphi(t) & \psi(t) \\ \dot{\varphi}(t) & \dot{\psi}(t) \end{bmatrix} = \begin{bmatrix} e^{-t} & te^{-t} \\ -e^{-t} & (1-t)e^{-t} \end{bmatrix}, \quad W(0) = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \quad \text{und} \quad W(0)^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Wir erhalten für die erste Zeile der Übergangsmatrix $\Phi(t, 0) = W(t)W(0)^{-1}$ die Einträge

$$e^{-t} + te^{-t} \quad \text{und} \quad te^{-t}.$$

In diesem Fall ist $\lambda = -1$ ein doppelter Eigenwert mit einem zwei-dimensionalen Jordanblock; es gibt keine Schwingungen.

4. Fall: $a > 1$ (starke Dämpfung): Mit den beiden reellen Nullstellen

$$\lambda_1 = -a - \sqrt{a^2 - 1} < \lambda_2 = -a + \sqrt{a^2 - 1} < 0$$

des charakteristischen Polynoms erhalten wir zwei Lösungen

$$\varphi(t) = e^{\lambda_1 t} \text{ und } \psi(t) = e^{\lambda_2 t}.$$

Aus der zugehörigen Wronski-Matrix

$$W(t) = \begin{bmatrix} \varphi(t) & \psi(t) \\ \dot{\varphi}(t) & \dot{\psi}(t) \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t} & e^{\lambda_2 t} \\ \lambda_1 e^{\lambda_1 t} & \lambda_2 e^{\lambda_2 t} \end{bmatrix}$$

erhalten wir

$$W(0) = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} \text{ und } W(0)^{-1} = \frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix}.$$

Dies liefert für die erste Zeile der Übergangsmatrix $\Phi(t, 0) = W(t)W(0)^{-1}$ die Einträge

$$\frac{1}{\lambda_2 - \lambda_1} [\lambda_2 e^{\lambda_1 t} - \lambda_1 e^{\lambda_2 t}] \text{ und } \frac{1}{\lambda_2 - \lambda_1} [e^{\lambda_2 t} - e^{\lambda_1 t}].$$

Als zweiten Spezialfall von (6.4.1) betrachten wir die inhomogene Differentialgleichung

$$\ddot{x} + b^2 x = \cos \omega t, \quad (6.4.3)$$

die von den zwei reellen Parametern $b, \omega \geq 0$ abhängt.

Zunächst eine auch allgemein wichtige **Vorüberlegung** zur Variation der Parameter Formel bei autonomen Differentialgleichungen der Form

$$\dot{x} = Ax + g(t).$$

Hier ist ja die Übergangsmatrix

$$\Phi(t, s) = e^{A(t-s)},$$

also

$$\varphi(t; x_0) = e^{At} x_0 + \int_0^t e^{A(t-s)} g(s) ds = e^{At} \left[x_0 + \int_0^t e^{-As} g(s) ds \right].$$

Für eine beliebige Fundamentalmatrix $W(t)$ gilt wie in Proposition 6.1.8 für alle t

$$e^{At} = \Phi(t, 0) = W(t)W(0)^{-1}.$$

Oben eingesetzt liefert dies

$$\varphi(t; x_0) = W(t)W(0)^{-1} \left[x_0 + \int_0^t W(-s)W(0)^{-1} g(s) ds \right]. \quad (6.4.4)$$

Für die zu (6.4.3) gehörende homogene Gleichung

$$\ddot{x} + b^2 x = 0 \quad (6.4.5)$$

kennen wir die Lösungen

$$\varphi(t) = \cos bt, \psi(t) = \sin bt,$$

und die zugehörige Fundamentalmatrix ist

$$W(t) = \begin{bmatrix} \cos bt & \sin bt \\ -b \sin bt & b \cos bt \end{bmatrix} \text{ mit } W(t)^{-1} = \begin{bmatrix} \cos bt & -\frac{1}{b} \sin bt \\ \sin bt & \frac{1}{b} \cos bt \end{bmatrix}.$$

Beachte, dass gilt

$$W(0) = \begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix} \text{ und } W(0)^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{b} \end{bmatrix}.$$

Um die Auswirkung der Anregung auf das Schwingungsverhalten zu klären, können wir die Lösung von (6.4.3) mit Hilfe der Variation der Parameter Formel (6.4.4) berechnen. Dies führt jedoch auf etwas komplizierte Integrale. Stattdessen benutzen wir, dass sich die Lösungen der inhomogenen Gleichung als Summe der Lösungen der homogenen Gleichung und einer partikulären Lösung der inhomogenen Gleichung darstellen lassen.

Die Lösungen der homogenen Gleichung sind

$$\varphi(t; x_0, y_0) = W(t)W(0)^{-1} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} \cos bt & \sin bt \\ -b \sin bt & b \cos bt \end{bmatrix} \begin{bmatrix} x_0 \\ \frac{1}{b}y_0 \end{bmatrix} = \begin{bmatrix} x_0 \cos bt + y_0 \sin bt \\ -x_0 b \sin bt + y_0 \cos bt \end{bmatrix}.$$

Offenbar haben alle Lösungen der homogenen Gleichung die gleiche Periode $2\pi/b$. Daher nennt man b auch die zur Differentialgleichung (6.4.3) gehörende Eigenfrequenz, während die Frequenz der Anregung gleich ω ist. Für $\omega \neq b$ ist eine Lösung der inhomogenen Gleichung gegeben durch

$$\varphi(t) = \frac{1}{\omega^2 - b^2} [\cos bt - \cos \omega t], t \in \mathbb{R}.$$

Denn die erste und zweite Ableitung sind

$$\dot{\varphi}(t) = \frac{1}{\omega^2 - b^2} (-b \sin bt + \omega \sin \omega t) \text{ und } \ddot{\varphi}(t) = \frac{1}{\omega^2 - b^2} (-b^2 \cos bt + \omega^2 \cos \omega t),$$

also

$$\begin{aligned} \ddot{\varphi} + b^2 \varphi &= \frac{1}{\omega^2 - b^2} [-b^2 \cos bt + \omega^2 \cos \omega t] + \frac{b^2}{\omega^2 - b^2} [\cos bt - \cos \omega t] \\ &= -\frac{b^2}{\omega^2 - b^2} \cos bt + \frac{\omega^2}{\omega^2 - b^2} \cos \omega t + \frac{b^2}{\omega^2 - b^2} \cos bt - \frac{b^2}{\omega^2 - b^2} \cos \omega t \\ &= \frac{\omega^2}{\omega^2 - b^2} \cos \omega t - \frac{b^2}{\omega^2 - b^2} \cos \omega t = \cos \omega t \end{aligned}$$

Für $\omega \neq b$ ist die erste Komponente der allgemeinen Lösung $\lambda(t; t_0, x_0, y_0)$ mit $t_0 = 0$ daher

$$\lambda_1(t; 0, x_0, y_0) = x_0 \cos bt + y_0 \sin bt + \frac{1}{\omega^2 - b^2} [\cos bt - \cos \omega t], t \in \mathbb{R}.$$

Für $\omega = b$ ist eine Lösung der inhomogenen Gleichung (mit Anfangswert $x_0 = y_0 = 0$) gegeben durch

$$\varphi(t) := \frac{t}{2b} \sin bt, t \in \mathbb{R},$$

denn

$$\dot{\varphi}(t) = \frac{1}{2b} \sin bt + \frac{t}{2} \cos bt \text{ und } \ddot{\varphi}(t) = \cos bt - \frac{t}{2} b \sin bt,$$

also

$$\ddot{\varphi} + b^2 \varphi = \cos bt - \frac{t}{2} b \sin bt + \frac{b}{2} t \sin bt = \cos bt.$$

Die erste Komponente der allgemeinen Lösung ist daher

$$\lambda_1(t; 0, x_0, y_0) = x_0 \cos bt + y_0 \sin bt + \frac{t}{2b} \sin \omega t.$$

Ist $\omega \neq b$, so liegt eine Überlagerung von zwei Schwingungen mit unterschiedlichen Frequenzen vor. Dies wird deutlich, wenn man die trigonometrische Identität

$$\cos bt - \cos \omega t = 2 \sin \frac{\omega - b}{2} t \cdot \sin \frac{\omega + b}{2} t$$

benutzt. Dann ist für $x_0 = y_0 = 0$

$$\lambda_1(t; 0, 0, 0) = \frac{1}{\omega^2 - b^2} \sin \frac{\omega + b}{2} t \cdot \sin \frac{\omega - b}{2} t, t \in \mathbb{R}.$$

Man hat (für ω nahe b) eine langsame Schwingung (mit kleiner Frequenz) $\sin \frac{\omega - b}{2} t$ mit einer sich schnell ändernden Amplitude $\frac{1}{\omega^2 - b^2} \sin \frac{\omega + b}{2} t$. Für $\omega \rightarrow b$ geht die Frequenz $|\omega - b|$ der langsamen Schwingung gegen 0, ihre Periode $\frac{\pi}{|\omega - b|}$ geht also gegen ∞ .

Im anderen Fall $\omega = b$ tritt Resonanz auf: wird das System mit seiner Eigenfrequenz angeregt, so führt dies dazu, dass alle Bewegungen des angeregten Systems unbeschränkt anwachsen. Dies beruht auf dem Faktor t in der partikulären Lösung: $\frac{t}{2b} \sin bt$.

Kapitel 7

Stabilität Autonomer Differentialgleichungen

In diesem Kapitel werden wir für autonome Differentialgleichungen das Verhalten der Lösungen für Zeit $t \rightarrow \infty$ analysieren. Dabei werden wir (für lineare autonome Differentialgleichungen) zunächst die expliziten Lösungsformeln aus dem vorangehenden Kapitel verwenden und dann Lyapunov-Funktionen.

7.1 Stabilität und Eigenwerte

Wir beginnen mit der Definition von Stabilität von Gleichgewichten für allgemeine nichtlineare autonome Differentialgleichungen.

Betrachte

$$\dot{x} = f(x) \tag{7.1.1}$$

mit einer Lipschitz-stetigen Funktion $f : D \rightarrow \mathbb{R}^n$ mit $D \subset \mathbb{R}^n$ offen. Es sei $x^* \in D$ ein Gleichgewicht, also

$$f(x^*) = 0.$$

Ferner sei $\varphi(t; x_0)$, $t \in (t^-(x_0), t^+(x_0))$, $x_0 \in D$, der zugehörige lokale Fluss.

Erinnerung: Eine Umgebung U eines Punktes $x \in \mathbb{R}^n$ ist eine Teilmenge des \mathbb{R}^n , die eine ε -Kugel um x enthält, also x im Innern enthält.

Definition 7.1.1 (i) Ein Gleichgewicht x^* heißt stabil, falls für jede Umgebung U von x^* eine Umgebung V von x^* existiert, so dass für alle $x \in V$ die zugehörige Lösung $\varphi(t; x)$ für alle $t \geq 0$ existiert und $\varphi(t; x) \in U$ gilt. Andernfalls heißt das Gleichgewicht instabil.

(ii) Ein Gleichgewicht x^* heißt attraktiv, falls es eine Umgebung W besitzt, so dass für alle $x \in W$ gilt

$$\lim_{t \rightarrow \infty} \varphi(t; x) = x^*.$$

Dann heißt $\{x \in D \mid \lim_{t \rightarrow \infty} \varphi(t; x) = x^*\}$ Attraktionbereich.

(iii) Ein Gleichgewicht x^* heißt asymptotisch stabil, falls es stabil und attraktiv ist.

Es reicht, an Stelle allgemeiner Umgebungen Kugeln um x^* zu betrachten: Stabilität ist daher gleichwertig mit der Eigenschaft, dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle x mit $\|x - x^*\| < \delta$ die Lösung existiert und $\|\varphi(t; x) - x^*\| < \varepsilon$ für alle $t \geq 0$ erfüllt. Denn eine beliebige Umgebung U von x^* enthält eine ε -Kugel $\mathbf{B}(x^*, \varepsilon)$, $\varepsilon > 0$, also ist die δ -Kugel $V = \mathbf{B}(x^*, \delta)$ die gesuchte Umgebung. Analog bedeutet Attraktivität, dass $\varepsilon_0 > 0$ existiert, so dass $\lim_{t \rightarrow \infty} \varphi(t; x) = x^*$ für alle x mit $\|x - x^*\| < \varepsilon_0$ gilt.

Offenbar gibt es Ruhelagen, die stabil, aber nicht attraktiv sind: Betrachte die Differentialgleichung $\dot{x} = 0$. Das folgende Beispiel (aus Hahn [8]) zeigt, dass es nichtlineare autonome Differentialgleichungen mit Ruhelagen gibt, die attraktiv, aber nicht stabil sind. Daher sind Stabilität und Attraktivität voneinander unabhängige Eigenschaften.

Beispiel 7.1.2 Betrachte das in Polarkoordinaten gegebene ebene System

$$\dot{r} = r(1 - r), \quad \dot{\phi} = \sin^2 \frac{\phi}{2}.$$

Als Fluss erhält man

$$(r(t; r_0, \phi_0), \phi(t; r_0, \phi_0)) = \left(\frac{r_0}{r_0 + (1 - r_0)e^{-t}}, 2 \arctan \left(\frac{2 \sin \phi_0}{2 \cos \phi_0 - t \sin \phi_0 + 2} \right) \right), t \in \mathbb{R}.$$

Die Ruhelage $x^* = (1, 0)$ ist attraktiv, aber instabil. Dies sieht man aber am Einfachsten nicht durch diese Formel, sondern durch Analyse des Phasenporträts.

Die Klassifikation linearer autonomer Differentialgleichungen $\dot{x} = Ax$ in der Ebene liefert uns sofort eine Reihe von weiteren Beispielen. Hier ist der Ursprung immer Gleichgewicht, und wir betrachten daher die Stabilität von $x^* = 0$.

Beispiel 7.1.3 Betrachte $\dot{x} = Ax$, $A \in \mathbb{R}^{2 \times 2}$ mit A in Jordannormalform. Hat diese Differentialgleichung ein Zentrum, so ist der Ursprung stabil, aber nicht asymptotisch stabil: Hier ist die Jordannormalform mit $0 \neq b \in \mathbb{R}$ gegeben durch

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Man beachte, dass auch unter einer Koordinatentransformation diese Eigenschaften erhalten bleiben.

Hat $\dot{x} = Ax$ einen stabilen (ein-, zwei- oder viel-tangentigen) Knoten oder einen stabilen Wirbel, so ist der Ursprung asymptotisch stabil. Hat $\dot{x} = Ax$ einen instabilen (ein-, zwei- oder viel-tangentigen) Knoten oder einen instabilen Wirbel, so ist der Ursprung instabil.

Ist $A = 0$, so ist der Ursprung (und jeder andere Punkt der Ebene) stabiles, aber nicht asymptotisch stabiles Gleichgewicht.

Ist $\lambda = 0$ Eigenwert mit algebraischer Vielfachheit 2 und geometrischer Vielfachheit 1 (also nicht halbeinfach), ist die Jordannormalform von A also

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

so ist der Ursprung instabil.

Die Inspektion dieser Beispiele in der Ebene \mathbb{R}^2 zeigt, dass das Vorzeichen des Realteils der Eigenwerte eine entscheidende Rolle spielt. Wir werden den folgenden allgemeinen Satz beweisen.

Satz 7.1.4 Betrachte für $A \in \mathbb{R}^{n \times n}$ die Differentialgleichung $\dot{x} = Ax$.

(i) Der Ursprung $0 \in \mathbb{R}^n$ ist genau dann stabil, wenn alle Eigenwerte von A nichtpositiven Realteil haben und diejenigen mit Realteil 0 halbeinfach sind.

