

Gewöhnliche Differentialgleichungen

Marius Müller

Wintersemester 2021/2022

Inhaltsverzeichnis

| | |
|--|------------|
| 1 Explizite Lösungstechniken | 8 |
| 1.1 Separation der Variablen | 8 |
| 1.2 Lineare Differentialgleichungen | 15 |
| 1.2.1 Lineare inhomogene Differentialgleichungen | 17 |
| 1.2.2 Lineare Systeme mit konstanten Koeffizienten | 22 |
| 1.2.3 Die Fundamentalmatrix | 30 |
| 1.2.4 Dekomplexifizierung | 34 |
| 1.2.5 DGLs höherer Ordnung | 37 |
| 1.2.6 Anwendung: Der harmonische Oszillator | 47 |
| 1.3 Integrale der Bewegung und weitere Tricks | 53 |
| 1.3.1 