(ii) Der Ursprung $0 \in \mathbb{R}^n$ ist genau dann asymptotisch stabil, wenn alle Eigenwerte von A negativen Realteil besitzen.

Beweis. Die Stabilitätseigenschaften sind unabhängig von der Koordinatendarstellung, weil für jede invertierbare Matrix T gilt $\|y\| = \|Tx\| \leq \|T\| \|x\|$ und $\|x\| \leq \|T^{-1}\| \|y\|$. Wir können also annehmen, dass A in reeller Jordan-Normalform ist. Weil sich die Blockstruktur auf die Lösungen überträgt, reicht es dann, die Jordanblöcke J zu einem Eigenwert λ einzeln zu betrachten:

$$\dot{y} = Jy.$$

Sei zunächst $\lambda \in \mathbb{R}$ und x ein Startwert in dem entsprechenden Unterraum. Dann gilt

$$e^{Jt}x = e^{\lambda t} \begin{bmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{d-1}}{(d-1)!} \\ 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & t & \cdot \\ 0 & \cdot & \cdot & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_d \end{bmatrix} = e^{\lambda t} \begin{bmatrix} x_1 + tx_2 + \dots + \frac{t^{d-1}}{(d-1)!}x_d \\ x_2 + tx_3 + \dots + \frac{t^{d-2}}{(d-2)!}x_d \\ \cdot \\ \cdot \\ x_d \end{bmatrix}. \quad (7.1.2)$$

Aus der Analysis I ist bekannt, dass für jedes Polynom $p(t)$ und $a < 0$

$$e^{at}p(t) \rightarrow 0 \text{ für } t \rightarrow \infty.$$

Insbesondere existiert das Maximum von $\{e^{at}|p(t)|, t \geq 0\}$, denn es gibt $T_0 > 0$ so dass $|e^{at}p(t)| \leq 1$ und $|e^{at}p(t)|$ ist für $t \in [0, T_0]$ beschränkt wegen der Stetigkeit. Betrachte die Maximumnorm in \mathbb{R}^d , $\|(y_i)\| := \max_i |y_i|$ und die zugehörige Matrizen-Norm

$$\|A\| = \max_i \sum_{k=1}^d |a_{ik}|.$$

Sei jetzt $\lambda < -\kappa$ für ein $\kappa > 0$. Dann folgt wegen $\lambda + \kappa < 0$, dass

$$e^{\kappa t} \|e^{Jt}\| = e^{(\kappa+\lambda)t} \max_i \sum_{k=0}^{i-1} \frac{t^k}{k!} \leq K \text{ für } t \geq 0$$

mit einer Konstanten $K > 0$ und daher

$$\|e^{Jt}x_0\| \leq \|e^{Jt}\| \|x_0\| \leq K e^{-\kappa t} \|x_0\| \text{ für alle } t \geq 0. \quad (7.1.3)$$

Aus dieser Formel liest man ab, dass Stabilität und Attraktivität gelten, also asymptotische Stabilität.

Solch eine Abschätzung gilt für jeden reellen Eigenwert kleiner als 0. Analog (mit etwas mehr Schreibarbeit) sieht man, dass dies auch für komplex-konjugierte Paare von Eigenwerten mit negativem Realteil gilt. Daher gilt für alle Matrizen A in reeller Jordan-Normalform, für die das Maximum der Realteile der Eigenwerte kleiner als $-\kappa < 0$ ist, dass für eine Konstante $K > 0$

$$\|e^{At}x_0\| \leq K e^{-\kappa t} \|x_0\| \text{ für alle } t \geq 0. \quad (7.1.4)$$

Aus dieser Formel liest man unmittelbar die Stabilität und asymptotische Stabilität ab. Damit haben wir gezeigt, dass asymptotische Stabilität und insbesondere Stabilität folgt, falls das Maximum der Realteile der Eigenwerte negativ ist. Offenbar impliziert die Existenz eines Eigenwertes mit positivem Realteil, dass Instabilität gilt.

Es bleibt nur noch das Stabilitätsverhalten für Eigenwerte mit Realteil 0 zu diskutieren. Sei jetzt $\lambda = 0$ ein Eigenwert mit zugehörigem Jordanblock J gegeben durch (7.1.2) mit $\lambda = 0$. Wenn die Dimension des Jordanblocks gleich 1 ist, so folgt

$$\|e^{Jt}x_0\| = e^{0t} \|x_0\| = \|x_0\| \text{ für alle } t \geq 0,$$

und die Stabilität ist unmittelbar ablesbar. Attraktivität gilt offenbar nicht. Analoges gilt für komplex-konjugierte Eigenwertpaare $\pm\sigma i$, $\sigma \in \mathbb{R}$. Es ist also nur noch zu zeigen, dass die Existenz eines Jordanblocks mit Dimension > 1 die Instabilität impliziert. Betrachte wieder $\lambda = 0$, also

$$e^{Jt}x = \begin{bmatrix} x_1 + tx_2 + \frac{t^2}{2!}x_3 + \dots + \frac{t^{d-1}}{(d-1)!}x_d \\ x_2 + tx_3 + \dots + \frac{t^{d-2}}{(d-2)!}x_d \\ \cdot \\ \cdot \\ x_d \end{bmatrix}.$$

Für jeden Anfangswert $x_1 = 0 = x_3 = \dots = x_d$, $x_2 = \delta$ mit $\delta > 0$ erhält man als Lösung $\delta(t, 1, 0, \dots, 0)^T$, also eine unbeschränkte Lösung und Instabilität folgt. Analog argumentiert man wieder für komplex-konjugierte Eigenwertpaare. ■

Die Eigenwerte und die zugehörige Struktur der Jordanblöcke legen also die Stabilitätseigenschaften einer autonomen, linearen Differentialgleichung fest.

Bemerkung 7.1.5 (i) Matrizen, für die alle Eigenwerte negativen Realteil haben, nennt man auch Hurwitz-Matrizen.

(ii) Der Beweis von Satz 7.1.4 zeigt auch die folgende stärkere Aussage: Haben alle Eigenwerte von A negativen Realteil, so folgt neben der Stabilität, dass für alle Anfangswerte $x_0 \in \mathbb{R}^n$ die Lösung $\varphi(t; x_0)$ gegen 0 konvergiert. Diese Eigenschaft heißt globale asymptotische Stabilität (im Unterschied dazu nennt man die Eigenschaft in Definition 7.1.1 auch lokale asymptotische Stabilität).

(iii) Wegen der Beziehung (7.1.4) gilt für autonome lineare Differentialgleichungen, dass die asymptotische Stabilität die folgende exponentielle Stabilitätseigenschaft impliziert: für alle asymptotisch stabilen Matrizen A gibt es Konstanten $K > 0$ und $\kappa > 0$ (hierbei müssen die Realteile aller Eigenwerte von A kleiner als $-\kappa$ sein), so dass für alle $x \in \mathbb{R}^n$

$$\|e^{At}x\| \leq K e^{-\kappa t} \|x\| \text{ für alle } t \geq 0.$$

Für autonome lineare Differentialgleichungen haben wir insbesondere gezeigt: Für alle $x_0 \in \mathbb{R}^n$ konvergiert die zugehörige Lösung gegen 0 für $t \rightarrow \infty$, wenn alle Eigenwerte von A negativen Realteil haben. Das folgende Beispiel zeigt, dass dies für nicht-autonome lineare Differentialgleichungen

$$\dot{x} = A(t)x$$

mit einer stetigen Matrixfunktion $A: \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ falsch ist. Hier haben für alle $t \in \mathbb{R}$ die Matrizen $A(t)$ nur Eigenwerte mit negativem Realteil, aber es gibt unbeschränkte Lösungen.

Beispiel 7.1.6 Betrachte die folgende nichtautonome lineare Differentialgleichung in der Ebene \mathbb{R}^2 :

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = A(t) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 + \frac{3}{2} \cos^2 t & 1 - \frac{3}{2} \cos t \sin t \\ -1 - \frac{3}{2} \cos t \sin t & -1 + \frac{3}{2} \sin^2 t \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Die Eigenwerte sind für jedes $t \in \mathbb{R}$ gegeben durch

$$\lambda_{1,2} = (-1 \pm i\sqrt{7})/4,$$

denn

$$\begin{aligned} \det [A(t) - \lambda I] &= \det \begin{bmatrix} -1 + \frac{3}{2} \cos^2 t - \lambda & 1 - \frac{3}{2} \cos t \sin t \\ -1 - \frac{3}{2} \cos t \sin t & -1 + \frac{3}{2} \sin^2 t - \lambda \end{bmatrix} \\ &= \left(-1 + \frac{3}{2} \cos^2 t - \lambda\right) \left(-1 + \frac{3}{2} \sin^2 t - \lambda\right) - \left(-1 + \frac{9}{4} \cos^2 t \sin^2 t\right) \\ &= \lambda^2 + \lambda \left(1 - \frac{3}{2} \cos^2 t + 1 - \frac{3}{2} \sin^2 t\right) + \left(1 - \frac{3}{2} \cos^2 t\right) \left(1 - \frac{3}{2} \sin^2 t\right) + 1 - \frac{9}{4} \cos^2 t \sin^2 t \\ &= \lambda^2 + \lambda \left[2 - \frac{3}{2} (\cos^2 t + \sin^2 t)\right] + \left[1 - \frac{3}{2} (\sin^2 t + \cos^2 t) + \frac{9}{4} \cos^2 t \sin^2 t\right] + 1 - \frac{9}{4} \cos^2 t \sin^2 t \\ &= \lambda^2 + \frac{1}{2} \lambda + \frac{1}{2}. \end{aligned}$$

Also sind die Eigenwerte

$$\lambda_{1,2} = -\frac{1}{4} \pm \sqrt{\frac{1}{16} - \frac{8}{16}} = (-1 \pm i\sqrt{7})/4.$$

Andererseits ist die Lösung für den Anfangswert

$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} -\varepsilon \\ 0 \end{bmatrix}, \varepsilon \neq 0.$$

gegeben durch

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \varepsilon \begin{bmatrix} -\cos t \\ \sin t \end{bmatrix} e^{t/2}, \quad (7.1.5)$$

denn man rechnet nach

$$\frac{1}{\varepsilon} \begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} \sin t \\ \cos t \end{bmatrix} e^{t/2} + \frac{1}{2} \begin{bmatrix} -\cos t \\ \sin t \end{bmatrix} e^{t/2} = \begin{bmatrix} \sin t - \frac{1}{2} \cos t \\ \cos t + \frac{1}{2} \sin t \end{bmatrix} e^{t/2}$$

und

$$\begin{aligned} & \begin{bmatrix} -1 + \frac{3}{2} \cos^2 t & 1 - \frac{3}{2} \cos t \sin t \\ -1 - \frac{3}{2} \cos t \sin t & -1 + \frac{3}{2} \sin^2 t \end{bmatrix} \begin{bmatrix} -\cos t \\ \sin t \end{bmatrix} e^{t/2} \\ &= \begin{bmatrix} \cos t - \frac{3}{2} \cos t \cos^2 t + \sin t - \frac{3}{2} \cos t \sin^2 t \\ \cos t + \frac{3}{2} \cos^2 t \sin t - \sin t + \frac{3}{2} \sin^2 t \sin t \end{bmatrix} e^{t/2} \\ &= \begin{bmatrix} -\frac{1}{2} \cos t + \sin t \\ \cos t + \frac{1}{2} \sin t \end{bmatrix} e^{t/2}. \end{aligned}$$

Die Lösung (7.1.5) wird offenbar unbeschränkt für $t \rightarrow \infty$.

Für Stabilitätseigenschaften nichtautonomer linearer Differentialgleichungen kann man also im Allgemeinen nicht mit den Eigenwerten der Matrizen für festes t arbeiten. Dieses Problem war (für periodische Matrixfunktionen $A(\cdot)$) schon Floquet Ende des 19. Jahrhunderts bekannt und ist etwas später von Lyapunov für allgemeine Matrixfunktionen näher untersucht worden. Dies führte ihn zur Definition der heute Lyapunov-Exponenten genannten Größen, die in den letzten Jahrzehnten intensiv untersucht worden sind und zu vielen neuen Entwicklungen in der Theorie dynamischer Systeme geführt haben.

Die genaue Analyse von Stabilität und asymptotischer Stabilität für nichtautonome Differentialgleichungen führt auf unterschiedliche Begriffe, je nachdem, ob Gleichmäßigkeit bezüglich t gefordert wird oder nicht.

7.2 Lyapunov-Funktionen

Ist die Dimension der Matrix A größer als drei, wird die Berechnung der Eigenwerte per Hand ausgesprochen unangenehm. Für höherdimensionale Systeme wird daher meist ein anderer Weg (Lyapunovs direkte Methode) beschritten, der ebenfalls auf A.M. Lyapunov (1892) zurückgeht und vor allem auch für viele nichtlineare Systeme gangbar ist.

Wir betrachten wieder ein Gleichgewicht x^* für das System (7.1.1). Wir können annehmen, dass $x^* = 0$ gilt. Denn für die Differentialgleichung

$$\dot{y} = f(y + x^*).$$

ist $0 \in \mathbb{R}^n$ offenbar ein Gleichgewicht, und es ist nicht schwer zu sehen, dass dies Gleichgewicht dieselben Stabilitätseigenschaften wie x^* hat.

Wir werden eine Funktion V , Lyapunov-Funktion genannt, einführen, die u.a. die folgende Eigenschaft besitzt.

Definition 7.2.1 Sei $V : U \rightarrow [0, \infty)$ eine C^1 -Funktion, die auf einer Umgebung $U \subset D$ von $0 \in \mathbb{R}^n$ definiert ist. Sie heißt positiv definit, falls

$$V(0) = 0 \text{ und } V(x) > 0 \text{ für } x \neq 0.$$

Zum Beispiel ist $V(x_1, x_2) = x_1^2 + x_2^2$ positiv-definit auf \mathbb{R}^2 ; die Funktion

$$V(x_1, x_2) = x_1^2 + x_2^2 - x_2^3$$

ist positiv definit auf einem Streifen um die x_1 -Achse.

Wir wollen garantieren, dass die Funktion V entlang von Lösungen der Differentialgleichung abnimmt, ohne explizit diese Lösungen zu verwenden. Die folgende Definition ist nützlich

$$\dot{V}(x) := \frac{\partial V(x)}{\partial x} \cdot f(x), \quad x \in U.$$

Man beachte, dass $\dot{V}(x)$ keine Funktion von t ist. Insbesondere muss man nur das Vektorfeld $f(x)$ kennen, aber keine Lösung der Differentialgleichung, um $\dot{V}(x)$ zu berechnen.

Definition 7.2.2 Eine positiv-definite C^1 -Funktion $V : U \rightarrow [0, \infty)$ heißt *Lyapunov-Funktion*, falls

$$\dot{V}(x) \leq 0 \text{ für alle } x \in U.$$

Sie heißt *strikte Lyapunov-Funktion*, falls

$$\dot{V}(x) < 0 \text{ für alle } x \in U, \quad x \neq 0.$$

Es gilt die folgende hinreichende Bedingung für Stabilität.

Satz 7.2.3 Es sei $0 \in \mathbb{R}^n$ ein Gleichgewicht für die Differentialgleichung (7.1.1).

(i) Existiert eine Lyapunov-Funktion V , so ist 0 stabil.

(ii) Existiert eine strikte Lyapunov-Funktion, so ist 0 asymptotisch stabil.

Beweis. (i) Sei $V : U \rightarrow [0, \infty)$ eine positiv-definite C^1 -Funktion mit $\dot{V}(x) \leq 0$ für alle $x \in U$. Betrachte $\varepsilon > 0$ mit $\mathbf{B}(0; 2\varepsilon) \subset U$. Definiere

$$m := \min_{\|x\|=\varepsilon} V(x) > 0.$$

Sei δ mit $0 < \delta < \varepsilon$ so, dass $V(x) < m$ für $\|x\| \leq \delta$. Dies existiert, weil V stetig ist mit $V(0) = 0$. Wähle einen Anfangswert x_0 mit $\|x_0\| < \delta$. Dann gilt, solange die Trajektorie in U bleibt, für $t \geq 0$ nach der Kettenregel

$$\frac{d}{dt} V(\varphi(t; x_0)) = \frac{\partial V(x)}{\partial x} \Big|_{x=\varphi(t; x_0)} \cdot f(\varphi(t; x_0)) = \dot{V}(\varphi(t; x_0)) \leq 0.$$

Durch Integration folgt aus dem Hauptsatz der Differential- und Integralrechnung

$$V(\varphi(t; x_0)) - V(x_0) \leq 0,$$

also

$$V(\varphi(t; x_0)) \leq V(x_0) < m. \tag{7.2.1}$$

Es folgt, dass $\|\varphi(t; x_0)\| < \varepsilon$ für alle $t \geq 0$ aus dem maximalen Existenzintervall; nach Satz 2.4.10 existiert dann die Lösung auf $[0, \infty)$. Die Abschätzung gilt, weil es andernfalls nach dem Zwischenwertsatz eine Zeit $s \in (0, t)$ gibt mit

$$\|\varphi(s; x_0)\| = \varepsilon.$$

Wählt man s als die minimale Zeit mit dieser Eigenschaft, so ist $\varphi(\tau; x_0) \in U$ für alle $\tau \in [0, s]$. Nach Definition von m ist dann aber $V(\varphi(s; x_0)) \geq m$, im Widerspruch zu (7.2.1). Damit ist die Stabilität nachgewiesen.

(ii) Da eine strikte Lyapunov-Funktion insbesondere eine Lyapunov-Funktion ist, folgt die Stabilität aus (i). Es ist also nur die Attraktivität zu zeigen. Wir verwenden dazu das folgende Lemma. ■

Lemma 7.2.4 *Es gibt $\delta_0 > 0$, so dass für alle $\|x_0\| < \delta_0$ gilt: Aus $x_n \rightarrow x_0$ folgt $\varphi(1; x_n) \rightarrow \varphi(1; x_0)$.*

Beweis. Sei L eine Lipschitz-Konstante in einer Umgebung von 0, also für alle x, y mit $\|x\|, \|y\| < \varepsilon$ mit $\varepsilon > 0$ gelte

$$\|f(x) - f(y)\| \leq L \|x - y\|.$$

Nach Teil (i) des Beweis ist 0 ein stabiles Gleichgewicht. Daher gibt es $\delta_0 > 0$ so dass für alle x mit $\|x\| < \delta_0$ gilt $\|\varphi(t; x)\| < \varepsilon$ für alle $t \geq 0$. Sei $\|x_0\| < \delta_0$.

Für n groß genug ist auch $\|x_n\| < \delta_0$. Dann folgt für alle $t \in [0, 1]$

$$\begin{aligned} \|\varphi(t; x_n) - \varphi(t; x_0)\| &\leq \|x_n - x_0\| + \left\| \int_0^t f(\varphi(s; x_n)) - f(\varphi(s; x_0)) ds \right\| \\ &\leq \|x_n - x_0\| + \int_0^t L \|\varphi(s; x_n) - \varphi(s; x_0)\| ds, \end{aligned}$$

also folgt mit der Gronwall-Ungleichung für jedes $t \in [0, 1]$

$$\|\varphi(t; x_n) - \varphi(t; x_0)\| \leq \|x_n - x_0\| e^{Lt} \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Damit ist das Lemma bewiesen.

Wir beginnen jetzt den **Beweis von (ii)**. Es sei $\varepsilon > 0$ so klein, dass $\mathbf{B}(0, 2\varepsilon) \subset U$. Dann existiert wegen der Stabilität ein $\delta > 0$, so dass für alle x_0 mit $\|x_0\| < \delta$ folgt, dass $\|\varphi(t; x_0)\| < \varepsilon$ für alle $t \geq 0$. Wir können annehmen, dass $\delta \leq \delta_0$ aus Lemma 7.2.4 ist. Es reicht zu zeigen, dass für diese x_0 die Trajektorie für $t \rightarrow \infty$ gegen 0 konvergiert. Gilt dies nicht, so existieren $t_k \rightarrow \infty$ mit $\varphi(t_k; x_0) \not\rightarrow 0$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $\varphi(t_k; x_0) \rightarrow x_1 \neq 0$ mit $\|x_1\| \leq \varepsilon$ und $t_{k+1} \geq t_k + 1$. Ferner ist $\varphi(t; x_0) \in U$ für alle $t \geq 0$.

Ähnlich wie in (i) folgt aus $\dot{V} < 0$, dass

$$V(\varphi(1; x)) - V(x) = \int_0^1 \frac{d}{ds} V(\varphi(s; x_1)) ds < 0,$$

also

$$V(\varphi(1; x)) < V(x). \quad (7.2.2)$$

Andererseits gilt wegen der Stetigkeit von V , dass $V(\varphi(t_{k+1}; x_0)) \rightarrow V(x_1)$ und nach Lemma 7.2.4 ist unter Verwendung der Fluss-Eigenschaft

$$\varphi(1 + t_k; x_0) = \varphi(1; \varphi(t_k; x_0)) \rightarrow \varphi(1; x_0),$$

also

$$V(\varphi(t_{k+1}; x_0)) \leq V(\varphi(1 + t_k; x_0)) \rightarrow V(\varphi(1; x_0)).$$

Wir erhalten im Widerspruch zu (7.2.2).

$$V(x_1) \leq V(\varphi(1; x_0)).$$

■

Der Beweis zeigt, dass eine strikte Lyapunov-Funktion verhindert, dass das System in positiver Distanz vom Gleichgewicht bleibt.

Bemerkung 7.2.5 *Die Methode, Lyapunov-Funktionen zum Nachweis von Stabilität einzusetzen, heißt auch die direkte Methode von Lyapunov, weil man sie anwenden kann, ohne Lösungen zu berechnen.*

Hier wenden wir uns jetzt der Frage zu, wie man für autonome lineare Differentialgleichungen Lyapunov-Funktionen konstruiert. Wir erinnern zunächst daran, dass eine symmetrische Matrix $B \in \mathbb{R}^{n \times n}$ positiv-definit ist, falls $x^T B x > 0$ für alle $0 \neq x \in \mathbb{R}^n$. Dies ist äquivalent dazu, dass $-B$ negativ definit ist. (Im weiteren steht positiv-definit für symmetrisch und positiv-definit).

Die Matrix B ist offenbar genau dann positiv definit, wenn die C^1 -Funktion

$$V(x) = x^T B x$$

eine positiv-definite Funktion auf $U = \mathbb{R}^n$ ist. Um eine Lyapunov-Funktion zu konstruieren, suchen wir daher nach einer positiv definiten Matrix. Der folgende Satz ist das Hauptresultat zur asymptotischen Stabilität von linearen autonomen Differentialgleichungen.

Satz 7.2.6 Für $A \in \mathbb{R}^{n \times n}$ sind äquivalent:

- (i) Die Differentialgleichung $\dot{x} = Ax$ ist asymptotisch stabil.
- (ii) A hat nur Eigenwerte mit negativem Realteil.
- (iii) Es gibt eine positiv-definite Matrix $C \in \mathbb{R}^{n \times n}$, so dass die Lyapunov-Gleichung

$$A^T B + B A = -C \tag{7.2.3}$$

eine positiv-definite Lösung $B \in \mathbb{R}^{n \times n}$ hat.

(iv) Für jede positiv-definite Matrix $C \in \mathbb{R}^{n \times n}$ hat die Lyapunov-Gleichung (7.2.3) eine positiv-definite Lösung $B \in \mathbb{R}^{n \times n}$.

Beweis. Die Äquivalenz von (i) und (ii) ist in Theorem 7.1.4 gezeigt und trivialerweise impliziert (iv) die Bedingung (iii).

Die Funktion $V(x) = x^T B x = \langle x, Bx \rangle$ ist positiv definit und

$$\frac{\partial V(x)}{\partial x} y = \langle y, Bx \rangle + \langle x, By \rangle, \quad y \in \mathbb{R}^n.$$

Also ergibt sich für die Differentialgleichung $\dot{x} = Ax$

$$\dot{V}(x) = \frac{\partial V(x)}{\partial x} Ax = \langle Ax, Bx \rangle + \langle x, B Ax \rangle = x^T A^T B x + x^T B A x = x^T [A^T B + B A] x.$$

Gilt also (iii), ist B also eine positiv-definite Lösung von (7.2.3) für eine negativ-definite Matrix $-C$, so ist \dot{V} negativ-definit, und daher V eine strikte Lyapunov-Funktion. Nach Theorem 7.2.3 folgt, dass $\dot{x} = Ax$ asymptotisch stabil ist. (iii) impliziert daher (i).

Gelte jetzt (ii), also alle Realteile von Eigenwerten seien negativ. Ferner sei C eine positiv-definite Matrix. Definiere

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

Die Matrix B ist wohldefiniert: nach (7.1.4) existieren Konstanten $K, \kappa > 0$, so dass

$$\|e^{A t}\| = \sup_{\|x\|=1} \|e^{A t} x\| \leq \sup_{\|x\|=1} K e^{-\kappa t} \|x\| = K e^{-\kappa t}, \quad t \geq 0; \tag{7.2.4}$$

dann ist

$$\left\| e^{A^T t} C e^{A t} \right\| \leq K^2 \|C\| e^{-2\kappa t}, \quad t \geq 0,$$

weil A und A^T die gleichen Eigenwerte haben und daher beide eine Abschätzung der Form (7.2.4) erfüllen. Daher ist diese Funktion integrierbar auf $[0, \infty)$ und dann existiert auch das Integral über die Matrizenfunktion. Ferner ist B symmetrisch und positiv-definit, denn

$$0 = x^T B x = x^T \int_0^\infty e^{A^T t} C e^{A t} dt x = \int_0^\infty x^T e^{A^T t} C e^{A t} x dt$$

impliziert, weil der Integrand nicht-negativ ist, dass $0 = x^T e^{A^T t} C e^{At} x = (e^{At} x)^T C (e^{At} x)$ für alle $t \geq 0$, was für $t = 0$ mit der Positiv-Definitheit von C die Gleichung $x = 0$ impliziert. Dann ist B Lösung von (7.2.3), weil nach der Produktregel

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

Daher folgt (iv) und der Satz ist vollständig bewiesen. ■

Bemerkung 7.2.7 Die einfachste Lyapunov-Gleichung erhält man für $C = I$, also

$$A^T B + BA = -I.$$

Wir werden jetzt Lyapunov-Funktionen anwenden, um eine fundamentale Aussage über nicht-lineare Differentialgleichungen zu beweisen.

Betrachte

$$\dot{x} = f(x)$$

mit einer C^1 -Funktion f und Ruhelage $x^* \in D$. Anstelle der nichtlinearen Funktion verwenden wir die Linearisierung in x^* , also der Jacobi-Matrix $Df(x^*) = \frac{\partial f(x)}{\partial x} \Big|_{x=x^*}$ und erhalten die autonome lineare Differentialgleichung

$$\dot{y} = Df(x^*)y.$$

Da die Stabilität dieser Differentialgleichung wesentlich einfacher zu bestimmen ist, möchten wir auf die Stabilität der Ruhelage der nichtlinearen Differentialgleichung zurückschließen. Solch eine Argumentation haben wir schon im einleitenden Kapitel bei der Diskussion des Pendels verwendet.

Satz 7.2.8 Für die Differentialgleichung (7.1.1) sei $f : D \rightarrow \mathbb{R}^n$ eine C^1 -Funktion, $x^* \in D$, und $f(x^*) = 0$. Ist $0 \in \mathbb{R}^n$ ein asymptotisch stabiles Gleichgewicht für die linearisierte Differentialgleichung

$$\dot{y} = Df(x^*)y \tag{7.2.5}$$

so ist $x^* \in \mathbb{R}^n$ ein asymptotisch stabiles Gleichgewicht für (7.1.1).

Beweis. Sei $A := Df(x^*)$. Nach Satz 7.2.6 existiert für (7.2.5) eine strikte Lyapunov-Funktion auf \mathbb{R}^n in der Form

$$V(y) = y^T B y,$$

wobei $A^T B + BA = -I$ gilt. Wir zeigen, dass

$$\hat{V}(x) = (x - x^*)^T B (x - x^*).$$

für (7.1.1) eine Lyapunov-Funktion auf einer Umgebung von x^* ist. Dafür ist nur zu zeigen, dass $\frac{\partial \hat{V}}{\partial x}(x) \cdot f(x) < 0$ gilt für $x \neq x^*$ in einer Umgebung von x^* . Nach Definition der Differenzierbarkeit gilt

$$f(x) = Df(x^*)(x - x^*) + h(x - x^*) \quad \text{mit} \quad \lim_{x \rightarrow x^*} \frac{h(x - x^*)}{\|x - x^*\|} = 0.$$

Für (7.1.1) berechnet man

$$\begin{aligned} \frac{\partial \hat{V}}{\partial x}(x) f(x) &= \langle f(x), B(x - x^*) \rangle + \langle x - x^*, Bf(x) \rangle \\ &= \langle A(x - x^*) + h(x - x^*), B(x - x^*) \rangle + \langle x - x^*, BA(x - x^*) + Bh(x - x^*) \rangle \\ &= (x - x^*)^T A^T B (x - x^*) + h(x - x^*)^T B (x - x^*) + (x - x^*)^T BA (x - x^*) + (x - x^*)^T Bh (x - x^*) \\ &= (x - x^*)^T [A^T B + BA] (x - x^*) + 2h(x - x^*)^T B (x - x^*) \\ &= -\|x - x^*\|^2 + 2h(x - x^*)^T B (x - x^*). \end{aligned}$$

Nun ist

$$|h(x-x^*)^T B(x-x^*)| \leq \|h(x-x^*)\| \|B\| \|x-x^*\| = \frac{\|h((x-x^*))\|}{\|x-x^*\|} \|B\| \|x-x^*\|^2;$$

hier geht der erste Faktor gegen 0 für $\|x-x^*\| \rightarrow 0$. Es folgt also $\frac{\partial \hat{V}(x)}{\partial x} f(x) < 0$ für $\|x-x^*\| < \alpha$ mit $\alpha > 0$ klein genug. Auf der α -Kugel $B(x^*, \alpha)$ um x^* ist \hat{V} also eine strikte Lyapunov-Funktion für (7.1.1). ■

Bemerkung 7.2.9 Aus der Existenz einer Lyapunov-Funktion folgt auch, dass die Lösungen für Anfangswerte nahe des Gleichgewichts auf $[0, \infty)$ existieren. Denn andernfalls werden sie auf einem endlichen Existenzintervall unbeschränkt.

Im Allgemeinen kann man von Eigenschaften der linearisierten Gleichung nur auf lokale Eigenschaften der nichtlinearen Gleichung schließen. Die Umkehrung von Theorem 7.2.8 gilt jedoch nicht, wie das folgende Beispiel zeigt.

Beispiel 7.2.10 Betrachte die nichtlineare eindimensionale Differentialgleichung

$$\dot{x} = -x^3.$$

Durch Nachrechnen sieht man leicht, dass die Lösungen dieser Gleichung gegeben sind durch

$$\varphi(t; x_0) = \begin{cases} 0 & \text{falls } x_0 = 0 \\ (2t + 1/x_0^2)^{-1/2} & \text{falls } x_0 > 0 \\ -(2t + 1/x_0^2)^{-1/2} & \text{falls } x_0 < 0 \end{cases}.$$

Daher ist das Gleichgewicht $x^* = 0$ asymptotisch stabil (sogar global asymptotisch stabil). Die Linearisierung dieser Gleichung in diesem Gleichgewicht ist durch

$$\dot{y} = 0$$

gegeben und offenbar ist $y^* = 0$ für diese Gleichung zwar stabil, jedoch nicht asymptotisch stabil.

Bemerkung 7.2.11 Für autonome lineare Differentialgleichungen folgt aus der asymptotischen Stabilität schon exponentielle Abschätzung (7.1.4). Daher kann man vermuten, dass exponentielle Stabilität einer Ruhelage einer nichtlinearen Differentialgleichung die asymptotische (also exponentielle) Stabilität der linearisierten Differentialgleichung impliziert. Dies ist in der Tat richtig (Grüne/Junge [7, Satz 9.8]).

Wir kehren jetzt zu Lyapunov-Funktionen bei allgemeinen nichtlinearen Differentialgleichungen zurück.

Definition 7.2.12 Es sei $x^* \in \mathbb{R}^n$ ein asymptotisch stabiles Gleichgewicht der Differentialgleichung (7.1.1). Der Einzugsbereich (oder Attraktionsbereich) $A(x^*)$ von x^* ist

$$A(x^*) := \{x_0 \in \mathbb{R}^n \mid \varphi(t; x_0) \rightarrow x^* \text{ für } t \rightarrow \infty\}.$$

Eine Folgerung aus dem Beweis von Satz 7.2.3 ist das folgende Korollar.

Korollar 7.2.13 Es sei $V : U \rightarrow [0, \infty)$ strikte Lyapunov-Funktion für das Gleichgewicht x^* und $\gamma > 0$. Die Subniveaumenge

$$N_\gamma := \{x \in U \mid V(x) < \gamma\}$$

ist enthalten im Attraktionsbereich, also $N_\gamma \subset A(x^*)$.

Beweis. Die Subniveaumenge N_γ ist positiv invariant, d.h., $x \in N_\gamma$ impliziert $\varphi(t; x) \in N_\gamma$ für alle $t \geq 0$, weil V längs Trajektorien abnimmt. Dann folgt die Behauptung unmittelbar aus dem Beweis von Satz 7.2.3. ■

Bemerkung 7.2.14 Manchmal kann man nur Funktionen V finden, für die überall $\dot{V}(x) < 0$ gilt sowie $V(x^*) = 0$, aber $V(x) > 0, x \neq x^*$, nur auf einer Teilmenge einer Subniveaumenge N_γ erfüllt ist. Ist diese Teilmenge positiv invariant und enthält sie das Gleichgewicht, so ist sie im Einzugsbereich des Gleichgewichts enthalten.

7.3 Lyapunov-Funktionen für das Pendel

Es gibt keine allgemeine Konstruktion von Lyapunov-Funktionen für stabile oder asymptotisch stabile Gleichgewichte. Für viele konkrete Systeme kann man aber (insbesondere physikalisch motivierte) Lyapunov-Funktionen finden.

Betrachte eine Pendelgleichung der Form

$$\ddot{x} + g(\dot{x}) + \sin x = 0 \quad (7.3.1)$$

mit einer C^1 -Funktion $g: \mathbb{R} \rightarrow \mathbb{R}$, für die $g(0) = 0$ und $yg(y) \geq 0$ gilt (zum Beispiel $g(y) \equiv 0$ oder $g(y) = y$ oder $g(y) = y^3, y \in \mathbb{R}$). Für das äquivalente System erster Ordnung

$$\dot{x} = y, \dot{y} = -g(y) - \sin x$$

sind die Ruhelagen gegeben durch die Punkte $(k\pi, 0), k \in \mathbb{Z}$. Ein Kandidat für eine Lyapunov-Funktion ergibt sich aus Summe von kinetischer Energie $\frac{y^2}{2}$ und potentieller Energie, also das Integral über die Kraft, die nötig ist, um aus der Position 0 in die Position x zu kommen:

$$V(x, y) := \frac{y^2}{2} + \int_0^x \sin z \, dz = \frac{y^2}{2} - \cos x + 1$$

Dies ist eine Lyapunov-Funktion für die Ruhelage $(0, 0)$, daher ist der Nullpunkt ein stabiles Gleichgewicht. In der Tat:

$$V(0, 0) = 0, \quad V(x, y) > 0 \text{ für alle } (x, y) \in (-\pi, \pi) \times \mathbb{R}, (x, y) \neq (0, 0),$$

und mit $f(x, y) = (y, -g(y) - \sin x)^T$ ist

$$\dot{V}(x, y) = \frac{\partial V}{\partial x}(x, y)f(x, y) = (\sin x, y) \begin{pmatrix} y \\ -g(y) - \sin x \end{pmatrix} = y \sin x - yg(y) - y \sin x = -yg(y) \leq 0.$$

Im Folgenden diskutieren wir einen Spezialfall der nichtlinearen Pendelgleichung mit Reibung (vgl. Abschnitt 1.1):

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} y \\ -\sin x - ky \end{bmatrix} \quad (7.3.2)$$

mit einer Konstanten $k > 0$. Wir betrachten die beiden Gleichgewichte

$$x_0^* = (0, 0)^T \text{ und } x_1^* = (\pi, 0)^T.$$

Physikalisch ist x_0^* gerade der Zustand des nach unten hängenden Pendels, während x_1^* der Zustand des senkrecht nach oben stehenden Pendels ist (offenbar sind auch alle Punkte $x_n^* = (n\pi, 0)^T, n \in \mathbb{Z}$, Gleichgewichte). Stellen wir uns vor, dass wir das Pendel zum Anfangszeitpunkt $t_0 = 0$ durch einen kleinen Stoß aus der Ruhelage x_0^* bringen, so werden wir erwarten, dass die Lösung nahe bei dem Gleichgewicht bleibt und, wegen des Reibungsterms $-ky$, für $t \rightarrow \infty$ gegen x_0^* läuft, also asymptotische Stabilität gilt. Für den Beweis dieser Aussage haben wir zwei Möglichkeiten zur Verfügung: Wir können eine Lyapunov-Funktion konstruieren oder mit Hilfe des Linearisierungssatzes, Satz 7.2.8, und unserer Analyse linearer autonomer Differentialgleichungen argumentieren. Mit dem Linearisierungssatz erhält man, dass die Lösungen für Anfangswerte nahe genug am Gleichgewicht dagegen laufen. Der Satz liefert aber keine Information darüber, wie groß diese Umgebung ist. Im Gegensatz dazu erlaubt die Konstruktion einer Lyapunov-Funktion Aussagen darüber, vgl. Bemerkung 7.2.14.

Um Information über den Einzugsbereich des Ursprungs für das Pendel zu bekommen, konstruieren wir eine strikte Lyapunov-Funktion. Oben hat wir die Lyapunov-Funktion

$$V(x, y) = \frac{1}{2}y^2 + 1 - \cos x$$

betrachtet mit

$$\dot{V}(x, y) = DV(x, y)f(x, y) = [\sin x, y] \begin{bmatrix} y \\ -\sin x - ky \end{bmatrix} = -ky^2.$$

Daher ist $\dot{V}(x, y) \leq 0$ für alle $x \in \mathbb{R}^2$, aber $\dot{V}(x, y) < 0$ nur für alle (x, y) mit $y \neq 0$. Wir können daher nicht auf die asymptotische Stabilität des Ursprungs schließen.

Mit etwas Probieren kommt man auf die modifizierte Lyapunov-Funktion

$$V_\alpha(x, y) = \frac{1}{2}y^2 + 1 - \cos x + \alpha y \sin x$$

für einen Parameter $\alpha > 0$. Dies ist ebenfalls eine C^1 -Funktion. Um die Positiv-Definitheit nachzuweisen, beachte, dass für $x \in (-\pi, \pi)$ gilt

$$|\alpha y \sin x| \leq |\alpha y x| \leq \frac{\alpha}{2}(x^2 + y^2). \quad (7.3.3)$$

Daher ist für $\alpha < \frac{4}{\pi^2}$

$$V(x, y) > 0 \text{ für } (x, y) \text{ mit } x \in (-\pi, \pi) \text{ und } V(x, \pm\pi) = 0.$$

In der Tat: Aus der Taylor-Entwicklung von \cos folgt $\cos x \leq 1 - \frac{2x^2}{\pi^2}$ für alle $x \in (-\pi, \pi)$, also

$$V(x, y) = \frac{1}{2}y^2 + 1 - \cos x + \alpha y \sin x \geq \frac{1}{2}(1 - \alpha)y^2 + \frac{2x^2}{\pi^2} - \frac{\alpha}{2}x^2 = \frac{1}{2}(1 - \alpha)y^2 + \left[\frac{2}{\pi^2} - \frac{\alpha}{2}\right]y^2 > 0.$$

Für die Ableitung in Richtung des Vektorfeldes f gilt

$$\begin{aligned} \dot{V}(x, y) &= DV(x, y)f(x, y) = [\sin x + \alpha y \cos x, y + \alpha \sin x] \begin{bmatrix} y \\ -\sin x - ky \end{bmatrix} \\ &= y \sin x + \alpha y^2 \cos x - y \sin x - ky^2 - \alpha \sin^2 x - \alpha k \sin x \\ &= -ky^2 + \alpha y^2 \cos x - \alpha \sin^2 x - \alpha y \sin x \\ &\leq -(k - \alpha)y^2 - \alpha \sin^2 x - \alpha ky \sin x \\ &= -[\sin x, y] \begin{bmatrix} \alpha & \alpha k/2 \\ \alpha k/2 & k - \alpha \end{bmatrix} \begin{bmatrix} \sin x \\ y \end{bmatrix}, \end{aligned}$$

denn Ausmultiplizieren ergibt

$$\begin{aligned} [\sin x, y] \begin{bmatrix} \alpha & \alpha k/2 \\ \alpha k/2 & k - \alpha \end{bmatrix} \begin{bmatrix} \sin x \\ y \end{bmatrix} &= [\alpha \sin x + y\alpha k/2, \sin x \alpha k/2 + y(k - \alpha)] \begin{bmatrix} \sin x \\ y \end{bmatrix} \\ &= \alpha \sin^2 x + y\alpha k/2 \sin x + \alpha k/2 y \sin x + y^2(k - \alpha). \end{aligned}$$

Die Determinante der Matrix $A = \begin{bmatrix} \alpha & \alpha k/2 \\ \alpha k/2 & k - \alpha \end{bmatrix}$ ist

$$\det A = \alpha(k - \alpha) - \alpha^2 k^2/4 = \alpha[k - \alpha(1 + k^2/4)] > 0,$$

falls

$$\alpha < \frac{4k}{4 + k^2}. \quad (7.3.4)$$

Die Spur von A ist $\text{Spur} A = k > 0$.

Die Determinante ist das Produkt der Eigenwerte, die Spur die Summe der Eigenwerte. Sind die Determinante und die Spur dieser zwei-dimensionalen symmetrischen, also diagonalisierbaren

Matrix positiv, so sind die beiden reellen Eigenwerte positiv. Die Matrix ist daher positiv definit, also $\dot{V}(x, y) < 0$ für alle (x, y) mit $0 \neq x \in (-\pi, \pi)$.

Damit ist die asymptotische Stabilität des Gleichgewichts $(0, 0)$ für die Pendelgleichung (7.3.1) nachgewiesen. Wir wollen aber die Lyapunov-Funktion auch benutzen, um Informationen über den Einzugsbereich von Gleichgewichten zu erhalten, vgl. Bemerkung 7.2.14. Dafür

Nun gilt $V(x, y) > 0$ nur auf dem Streifen $(-\pi, \pi) \times \mathbb{R}$, der nicht positiv invariant ist. Wähle jetzt $0 < \gamma < \frac{4}{3}$ und betrachte die (positiv invariante) Subniveaumenge

$$N_\gamma = \{(x, y) \mid V(x, y) < \gamma\}.$$

Dann ist wegen $\frac{y^2}{2} \geq 0$ und (7.3.3) für $|x| \leq \pi, |y| \leq 1/k$

$$\gamma > V(x, y) \geq -\frac{\alpha}{2}(x^2 + y^2) + 1 - \cos x \geq -\frac{\alpha}{2}(\pi^2 + 1/k^2) + 1 - \cos x \geq -\frac{\gamma}{2} + 1 - \cos x,$$

falls α so klein ist, dass $\alpha(\pi^2 + 1/k^2) < \gamma$. Gilt also $(x, y) \in \tilde{N}_\gamma$ und $|x| \leq \pi, |y| \leq 1/k$, insbesondere also $V(x, y) < \gamma$ so folgt

$$1 - \cos x < \frac{3}{2}\gamma \text{ d.h., } \cos x > 1 - \frac{3}{2}\gamma,$$

und $1 - \frac{3}{2}\gamma > -1$ wegen $\frac{3}{2}\gamma < 2$. Die letzte Ungleichung ist äquivalent zu

$$x \in \left(-\arccos\left(1 - \frac{3}{2}\gamma\right), \arccos\left(1 - \frac{3}{2}\gamma\right) \right) \subset (-\pi, \pi) \quad (7.3.5)$$

Gilt die Bedingung $|y| \leq k$ in $t = 0$, so auch für $t > 0$, weil für die zweite Komponente f der Differentialgleichung (7.3.2), also

$$g(x, y) := -\sin x - ky,$$

gilt: $g(x, 1/k) = -\sin x - 1 \leq 0$ und $g(x, -1/k) \geq 0$ für alle x .

Wir erhalten daher: Gilt (7.3.5) und $|y| \leq 1/k$ sowie $(x, y) \in \tilde{N}_\gamma$, so sind alle Bedingungen aus Bemerkung 7.2.14 erfüllt. Daher ist die Menge dieser Punkte (x, y) im Einzugsbereich $A(0, 0)$ des asymptotisch stabilen Gleichgewichts $(0, 0)$ enthalten.

7.4 Diskussion der Stabilität

Stabilitätstheorie ist ein sehr umfangreiches Teilgebiet der Theorie von Differentialgleichungen und wächst immer noch rasch an, mit hunderten von wissenschaftlichen Arbeiten in jedem Jahr (nicht alle sind wichtig!). Alle Konzepte und Beweise aus den vorangehenden Abschnitten können in viel allgemeinerem Rahmen und unter Heranziehung vieler weiterer Ideen und Methoden entwickelt werden. Wir gehen hier nur auf einige dieser Verallgemeinerungen kurz ein.

Aus der asymptotischen Stabilität der in einem Gleichgewicht x^* linearisierten Differentialgleichung folgt die asymptotische Stabilität von x^* . Hierbei ist die asymptotische Stabilität der linearisierten Differentialgleichung äquivalent zu der Eigenschaft, dass die entsprechende Jacobi-Matrix nur Eigenwerte mit negativem Realteil hat. Es gilt auch: Hat die Jacobi-Matrix einen Eigenwert mit positivem Realteil, so ist das nichtlineare Gleichgewicht nicht asymptotisch stabil (vgl. Grüne und Junge [7, Satz 7.7 und die anschließende Bemerkung über instabile Mannigfaltigkeiten])

Häufig ist nicht nur die Stabilität einer Ruhelage interessant, sondern man betrachtet eine nichttriviale Lösung $\varphi(t; x_0)$, $t \geq 0$, und möchte wissen, wie sich Lösungen verhalten, deren Anfangswerte nahe an x_0 liegen. Eine naheliegende Verallgemeinerung wäre die folgende: Man könnte etwa eine Lösung $\varphi(t; x_0)$, $t \geq 0$, von (7.1.1) stabil nennen, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass

$$\|x_0 - x_1\| < \delta \text{ impliziert } \|\varphi(t; x_0) - \varphi(t; x_1)\| < \varepsilon \text{ für alle } t \geq 0.$$

Im einfachsten Fall, der über ein Gleichgewicht hinausgeht, ist $\varphi(t; x_0)$, $t \geq 0$, eine periodische Lösung, also $\varphi(t; x_0) = \varphi(t + T; x_0)$ für ein $T > 0$ und $\varphi(t; x_0) \neq \varphi(s; x_0)$ für alle $t, s \in (0, T)$. Das folgende Beispiel zeigt eine der Schwierigkeiten, die hierbei auftreten.

Beispiel 7.4.1 Betrachte für $\alpha > 0$ das ebene autonome System

$$\dot{x} = f(x, y) = y(x^2 + y^2)^\alpha, \quad \dot{y} = g(x, y) = -x(x^2 + y^2)^\alpha.$$

Die Lösung für Anfangswert $(x_0, y_0) \in \mathbb{R}^2$ ist

$$\varphi(t; x_0, y_0) = (x_0 \cos [t(x_0^2 + y_0^2)^\alpha] + y_0 \sin [t(x_0^2 + y_0^2)^\alpha], y_0 \cos [t(x_0^2 + y_0^2)^\alpha] - x_0 \sin [t(x_0^2 + y_0^2)^\alpha]).$$

Dies rechnet man leicht nach. Es sei erwähnt, dass man diese Lösung durch Einführung von Polarkoordinaten

$$x = r \cos \phi, \quad y = r \sin \phi$$

erhalten kann. Dann ist $r^2 = x^2 + y^2$, also erhält man nach der Kettenregel und Einsetzen

$$r \cdot \dot{r} = x \cdot \dot{x} + y \cdot \dot{y} = x \cdot f(x, y) + y \cdot g(x, y) = r \cdot f(x, y) \cdot \cos \phi + r \cdot g(x, y) \cdot \sin \phi,$$

die Gleichung für die r -Komponente. Analog erhält man die Gleichung für ϕ aus $\tan \phi = \frac{y}{x}$. Rücktransformation in cartesische Koordinaten liefert die obige Lösung.

Für jedes α besteht das Phasenporträt aus konzentrischen Kreisen um das Gleichgewicht $0 \in \mathbb{R}^2$. Ferner ist jede Lösung periodisch mit Periode T , gegeben durch

$$2\pi = T(x_0^2 + y_0^2)^\alpha, \quad \text{also } T = \frac{2\pi}{(x_0^2 + y_0^2)^\alpha}.$$

Das Gleichgewicht 0 ist für jedes α stabil. Für $\alpha = 0$ sind auch alle anderen Lösungen stabil. Für $\alpha > 0$ sind die echt periodischen Lösungen jedoch nicht stabil, weil dann die Periode vom Anfangswert (x_0, y_0) abhängt, die Trajektorien also verschieden schnell durchlaufen werden. Die Distanz zwischen

$$\varphi(t; x_0, y_0) \text{ und } \varphi(t; x_1, y_1) \text{ für } (x_0, y_0) \neq (x_1, y_1),$$

geht also für $t \rightarrow \infty$ nicht gegen 0 .

Bei periodischen Lösungen unterscheidet man daher zwischen der oben definierten Stabilität und der sog. orbitalen Stabilität, bei der die Parametrisierung der Trajektorien durch die Zeit t nicht berücksichtigt wird.

Wie bei Gleichgewichten kann man auch bei periodischen Lösungen oder allgemeinen Lösungen Linearisierungsmethoden verwenden. Sei etwa $x(t) = \varphi(t; x^*), t \geq 0$, eine Lösung. Für einen Anfangswert $x_0 \neq x^*$ erhält man für $y(t) = \varphi(t; x_0)$, dass für die Abweichung $z(t) = y(t) - x(t)$ gilt

$$\dot{z}(t) = \dot{y}(t) - \dot{x}(t) = f(y(t)) - f(x(t)).$$

Wie im Fall des Gleichgewichts liefert Linearisierung (nun in jedem Punkt $x(t)$, $t \geq 0$) die linearisierte Differentialgleichung

$$\dot{z} = \frac{\partial f(x(t))}{\partial x} [y(t) - x(t)] = \frac{\partial f(x(t))}{\partial x} z(t).$$

mit der Jacobi-Matrix $\frac{\partial f(x(t))}{\partial x}$. Dies ist wieder eine lineare, nun aber nicht-autonome Differentialgleichung. Wie wir in Beispiel 7.1.6 gesehen haben, kann man die Stabilität nicht durch Eigenwert-Analyse charakterisieren. Einen Ausweg bieten Lyapunov-Exponenten, die folgendermaßen definiert sind.

Betrachte eine Lösung $z(t)$, $t \in \mathbb{R}$, von $\dot{z} = A(t)z$ mit einer stetigen Matrix-Funktion $t \mapsto A(t) : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$. Dann ist der zugehörige Lyapunov-Exponent oder die exponentielle Wachstumsrate

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|z(t)\|.$$

Man kann zeigen, dass sich im autonomen Fall gerade die Realteile der Eigenwerte von A ergeben. Im Allgemeinen muss dieser Limes aber nicht existieren und man betrachtet stattdessen etwa

den Limes superior. Schwierige Fragen ergeben sich, wenn man studiert, wann (etwa für eine linearisierte Differentialgleichung) diese Limes existieren, und ob es dazu eine Zerlegung des \mathbb{R}^n gibt, die analog zur Zerlegung in verallgemeinerte Eigenräume im autonomen Fall ist (vgl. Abschnitt 6.2). Dies hängt auch mit der Frage zusammen, ob die Attraktivität einer Lösung der nichtlinearen Gleichung folgt, wenn die entlang der Lösung linearisierte Gleichung asymptotisch stabil ist.

7.5 Regelungsprobleme

Für technische Systeme sind Instabilitäten unerwünscht. Man versucht daher, das Stabilitätsverhalten von Systemen zu verändern, also insbesondere Gleichgewichte stabil zu machen. In diesem Abschnitt werden wir einige grundlegende Ideen dazu diskutieren. Eine allgemeine Einführung in dieses Gebiet gibt zum Beispiel E. Sontag [14].

Das Analogon zu linearen autonomen Differentialgleichungen sind lineare autonome Kontrollsysteme, die die Form

$$\dot{x}(t) = Ax(t) + Bu(t)$$

mit Matrizen $A \in \mathbb{R}^{n \times n}$ und $B \in \mathbb{R}^{n \times m}$ haben. Dabei ist $\dot{x} = Ax$ das unkontrollierte (auch ungesteuerte, unregelte) System, das durch Wahl von Steuerungen u beeinflusst werden soll.

Der Einfachheit halber betrachten wir hier nur ein einfaches Beispiel, das Pendelmodell ohne Reibung

$$\begin{aligned}\dot{y}_1 &= y_2 \\ \dot{y}_2 &= -g \sin y_1.\end{aligned}$$

Linearisierung im instabilen Gleichgewicht $(\pi, 0)^T$ ergibt die Matrix

$$\begin{bmatrix} 0 & 1 \\ -g \cos \pi & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ g & 0 \end{bmatrix},$$

wir erhalten also die lineare autonome Differentialgleichung

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= gx_1\end{aligned}$$

Wir nehmen nun an, dass das Pendel auf einem Wagen montiert ist, dessen Beschleunigung u wir durch einen Motor beeinflussen können. Die Kontrolle ist hier also eindimensional.

Um den Wagen in unsere Differentialgleichung einzubeziehen, müssen wir zwei weitere Zustände hinzunehmen: Die Position x_3 des Wagens und seine Geschwindigkeit x_4 . Die Differentialgleichungen für den Wagen lauten dann

$$\dot{x}_3 = x_4, \quad \dot{x}_4 = u.$$

Ebenso beeinflusst u natürlich das Pendel: wird der Wagen nach links beschleunigt, so bewegt sich das Pendel nach rechts, weswegen die Differentialgleichung $\dot{x}_2 = -gx_1$ zu

$$\dot{x}_2 = -gx_1 + u$$

wird. Insgesamt führt dies für $x = (x_1, x_2, x_3, x_4)^T \in \mathbb{R}^4$ auf das Kontroll- oder Regelungssystem

$$\dot{x} = Ax + Bu = \begin{bmatrix} 0 & 1 & 0 & 0 \\ g & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u.$$

Betrachtet man die Lösung der linearisierten Pendelgleichung mit Anfangswert nahe 0, so stellt man fest, dass sich die Lösungen vom Ursprung entfernen. Das gleiche gilt für die nichtlineare

Pendelgleichung für Startwerte nahe $(\pi, 0)^T$. Unser Ziel ist nun, die Beschleunigung u so zu wählen, dass dies nicht passiert, d.h., wir wollen das Pendel in der instabilen umgekehrten Position balancieren. Sicherlich müssen wir dabei u mit der Zeit variieren, damit das klappt. Es ist aber im Allgemeinen nicht ratsam, die dafür notwendige Beschleunigung u direkt als Funktion der Zeit, also $u = u(t)$, zu berechnen: Wir müssten die Funktion dann für alle Zeiten $t \geq 0$ im Voraus berechnen und implementieren, was - wegen der Instabilität des Gleichgewichts - bereits bei kleinen Fehlern dazu führen würde, dass das Pendel am Ende doch wieder umkippt.

Stattdessen nehmen wir an, dass wir die Positionen und Geschwindigkeiten des Pendels und des Wagens - mithin den Vektor $x(t)$ - kontinuierlich messen können und die zu jedem Zeitpunkt anzuwendende Beschleunigung abhängig von diesem gemessenen Vektor $x(t)$ als $u(t) = F(x(t))$ für eine geeignete Abbildung F berechnen. Solch eine Abbildung F heißt Zustandsfeedback oder auch Zustandsrückführung. Da unser Modell linear ist, liegt es nahe, auch die Abbildung F als lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m , hier von \mathbb{R}^4 nach \mathbb{R}^1 , zu wählen. Das Feedback lässt sich daher als eine 1×4 -Matrix

$$F = (f_1, f_2, f_3, f_4) \text{ mit } f_i \in \mathbb{R}$$

schreiben. Statt Funktionen $u(t)$ für alle t zu berechnen, müssen wir nur noch die vier Einträge der Matrix F berechnen.

Setzt man $u = Fx$ in das Kontrollsystem ein, so erhält man

$$\dot{x} = Ax + Bu = Ax + BFx = [A + BF]x, \quad (7.5.1)$$

mit der Matrix

$$A + BF = \begin{bmatrix} 0 & 1 & 0 & 0 \\ f_1 + g & f_2 & f_3 & f_4 \\ 0 & 0 & 0 & 1 \\ f_1 & f_2 & f_3 & f_4 \end{bmatrix}. \quad (7.5.2)$$

Um das linearisierte Pendel in der aufrechten Position $x = 0$ zu balancieren, müssen wir nun erreichen, dass $x^* = 0$ asymptotisch stabil wird, also diese Matrix nur Eigenwerte mit negativem Realteil hat. Zum Beispiel können wir versuchen, F so zu wählen, dass die Matrix die vier Eigenwerte $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = -1$ besitzt. Um dies zu erreichen, gibt es in der Kontrolltheorie Techniken, die die Matrix mit Hilfe von Koordinatentransformationen in eine geeignete Normalform bringen, vgl. Sontag [14, Section 5.1]. Unser Beispiel hier ist aber so einfach, dass wir diese Techniken nicht brauchen, da man das charakteristische Polynom von $A + BF$ in Abhängigkeit von f_1, f_2, f_3, f_4 ganz einfach, zum Beispiel mit MAPLE, berechnen kann. Man erhält

$$\chi_{A+BF}(\lambda) = \det(\lambda I - A - BF) = \lambda^4 + (-f_4 - f_2)\lambda^3 + (-f_3 - f_1 - g)\lambda^2 + f_4g\lambda + f_3g.$$

Um nun den vierfachen Eigenwert -1 zu erzeugen, müssen wir f_1, f_2, f_3, f_4 so wählen, dass

$$\chi_{A+BF}(\lambda) = (\lambda + 1)^4 = \lambda^4 + 4\lambda^3 + 6\lambda^2 + 4\lambda + 1$$

gilt. Durch sukzessive Bestimmung der Koeffizienten sieht man, dass dies mit

$$F = \left[-g - \frac{1}{g} - 6, -\frac{4}{g} - 4, \frac{1}{g}, \frac{4}{g} \right]$$

gelingt. Dies Feedback stabilisiert den Ursprung für das linearisierte Kontrollsystem, also für $\dot{x} = (A + BF)x$ gegeben durch (7.5.2). Wir können das gleiche Feedback (angewendet auf die Abweichung vom Gleichgewicht in der nichtlinearen Gleichung für das auf einem Wagen montierte Pendel verwenden (die Variablen für die Position und die Geschwindigkeit des Wagens bezeichnen

wir jetzt mit y_1 und y_2):

$$\dot{y}_1 = y_2 \tag{7.5.3}$$

$$\dot{y}_2 = -g \sin y_1 + u = -g \sin y_1 + Fy = -g \sin y_1 + \left[-g - \frac{1}{g} - 6, -\frac{4}{g} - 4, \frac{1}{g}, \frac{4}{g} \right] [y_1 - \pi, x_2, y_3, y_4]^T$$

$$\dot{y}_3 = y_4$$

$$\dot{y}_4 = u = Fx = \left[-g - \frac{1}{g} - 6, -\frac{4}{g} - 4, \frac{1}{g}, \frac{4}{g} \right] [y_1 - \pi, x_2, y_3, y_4]^T$$

Man rechnet leicht nach, dass die Linearisierung dieser Differentialgleichung im Gleichgewicht $(\pi, 0, 0, 0)^T$ die lineare Differentialgleichung (7.5.1) für das linearisierte Kontrollsystem mit Feedback F ist. Der Linearisierungssatz, Satz 7.2.8, zeigt, dass für dieses Feedbacksystem das Gleichgewicht $(\pi, 0, 0, 0)^T$ asymptotisch stabil wird, also für kleine Auslenkungen aus der Ruhelage.

Es sei noch bemerkt, dass physikalisch korrekt für das nichtlineare System auch die Kontrollwirkung in der zweiten Gleichung nichtlinear angesetzt werden muss, wir haben hier der Einfachheit halber in der zweiten Gleichung von (7.5.3) gleich den linearisierten Term u verwendet.

Kapitel 8

Abhängigkeit von Parametern und Anfangswerten

In diesem Kapitel diskutieren wir die Änderung in Lösungen von Anfangswertproblemen, wenn sich die Differentialgleichung in Abhängigkeit von Parametern ändert. Typische Beispiele für Parameter sind l, g, k in der Pendelgleichung

$$\ddot{x} = -\frac{k}{l}\dot{x} - \frac{g}{l}x,$$

also die Pendellänge, die Proportionalitätskonstante der Reibung und die Erdbeschleunigung g . Man kann diese Parameter einzeln oder zusammen betrachten. Wir werden ihren Einfluss auf die Lösungen analysieren. Ebenso werden wir klären, wie die Lösungen von Anfangswerten abhängen, also die Eigenschaften der allgemeinen Lösung $\lambda(t; t_0, x_0), t \in I_{\max}(t_0, x_0)$.

8.1 Stetigkeit

Allgemein betrachten wir für Parameter α in \mathbb{R}^m

$$\dot{x} = f(t, x, \alpha) \tag{8.1.1}$$

mit $f : D \rightarrow \mathbb{R}^n$ definiert auf einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$. Wir setzen voraus, dass f stetig und bezüglich x Lipschitz-stetig ist. Dann existiert für jedes α und für jede Anfangsbedingung

$$x(t_0) = x_0 \text{ mit } (t_0, x_0, \alpha) \in D$$

die Lösung auf dem maximalen Existenzintervall I_{\max} . Um die Abhängigkeit von α zu verdeutlichen schreiben wir diese maximale Lösung als

$$\lambda(t; t_0, x_0, \alpha), t \in I_{\max}(t_0, x_0, \alpha), (t_0, x_0, \alpha) \in D.$$

Dies bezeichnen wir auch als die allgemeine Lösung der parameterabhängigen Gleichung (8.1.1).

Beispiel 8.1.1 Betrachte die skalare Differentialgleichung

$$\dot{x} = \frac{1}{t^2 + \alpha^2}$$

Die allgemeine Lösung ist, wie man sofort nachrechnet

$$\lambda(t; t_0, x_0) = \begin{cases} x_0 + \frac{1}{\alpha} \left[\arctan \frac{t}{\alpha} - \arctan \frac{t_0}{\alpha} \right] & \text{für } \alpha \neq 0 \\ x_0 + \frac{1}{t_0} - \frac{1}{t} & \text{für } \alpha = 0. \end{cases}$$

Für $\alpha \neq 0$ ist das maximale Lösungsintervall jeder Lösung gleich $(-\infty, \infty)$, für $\alpha = 0$ ist es gleich $(0, \infty)$ oder gleich $(-\infty, 0)$. Das Lösungsintervall ändert sich hier also drastisch in Abhängigkeit von dem Parameter α : Bei einer kleinen Änderung von α wird es sprunghaft größer.

Die folgende einfache Proposition zeigt, dass wir die Abhängigkeit von dem Parameter als Abhängigkeit von einem Anfangswert auffassen können.

Proposition 8.1.2 *Ist $\lambda(t), t \in I$, eine Lösung von (8.1.1) mit Anfangsbedingung $\lambda(t_0) = x_0$, so ist $(\lambda(t), \alpha), t \in I$, eine Lösung von*

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f(t, x, y) \\ 0 \end{bmatrix} \quad (8.1.2)$$

mit Anfangsbedingung

$$\begin{bmatrix} x(t_0) \\ y(t_0) \end{bmatrix} = \begin{bmatrix} x(t_0) \\ \alpha \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R}^m.$$

Ist umgekehrt $(\lambda(t), \mu(t)), t \in I$, eine Lösung dieses Anfangswertproblems, so ist $\lambda(t), t \in I$, eine Lösung von (8.1.1) mit $\lambda(t_0) = x_0$.

Beweis. Dies ist offensichtlich, weil die Lösungen von $\dot{y} = 0$ konstant sind. ■

Offenbar brauchen wir für die Differentialgleichung (8.1.1) Lipschitz-Stetigkeit bezüglich α nicht, um die Existenz und Eindeutigkeit der Lösungen auf maximalen offenen Intervallen zu zeigen. Die Proposition zeigt, dass die Stetigkeit von Lösungen bezüglich Parametern in der rechten Seite folgt, wenn wir die Stetigkeit bezüglich Anfangswerten zeigen können. Deswegen diskutieren wir jetzt dieses Problem.

Satz 8.1.3 *Es sei $f : D \rightarrow \mathbb{R}^n$ mit einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n$ eine stetige und bezüglich x Lipschitz-stetige Funktion. Dann ist für $\dot{x} = f(t, x)$ jede Lösung $\lambda(t; t_0, x_0)$ für alle t aus ihrem maximalen Existenzintervall $I(t_0, x_0)$ stetig im Anfangswert x_0 . Darüberhinaus ist sie Lipschitz-stetig in x_0 im folgenden Sinn:*

Für jedes kompakte Intervall $I \subset I(t_0, x_0)$ gibt es Konstanten $\delta > 0$ und $L_I > 0$, so dass für alle $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\| \leq \delta$ die Inklusion $I \subset I(t_0, y_0)$ und die Ungleichung

$$\|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0)\| \leq L_I \|y_0 - x_0\| \quad (8.1.3)$$

für alle $t \in I$ gilt.

Beweis. Wir beweisen die angegebene Ungleichung für alle $t \in I$ mit $t \geq t_0$, der Beweis für $t \leq t_0$ funktioniert analog.

Da der Definitionsbereich D von f offen ist, können wir für das gegebene kompakte Intervall I ein $\delta_1 > 0$ finden, so dass die kompakte Menge

$$K := \{(t, y) \in I \times \mathbb{R}^n \mid \|y - \lambda(t; t_0, x_0)\| \leq \delta_1\}$$

in D liegt.

Wir beweisen die angegebene Abschätzung nun zunächst für alle $y_0 \in \mathbb{R}^n$ und $t \in I$, für die

$$(s, \lambda(s; t_0, y_0)) \in K \text{ für alle } s \in [0, t] \quad (8.1.4)$$

gilt. Beachte, dass wegen der Stetigkeit der Lösung in t für jedes y_0 mit $\|y_0 - x_0\| < \delta_1$ ein $t > t_0$ existiert, für das diese Bedingung erfüllt ist.

Zum Beweis der Abschätzung für diese y_0 und t sei L eine Lipschitz-Konstante von f in x auf der kompakten Menge K (siehe Proposition 2.2.6). Damit gilt

$$\begin{aligned} \|\lambda(t; t_0, x_0) - \lambda(t; t_0, y_0)\| &= \left\| x_0 + \int_{t_0}^t f(s, \lambda(s; t_0, x_0)) ds - y_0 - \int_{t_0}^t f(s, \lambda(s; t_0, y_0)) ds \right\| \\ &\leq \|x_0 - y_0\| + \int_{t_0}^t \|f(s, \lambda(s; t_0, x_0)) - f(s, \lambda(s; t_0, y_0))\| ds \\ &\leq \|x_0 - y_0\| + \int_{t_0}^t L \|\lambda(s; t_0, x_0) - \lambda(s; t_0, y_0)\| ds. \end{aligned}$$

Mit diesen Umformungen haben wir eine Ungleichung gefunden, auf die wir das Gronwall-Lemma, Lemma 2.2.7, anwenden können: Mit $u(t) := \|\lambda(t; t_0, x_0) - \lambda(t; t_0, y_0)\|$, $\alpha := \|x_0 - y_0\|$, $\beta := L$ gilt

$$u(t) \leq \alpha + \beta \int_{t_0}^t u(s) ds \text{ für alle } t \in I \text{ mit } t \geq t_0,$$

und es folgt $u(t) \leq \alpha e^{\beta(t-t_0)}$, also

$$\|\lambda(t; t_0, x_0) - \lambda(t; t_0, y_0)\| \leq \|x_0 - y_0\| e^{L(t-t_0)}. \quad (8.1.5)$$

Für $t \leq t_0$ gilt dieselbe Ungleichung mit $\|x_0 - y_0\| e^{L(t_0-t)}$. Damit erhalten wir die gewünschte Ungleichung für

$$L_I := \max_{t \in I} e^{L|t-t_0|}.$$

Wir müssen noch die Voraussetzung (8.1.4) loswerden. Dafür müssen wir $\delta > 0$ konstruieren, so dass die Ungleichung für alle y_0 mit $\|x_0 - y_0\| \leq \delta$ gilt. Dies erreichen wir, indem wir $\delta > 0$ so wählen, dass die Lösung $\lambda(t; t_0, y_0)$ für alle diese y_0 und alle $t \in I, t \geq t_0$, in K liegt. Wir setzen dazu

$$\delta := \frac{\delta_1}{2L_I}.$$

Nehmen wir nun an, dass ein Anfangswert $y_0 \in \mathbb{R}^n$ mit $\|x_0 - y_0\| \leq \delta$ existiert, dessen zugehörige Lösung $\lambda(t; t_0, y_0)$ zu einer Zeit $t \in I, t \geq t_0$ nicht mehr in K liegt. Dann gibt es wegen der Stetigkeit der Lösung in t eine minimale Zeit $t_1 \geq t_0$, zu der die Lösung gerade am Rand der Menge K liegt, also

$$\|\lambda(t_1; t_0, x_0) - \lambda(t_1; t_0, y_0)\| = \delta_1 \quad (8.1.6)$$

und

$$\|\lambda(t; t_0, x_0) - \lambda(t; t_0, y_0)\| < \delta_1 \text{ für alle } t \in [t_0, t_1].$$

Auf dem Intervall $[t_0, t_1]$ liegt die Lösung also in K , weswegen wir die Ungleichung (8.1.5) für $t = t_1$ anwenden können. Wegen $\|x_0 - y_0\| \leq \delta$ und der Wahl von δ ergibt dies

$$\|\lambda(t_1; t_0, x_0) - \lambda(t_1; t_0, y_0)\| \leq L_I \|x_0 - y_0\| \leq \frac{\delta_1}{2},$$

im Widerspruch zu (8.1.6). Also liegen alle Lösungen mit Anfangswert $\|x_0 - y_0\| \leq \delta$ für alle Zeiten $t \in I, t \geq t_0$, in K und erfüllen daher die gewünschte Ungleichung.

Analog argumentiert man für $t < t_0$. ■

Eine unmittelbare Folgerung aus Proposition 8.1.2 und Satz 8.1.3 ist die Stetigkeit bezüglich Parametern auf der rechten Seite. Wir erhalten das folgende wichtige Korollar.

Korollar 8.1.4 *Betrachte die parameterabhängige Differentialgleichung (8.1.1)*

$$\dot{x} = f(t, x, \alpha)$$

mit einer Funktion $f : D \rightarrow \mathbb{R}^n$, definiert auf einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ und stetig und bezüglich x Lipschitz-stetig. Dann ist jede Lösung $\lambda(t; t_0, x_0, \alpha)$ für alle t aus ihrem maximalen Existenzintervall $I(t_0, x_0, \alpha)$ stetig in (x_0, α) .

Beweis. Wende Satz 8.1.3 auf die Differentialgleichung (8.1.2) an. ■

Diskussion der Lipschitz-Konstante. Beachte, dass die im Beweis errechnete Lipschitz-Konstante

$$L_I := \max_{t \in I} e^{L|t-t_0|}$$

um so größer wird, je größer das Intervall I gewählt wird. Dies erklärt, warum die Instabilität von Gleichgewichten kein Widerspruch zur Stetigkeit der Lösungen im Anfangswert ist: Für große

Intervalle kann die Lipschitz-Konstante so groß werden, dass selbst Lösungen zu sehr nahe beieinanderliegenden Anfangswerten nach einer hinreichend langen Zeit stark voneinander abweichen können.

Wie nah ist die im Beweis berechnete Konstante aber nun an der tatsächlichen Lipschitz-Konstante der Lösung? Liefert L_I eine scharfe Abschätzung, oder überschätzen wir die richtige Konstante damit? Die Antwort auf diese Frage hängt ganz von der betrachteten Differentialgleichung ab. Um dies zu veranschaulichen, genügt die bekannte einfache skalare lineare Differentialgleichung

$$\dot{x} = cx \tag{8.1.7}$$

für $c \in \mathbb{R}$. Für diese rechnet man leicht nach, dass die Lipschitz-Konstante des Vektorfeldes $f(x) = cx$ gerade $L = |c|$ ist. Betrachten wir die Anfangszeit $t_0 = 0$ und das Intervall $I = [0, T]$, so erhalten wir aus dem Beweis von Satz 8.1.3 die Lipschitz-Konstante $L_I = e^{|c|T}$. Die Lösungen der Gleichung (8.1.7) sind bekanntlich gegeben durch

$$x(t) = x_0 e^{ct}.$$

Auf dem Intervall I können wir die exakte Lipschitz-Konstante L_{exakt} mittels

$$\|\lambda(t; 0, x_0) - \lambda(t; 0, y_0)\| = \|x_0 e^{ct} - y_0 e^{ct}\| \leq |x_0 - y_0| e^{ct} = |x_0 - y_0| \max_{t \in I} e^{ct}$$

als

$$L_{\text{exakt}} = \begin{cases} e^{cT}, & c \geq 0 \\ 1, & c < 0 \end{cases}$$

errechnen. Wir sehen: Für $c \geq 0$ ist $L_I = e^{LT} = e^{cT} = L_{\text{exakt}}$ und wir erhalten aus Satz 8.1.3 tatsächlich eine scharfe Abschätzung der exakten Lipschitz-Konstante. Für $c < 0$ hingegen folgt

$$L_I = e^{LT} = e^{-cT} \text{ und } L_{\text{exakt}} = 1.$$

d.h. wenn $|c|T$ groß ist, überschätzen wir die exakte Lipschitz-Konstante beträchtlich!

Stetigkeit in der Anfangszeit. Zum Abschluss dieses Abschnitts betrachten wir noch die Stetigkeit der Lösung bezüglich der Anfangszeit t_0 , die sich mit Hilfe der bereits bekannten Stetigkeit bezüglich x_0 beweisen lässt.

Satz 8.1.5 *Es sei $f : D \rightarrow \mathbb{R}^n$ mit einer offenen Menge $D \subset \mathbb{R} \times \mathbb{R}^n$ eine stetige und bezüglich x Lipschitz-stetige Funktion. Dann ist jede Lösung $\lambda(t; t_0, x_0)$ von $\dot{x} = f(t, x)$ für alle t aus ihrem maximalen Existenzintervall $I(t_0, x_0)$ stetig in der Anfangszeit t_0 .*

Beweis. Wir beweisen die Stetigkeit, indem wir für eine beliebige Folge von Anfangszeiten t_k die Implikation

$$t_k \rightarrow t_0 \text{ impliziert } \lambda(t; t_k, x_0) \rightarrow \lambda(t; t_0, x_0)$$

zeigen, vorausgesetzt, dass die t_k hinreichend nahe an t_0 liegen. Zum Beweis der Implikation nutzen wir die Kozykluseigenschaft

$$\lambda(t; t_k, x_0) = \lambda(t; t_0, \lambda(t_0; t_k, x_0)).$$

aus. Nach Satz 8.1.3 ist $\lambda(t; t_0, \cdot)$ stetig, also genügt es, $\lambda(t_0; t_k, x_0) \rightarrow x_0$ zu zeigen. Aus der Integraldarstellung der Lösung folgt

$$\lambda(t_0; t_k, x_0) = x_0 + \int_{t_k}^{t_0} f(s, \lambda(s; t_k, x_0)) ds.$$

Wie im Beweis des Satzes von Picard-Lindelöf, Satz 2.4.4, zeigt man, dass $\|f(s, \lambda(s; t_k, x_0))\|$ für alle s aus dem Integrationsintervall und alle t_k hinreichend nahe an t_0 durch eine Konstante M beschränkt ist. Daraus folgt

$$\|\lambda(t_0; t_k, x_0) - x_0\| = \left\| \int_{t_k}^{t_0} f(s, \lambda(s; t_k, x_0)) ds \right\| \leq |t_0 - t_k| M.$$

und daher $\|\lambda(t_0; t_k, x_0) - x_0\| \rightarrow 0$ für $k \rightarrow \infty$, also $\lambda(t_0; t_k, x_0) \rightarrow x_0$. Dies zeigt die Behauptung. ■

8.2 Differenzierbarkeit

In diesem Abschnitt diskutieren wir die Differenzierbarkeit der Lösung von $\dot{x} = f(t, x)$ nach Anfangswerten. Wir haben im vorhergehenden Abschnitt gesehen, dass die Lipschitz-Stetigkeit von f bezüglich x die Lipschitz-Stetigkeit der Lösung bezüglich des Anfangswertes x_0 impliziert. Es stellt sich daher in natürlicher Weise die Frage, ob man auch Differenzierbarkeit der Lösung bezüglich x_0 erwarten kann, wenn man Differenzierbarkeit von f bezüglich x voraussetzt. Dies ist tatsächlich der Fall, wie wir in diesem Abschnitt sehen werden. Bevor wir die stetige Differenzierbarkeit nachweisen, überlegen wir uns zunächst, wie die Ableitung aussehen muss, falls sie existiert. Eine formale Rechnung zeigt für die Ableitung $\frac{d}{dx_0} \lambda(t; t_0, x_0) \in \mathbb{R}^{n \times n}$ der Lösung $\lambda(t; t_0, x_0)$ nach dem Anfangswert x_0 , dass

$$\frac{d}{dt} \frac{d}{dx_0} \lambda(t; t_0, x_0) = \frac{d}{dx_0} \frac{d}{dt} \lambda(t; t_0, x_0) = \frac{d}{dx_0} f(t, \lambda(t; t_0, x_0)) = \frac{\partial}{\partial x} f(t, \lambda(t; t_0, x_0)) \frac{d}{dx_0} \lambda(t; t_0, x_0).$$

Wir erwarten daher, dass die Ableitung nach dem Anfangswert die entlang der Lösung linearisierte Differentialgleichung erfüllt, und beginnen mit einer Differenzierbarkeitseigenschaft von f .

Proposition 8.2.1 *Es sei $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, stetig, für alle t nach x differenzierbar, und die Ableitung sei stetig in (t, x) . Dann ist die Jacobi-Matrix $D_x f(t, x) = \frac{\partial f}{\partial x}(t, x) : D \rightarrow \mathbb{R}^{n \times n}$ stetig und für jede kompakte Teilmenge $K \subset D$ existiert eine Funktion $r : \mathbb{R} \rightarrow \mathbb{R}$ mit*

$$\|f(t, y) - f(t, x) - D_x f(t, x)(y - x)\| \leq r(\|y - x\|)$$

für alle $(t, x), (t, y) \in K$ und

$$\lim_{\alpha \rightarrow 0, \alpha \neq 0} \frac{r(\alpha)}{\alpha} = 0.$$

Dies beweist man ähnlich wie Proposition 2.2.6, die eine ähnliche Aussage für Lipschitz-Stetigkeit enthält. Ferner impliziert nach Proposition 2.2.4 die obige Bedingung, dass f Lipschitz-stetig in x ist.

Im Folgenden betrachten wir eine Lösung $\lambda(t; t_0, x_0)$ der Differentialgleichung $\dot{x} = f(t, x)$ und ein kompaktes Intervall $I \subset I(t_0, x_0)$. Entlang der Lösung definieren wir

$$A(t) := \frac{\partial}{\partial x} f(t, \lambda(t; t_0, x_0)), \quad t \in I. \quad (8.2.1)$$

Mit Hilfe dieser zeitabhängigen Matrix können wir nun die lineare nichtautonome Differentialgleichung

$$\dot{z}(t) = A(t)z(t) \quad (8.2.2)$$

aufstellen, die sogenannte **Linearisierung oder Variationsgleichung** von $\dot{x} = f(t, x)$ entlang der Lösung $\lambda(t; t_0, x_0)$. Ihre Lösung mit Anfangsbedingung $z(t_0) = z_0$ ist durch $\Phi(t, t_0)z_0$ gegeben, wobei $\Phi(t, t_0)$ die Übergangsmatrix aus (6.1.5) ist.

Um die Differenzierbarkeit nach x_0 nachzuweisen und die Ableitung zu berechnen, zeigen wir im folgenden Lemma zunächst eine explizite Abschätzung über den Zusammenhang zwischen $\lambda(t; t_0, y_0)$, $\lambda(t; t_0, x_0)$ und $\Phi(t, t_0)(y_0 - x_0)$.

Lemma 8.2.2 *Betrachte eine nichtlineare Differentialgleichung $\dot{x} = f(t, x)$, deren Vektorfeld f die Eigenschaften aus Proposition 8.2.1 erfüllt. Es sei $\lambda(t; t_0, x_0)$ eine Lösung der Gleichung und $\Phi(t, t_0)$ die Übergangsmatrix der Linearisierung gemäß (8.2.1). Zudem sei ein kompaktes Intervall $I \subset I(t_0, x_0)$ und ein $\varepsilon > 0$ gegeben. Dann gibt es ein $\delta > 0$, so dass für jeden Anfangswert $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\| < \delta$ und alle $t \in I$ die folgende Abschätzung gilt:*

$$\|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0) - \Phi(t, t_0)(y_0 - x_0)\| \leq \varepsilon \|y_0 - x_0\|.$$

Beweis. Wir wählen eine Anfangszeit $t_0 \in I$ und einen Anfangswert x_0 . Aus Satz 8.1.3 folgt die Existenz von $\delta > 0$ und $L_I > 0$, so dass für alle Anfangswerte y_0 mit $\|y_0 - x_0\| < \delta$ und alle $t \in I$

$$\|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0)\| \leq L_I \|y_0 - x_0\|.$$

Wir definieren die kompakte Menge

$$K = \{(t, y) \mid t \in I \text{ und } \|y - \lambda(t; t_0, x_0)\| \leq L_I \delta\}$$

und die Funktion

$$\rho(t, y) := f(t, y) - f(t, \lambda(t; t_0, x_0)) - A(t) [y - \lambda(t; t_0, x_0)].$$

Für $(t, y) \in K$ gilt dann nach Proposition 8.2.1 die Abschätzung

$$\rho(t, y) \leq r(\|y - \lambda(t; t_0, x_0)\|).$$

Wähle nun $\varepsilon > 0$ und setze $T := \max_{t \in I} |t - t_0|$ und $M := \max_{t, s \in I} \|\Phi(t; s)\|$, wobei wir die von der Euklidischen Norm induzierte Matrixnorm verwenden, also $\|\Phi(t; s)\| := \sup_{\|x\|_2 \leq 1} \|\Phi(t; s)x\|_2$; dann $\|\Phi(t; s)x\|_2 \leq \|\Phi(t; s)\| \|x\|_2$. Aus den Eigenschaften von r folgt, dass ein $\delta_1 > 0$ existiert mit

$$r(\|y - x\|) \leq \frac{\varepsilon}{MTL_I} \|y - x\|$$

für alle $x, y \in \mathbb{R}^n$ mit $\|y - x\| \leq \delta_1$. Wir können annehmen, $\delta \leq \delta_1$ und $\delta \leq \delta_1/L_I$ gilt. Für jeden Anfangswert $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\| < \delta$ und alle $t \in I$ gilt dann

$$\begin{aligned} \dot{\lambda}(t; t_0, y_0) - \dot{\lambda}(t; t_0, x_0) &= f(t, \lambda(t; t_0, y_0)) - f(t, \lambda(t; t_0, x_0)) \\ &= A(t) [\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0)] + \rho(t, \lambda(t; t_0, y_0)). \end{aligned}$$

Die Funktion

$$z(t) := \lambda(t; t_0, y_0) - \lambda(t; t_0, x_0)$$

erfüllt also die lineare inhomogene Differentialgleichung

$$\dot{z}(t) = A(t)z(t) + \rho(t; \lambda(t; t_0, y_0)) \text{ mit } z(t_0) = y_0 - x_0.$$

Mit der Variation-der-Parameter-Formel für die Lösung dieser Gleichung gilt

$$z(t) = \Phi(t, t_0)(y_0 - x_0) + \int_{t_0}^t \Phi(t, s)\rho(s, \lambda(s; t_0, y_0))ds.$$

Also folgt für $t \in I$, wie behauptet,

$$\begin{aligned} \|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0) - \Phi(t, t_0)(y_0 - x_0)\| &= \|z(t) - \Phi(t, t_0)(y_0 - x_0)\| \\ &= \left\| \int_{t_0}^t \Phi(t, s)\rho(s, \lambda(s; t_0, y_0))ds \right\| \\ &\leq \int_{t_0}^t \|\Phi(t, s)\| \|\rho(s, \lambda(s; t_0, y_0))\| ds \\ &\leq MT \sup_{s \in I} r(\|\lambda(s; t_0, y_0) - \lambda(s; t_0, x_0)\|) \\ &\leq MT \frac{\varepsilon}{MTL_I} L_I \|y_0 - x_0\| \leq \varepsilon \|y_0 - x_0\|. \end{aligned}$$

■

Bemerkung 8.2.3 *Mit anderen Worten liefert dieser Satz die Approximationseigenschaft*

$$\lambda(t; t_0, x_0) + \Phi(t, t_0)(y_0 - x_0) \approx \lambda(t; t_0, y_0),$$

deren Güte von der Anfangsdifferenz $z_0 := y_0 - x_0$ abhängt und deren Approximationsfehler für $z_0 \rightarrow 0$ schneller als linear abnimmt, also schneller als $c\|z_0\|$ für jede Konstante $c > 0$.

Der folgende Satz zeigt nun, dass sich mit der Abschätzung aus Lemma 8.2.2 leicht die Differenzierbarkeit der Lösung nach x_0 folgern lässt.

Satz 8.2.4 *Es sei $f : D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R} \times \mathbb{R}^n$ offen, stetig, für alle t nach x differenzierbar, und die Ableitung sei stetig in (t, x) . Dann ist die Lösung $\lambda(t; t_0, x_0)$ von $\dot{x} = f(t, x)$ für jeden Zeitpunkt $t \in I$ stetig differenzierbar im Anfangswert x_0 mit Ableitung*

$$\frac{d}{dx_0} \lambda(t; t_0, x_0) = \Phi(t, t_0),$$

wobei $\Phi(t, t_0)$ die Übergangsmatrix der Linearisierung (8.2.2) ist.

Beweis. Sei $\varepsilon > 0$ beliebig, $\delta > 0$ aus Lemma 8.2.2 und $y_0 \in \mathbb{R}^n$ mit $\|y_0 - x_0\| < \delta$. Dann gilt

$$\frac{\|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0) - \Phi(t, t_0)(y_0 - x_0)\|}{\|y_0 - x_0\|} \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, ergibt sich

$$\lim_{y_0 \rightarrow x_0, y_0 \neq x_0} \frac{\|\lambda(t; t_0, y_0) - \lambda(t; t_0, x_0) - \Phi(t, t_0)(y_0 - x_0)\|}{\|y_0 - x_0\|} = 0,$$

also die Differenzierbarkeit nach dem Anfangswert mit Ableitung $\Phi(t, t_0)$. Wegen (8.2.1) und Korollar 8.1.4 hängt $A(t) = \frac{\partial}{\partial x} f(t; \lambda(t; t_0, x_0))$ stetig von x_0 ab. Damit hängt, wieder nach Korollar 8.1.4, auch die Lösung der linearisierten Differentialgleichung stetig von x_0 ab. Schließlich folgt, dass auch die Matrix $\Phi(t, t_0)$ stetig von x_0 abhängt, also gilt die stetige Differenzierbarkeit. ■

Bemerkung 8.2.5 *Für alle $z_0 \in \mathbb{R}^n$ ist die Ableitung in Richtung z_0*

$$\frac{d}{dx_0} \lambda(t; t_0, x_0) z_0 = \Phi(t, t_0) z_0.$$

Kurz kann man das auch so ausdrücken: Die Linearisierung nach dem Anfangswert ist durch die linearisierte Differentialgleichung gegeben.

Bemerkung 8.2.6 *In Zusammenhang mit asymptotischer Stabilität haben wir für autonome Differentialgleichungen schon die Linearisierung in einem Gleichgewicht x^* betrachtet. Dies ist einerseits ein Spezialfall der hier betrachteten Linearisierung, nämlich entlang der konstanten Lösung $\lambda(t; 0, x^*) = x^*$. Andererseits haben wir dort das unbeschränkte Zeitintervall $[0, \infty)$ betrachtet.*

Bemerkung 8.2.7 *Die Lösung einer parameterabhängigen Differentialgleichung (8.1.1)*

$$\dot{x} = f(t, x, \alpha)$$

ist auch C^1 in α , wenn die Bedingungen von Satz 8.2.4 für (x, α) (anstelle von x) ist.

Bemerkung 8.2.8 *Indem man Satz 8.2.4 iterativ auf die Variationsgleichungen anwendet, kann man beweisen, dass die Lösung p -mal stetig differenzierbar nach x_0 (sowie nach eventuellen Parametern) ist, wenn das Vektorfeld f eine C^p -Funktion ist (vgl. Amann [2]).*

Im folgenden Korollar schreiben wir den autonomen Spezialfall von Satz 8.2.4 noch einmal auf.

Korollar 8.2.9 *Betrachte $\dot{x} = f(x)$ mit einer C^1 -Funktion $f : D \rightarrow \mathbb{R}^n$ definiert auf einer offenen Teilmenge $D \subset \mathbb{R}^n$. Dann ist der Fluss $\varphi(t; x_0)$ stetig nach x_0 differenzierbar und $\frac{d}{dx_0} \varphi(\cdot, x_0) z_0$ ist die Lösung von*

$$\dot{z} = \frac{\partial}{\partial x} f(\varphi(t; x_0)) z, \quad z(0) = z_0.$$

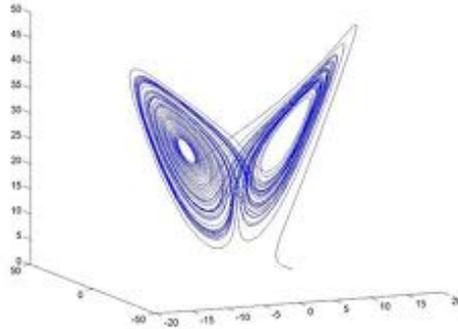


Abbildung 8.1: Eine Trajektorie der Lorenz-Gleichung

Der linearisierte Fluss ist die allgemeine Lösung der linearisierten Differentialgleichung. Man beachte, dass auch hier der Übergang zu einer nichtautonomen (linearen) Differentialgleichung nötig ist, falls nicht in einem Gleichgewicht x^* linearisiert wird. Wenn man

$$\begin{aligned}\dot{x} &= f(x) \\ \dot{y} &= \frac{\partial}{\partial x} f(x)y\end{aligned}$$

als **gekoppeltes System** betrachtet, erhält man allerdings wieder eine autonome Differentialgleichung; hierbei hängt die erste Komponente x nicht von der zweiten Komponente y ab. Den zugehörigen Fluss bezeichnet man auch als Schiefproduktfluss.

Die linearisierte Differentialgleichung beschreibt auf endlichen Zeitintervallen, wie sich nahe beieinander startende Lösungen verhalten. Es gibt Beispiele von Differentialgleichungen, bei denen für jeden Anfangswert nahe dabei startende Lösungen wegeinanderlaufen (weil die linearisierten Differentialgleichungen instabil sind), obwohl alle Lösungen für $t \geq 0$ beschränkt bleiben. Dies führt dazu, dass längerfristige Voraussagen (bei nicht exakt bekanntem Anfangswert) unmöglich sind. Ein berühmtes Beispiel ist die Lorenz-Gleichung in \mathbb{R}^3 , die von Lorenz (1963) als (stark) vereinfachtes Modell von Wetterphänomenen untersucht wurde:

$$\begin{aligned}\dot{x}_1 &= 10(x_2 - x_1) \\ \dot{x}_2 &= 28x_1 - x_2 - x_1x_3 \\ \dot{x}_3 &= x_1x_2 - 8/3x_3.\end{aligned}$$

Figur 8.1 präsentiert eine typische (numerisch berechnete) Trajektorie. Für die Lorenz-Gleichung zeigen die Trajektorien allerdings auch bei sehr kleinen Änderungen des Anfangswerts sehr unterschiedliches Verhalten..

Wie oben diskutiert, führt die Linearisierung entlang einer Lösung auf eine nichtautonome lineare Differentialgleichung der Form

$$\dot{x} = A(t)x \tag{8.2.3}$$

mit einer stetigen Abbildung $t \mapsto A(t) : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ (mit $n = 3$ für die Lorenz-Gleichung). Die Stabilitätseigenschaften solcher Differentialgleichungen kann man häufig durch ihre Lyapunov-Exponenten beschreiben: Für einen Anfangswert $x_0 \in \mathbb{R}^n$, $x_0 \neq 0$, zur Zeit $t_0 = 0$ ist der Lyapunov Exponent der Lösung $\varphi(t; 0, x_0)$ definiert als ihre exponentielle Wachstumsrate, gegeben durch

$$\lim_{t \rightarrow \infty} \frac{1}{t} \ln \|\varphi(t; 0, x_0)\|$$

(der Limes muss nicht existieren!). Ist der Lyapunov-Exponent negativ, so konvergiert die Trajektorie gegen 0, ist er positiv, so divergiert sie. Häufig lässt sich daraus schließen, dass Trajektorien

der nichtlinearen Differentialgleichung lokal exponential schnell auseinanderlaufen. Die Theorie der Lyapunov-Exponenten ist ein Ausgangspunkt der allgemeinen Theorie dynamischer Systeme.

8.3 Verzweigung von Ruhelagen

In diesem Abschnitt diskutieren wir kurz das Verzweigungs- oder Bifurkationsverhalten von Differentialgleichungen. Wir betrachten eine Parameter-abhängige Differentialgleichung der Form

$$\dot{x} = f(x, \alpha) \quad (8.3.1)$$

mit Parameter $\alpha \in \mathbb{R}$ und eine Funktion $f : D \rightarrow \mathbb{R}^n$, definiert auf einer offenen Teilmenge $D \subset \mathbb{R}^{n+1}$. Für α_0 besitze die Differentialgleichung eine Ruhelage x_0 , also $f(x_0, \alpha_0) = 0$.

Unter welchen Voraussetzungen besitzt die Differentialgleichung für Parameter α nahe α_0 weitere Ruhelagen in der Nähe von x_0 ? Wieviele Ruhelagen gibt es, und welche Stabilitätseigenschaften haben sie? Diese Fragen sind Gegenstand der ausgedehnten Bifurkationstheorie von Differentialgleichungen.

Allgemein versteht man unter einer Verzweigung eines von einem "externen" Parameter abhängigen dynamischen Systems die Änderung von qualitativen Eigenschaften seiner Lösungen. Die Detektion (und Klassifizierung) von Verzweigungen ist in Anwendungen von größtem Interesse, da so gut wie jedes (technische) System von Parametern abhängt, deren Wahl entscheidend das Systemverhalten beeinflussen kann.

Wir betrachten also eine autonome gewöhnliche Differentialgleichung (8.3.1), bei der das Vektorfeld f zusätzlich von einem externen Parameter $\alpha \in \mathbb{R}$ abhängt. In diesem Abschnitt nehmen wir an, dass f (nach x und α) stetig differenzierbar ist.

Die Sattel-Knoten-Verzweigung

Angenommen, das System hat für einen Parameterwert $\alpha = \alpha_0$ eine Gleichgewichtslösung $x_0 \in \mathbb{R}^n$, d.h. $0 = f(x_0, \alpha_0)$. Aus dem Satz über implizite Funktionen erhalten wir das folgende Lemma.

Lemma 8.3.1 *Es sei $f(x_0, \alpha_0) = 0$ und $Df(x_0, \alpha_0) \in \mathbb{R}^{n \times (n+1)}$ habe maximalen Rang n . Dann ist*

$$f^{-1}(0) = \{(x, \alpha) \mid f(x, \alpha) = 0\}$$

in der Nähe von (x_0, α_0) eine Kurve $(x(s), \alpha(s))$ mit s in einem offenen Intervall.

Falls $D_x f(x_0, \alpha_0)$ regulär ist, dann ist diese Kurve lokal durch α parametrisierbar, d.h. dann gibt es für α nahe α_0 eine Kurve $x(\alpha)$ mit $f(x(\alpha), \alpha) = 0$, und $D_x f(x(\alpha), \alpha)$ ist regulär für diese α .

Interessant ist nun der andere Fall, wenn $D_x f(x_0, \alpha_0)$ singulär ist. Es gebe genau einen Eigenwert gleich 0. Ferner sei wieder vorausgesetzt, dass der Rang von $Df(x_0, \alpha_0)$ maximal ist. Dann gilt $D_\alpha f(x_0, \alpha_0) \neq 0$ und im eindimensionalen Fall ($n = 1$) ist die Kurve lokal durch x parametrisierbar. Außerdem folgt in diesem Fall aus der Kettenregel angewandt auf $0 = f(x, \alpha(x))$

$$0 = D_x f(x_0, \alpha_0) + D_\alpha f(x_0, \alpha_0) \alpha'(x_0) = D_\alpha f(x_0, \alpha_0) \alpha'(x_0). \quad (8.3.2)$$

Es folgt $\alpha'(x_0) = 0$. Gelte darüberhinaus

$$D_{xx} f(x_0, \alpha_0) \neq 0,$$

Dann ergibt sich wie in (8.3.2) aus der Kettenregel

$$\begin{aligned} 0 &= \frac{d^2}{dx^2} f(x, \alpha(x)) \Big|_{x=x_0} \\ &= \frac{d}{dx} [D_x f(x, \alpha(x)) + D_\alpha f(x, \alpha(x)) \alpha'(x)] \Big|_{x=x_0} \\ &= D_{xx} f(x_0, \alpha_0) + D_{\alpha x} f(x_0, \alpha_0) \alpha'(x_0) + D_{x\alpha} f(x_0, \alpha_0) \alpha'(x_0) + D_{\alpha\alpha} f(x_0, \alpha_0) \alpha'(x_0)^2 \\ &\quad + D_\alpha(x_0, \alpha_0) \alpha''(x_0) \\ &= D_{xx} f(x_0, \alpha_0) + D_\alpha(x_0, \alpha_0) \alpha''(x_0). \end{aligned}$$

Es folgt, dass $\alpha''(x_0) \neq 0$. Die Kurve $(x, \alpha(x))$ von Gleichgewichten hat folglich in x_0 ein lokales Extremum bezüglich α . Aus der Sicht des Parameters α bedeutet dies: für $\alpha > \alpha_0$ (oder $\alpha < \alpha_0$, je nach Vorzeichen von $\alpha''(x_0)$) gibt es zwei Gleichgewichte in der Nähe von x_0 , für $\alpha < \alpha_0$ (bzw. für $\alpha > \alpha_0$) dagegen keines.

Beispiel 8.3.2 Die Differentialgleichung in \mathbb{R}

$$\dot{x} = x^2 - \alpha, \quad \text{mit } \alpha \in \mathbb{R}, \quad (8.3.3)$$

besitzt für $\alpha > 0$ die Gleichgewichte $x^\pm = \pm\sqrt{\alpha}$, für $\alpha = 0$ nur $x = 0$, und für $\alpha < 0$ keine.

Eine solche Änderung der Dynamik in der Nähe von x_0 bei Variation des Parameters α in der Nähe von α_0 nennt man Sattel-Knoten-Verzweigung (bei α_0) – der Name wird verständlich, wenn wir diese Verzweigung bei höherdimensionalen Systemen untersuchen.

Zunächst wollen wir uns aber noch die Stabilitätseigenschaften der beiden Gleichgewichte für $\alpha > \alpha_0$ anschauen. Im Beispiel 8.3.2 ist wegen $D_x f(x^\pm, \alpha) = \pm 2\sqrt{\alpha}$ das Gleichgewicht x^+ instabil, x^- dagegen stabil.

Der Name "Sattel-Knoten-Verzweigung" wird verständlich, wenn wir das Vektorfeld (8.3.3) um Koordinaten erweitern, deren Dynamik nicht von α abhängt.

Beispiel 8.3.3 Wir betrachten das System

$$\begin{aligned} \dot{x} &= x^2 - \alpha \\ \dot{y} &= -y \end{aligned}$$

Offenbar besitzt dieses System für $\alpha > 0$ die beiden Gleichgewichte $(\pm\sqrt{\alpha}, 0)$. Das Gleichgewicht $(-\sqrt{\alpha}, 0)$ ist stabil (ein "Knoten"), $(\sqrt{\alpha}, 0)$ dagegen ist ein "Sattel", d.h. besitzt jeweils eine ein-dimensionale stabile und instabile Mannigfaltigkeit. Für $\alpha \rightarrow 0$ laufen die beiden Gleichgewichte auf der x -Achse im Ursprung zusammen, für $\alpha < 0$ besitzt das System keine Gleichgewichte mehr. Daher erklärt sich der Name Sattel-Knoten-Verzweigung.

Die Pitchfork-(Heugabel-)Verzweigung. Die Sattel-Knoten-Verzweigung ist die „typische“ Verzweigung, wenn man nur einen Parameter α im Vektorfeld variiert: gilt $D_x f(x_0, \alpha_0) = 0$, dann ist „typischerweise“ nicht gleichzeitig auch $D_\alpha f(x_0, \alpha_0) = 0$ oder $D_{xx} f(x_0, \alpha_0) = 0$ – außer in Spezialfällen. Ein solcher Spezialfall ist die Gleichung

$$\dot{x} = \alpha x - x^3.$$

Hier gilt tatsächlich $D_x f(0, 0) = 0$ und $D_{xx} f(0, 0) = 0$. Die Gleichung besitzt das Gleichgewicht $x = 0$ für alle α , sowie die beiden Gleichgewichte $x^\pm = \pm\sqrt{\alpha}$ für $\alpha > 0$. Offenbar verliert $x = 0$ bei $\alpha = 0$ seine Stabilität an die beiden anderen Gleichgewichte.

Die Hopf-Verzweigung. Charakteristisch für die Sattel-Knoten-Verzweigung ist, dass bei Variation des Parameters α ein reeller Eigenwert der Jacobi-Matrix $D_x f(x, \alpha)$ die Null durchläuft. Analog erhält man eine qualitative lokale Änderung der Dynamik (also eine Verzweigung), wenn unter Variation von α der Realteil eines Paares komplexer Eigenwerte von $D_x f(x, \alpha)$ die imaginäre Achse passiert.

Beispiel 8.3.4 Betrachten wir das System

$$\begin{aligned} \dot{x} &= -y - x(\alpha + x^2 + y^2) \\ \dot{y} &= x - y(\alpha + x^2 + y^2) \end{aligned}$$

mit dem reellen Parameter α . Offenbar ist $(0, 0)$ für alle α ein Gleichgewicht. Die x -Ableitung des Vektorfelds in $(0, 0)$ ist

$$\begin{bmatrix} -\alpha & -1 \\ 1 & -\alpha \end{bmatrix}$$

mit den Eigenwerten $-\alpha \pm i$. Für $\alpha > 0$ ist der Ursprung also stabil, für $\alpha < 0$ instabil. Für $\alpha < 0$ existiert in der Nähe des Ursprungs ein (stabiler) periodischer Orbit. Tatsächlich lässt sich die Existenz dieses periodischen Orbits leicht nachweisen, wenn man das System in Polarkoordinaten $x = r \cos \theta$, $y = r \sin \theta$ betrachtet:

$$\begin{aligned}\dot{r} &= -r(\alpha + r^2) \\ \dot{\theta} &= 1.\end{aligned}$$

Offenbar liegt für die r -Gleichung eine Pitchfork-Verzweigung bei $\alpha = 0$ vor, für $\alpha < 0$ hat diese Gleichung das stabile Gleichgewicht $r_0 = \sqrt{-\alpha}$, das dem stabilen periodischen Orbit des Gesamtsystems entspricht.

Die wichtigste Bedingung für das Vorliegen einer Hopf-Verzweigung in x^0 für $\alpha = \alpha^0$ ist, dass für Gleichgewichte $x(\alpha)$ von $\dot{x} = f(x, \alpha)$ mit $x(\alpha^0) = x^0$ die Linearisierung $D_x f(x, \alpha)$ ein Paar von komplexen Eigenwerten $\mu(\alpha), \overline{\mu(\alpha)}$ mit den folgenden Eigenschaften hat:

$$\mu(\alpha^0) = i\omega, \overline{\mu(\alpha^0)} = -i\omega \text{ mit } 0 \neq \omega \in \mathbb{R}, \text{ und } \frac{d}{d\alpha} \operatorname{Re} \mu(\alpha) |_{\alpha=\alpha^0} \neq 0.$$

Eine exakte Formulierung findet man zum Beispiel in Robinson [13, Chapter IV, Theorem 5.2] oder in Grüne/Junge [7, Satz 10.9], Aulbach [4, Abschnitt 7.11].

Literaturverzeichnis

- [1] D. ACHESON, *From Calculus to Chaos. An Introduction to Dynamics*, Oxford University Press, 1997.
- [2] H. AMANN, *Gewöhnliche Differentialgleichungen*, De Gruyter, Berlin, New York, 1995. 2. Auflage.
- [3] H. AMANN AND J. ESCHER, *Analysis II*, Birkhäuser, 2006.
- [4] B. AULBACH, *Gewöhnliche Differenzialgleichungen*, Spektrum Akademischer Verlag, 2004. Zweite Auflage.
- [5] W. L. BOOTHBY, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, Orlando, 1986. Second Edition.
- [6] M. BRAUN, *Differentialgleichungen und ihre Anwendungen*, Springer-Verlag, 1991.
- [7] L. GRÜNE AND O. JUNGE, *Gewöhnliche Differentialgleichungen. Eine Einführung aus der Perspektive der dynamischen Systeme*, Vieweg und Teubner, 2009.
- [8] W. HAHN, *Stability of Motion*, Springer-Verlag, 1967.
- [9] E. HAIRER, S. P. NORSETT, AND G. WANNER, *Solving Ordinary Differential Equations I, Nonstiff Problem*, Springer-Verlag, 1987.
- [10] ———, *Solving Ordinary Differential Equations I, Nonstiff Problem*, Springer-Verlag, 1987.
- [11] M. HIRSCH, S. SMALE, AND R. DEVANEY, *Differential Equations, Dynamical Systems and an Introduction to Chaos*, Elsevier, 2004.
- [12] H. KNOBLOCH AND F. KAPPEL, *Gewöhnliche Differentialgleichungen*, Teubner, Stuttgart, 1974.
- [13] C. ROBINSON, *Dynamical Systems. Stability, Symbolic Dynamics, and Chaos*, CRC Press, 1999. Second edition.
- [14] E. D. SONTAG, *Mathematical Control Theory*, Springer-Verlag, 1998. Second Edition (First Edition, 1990).
- [15] W. WALTER, *Gewöhnliche Differentialgleichungen*, Springer Verlag, 2000.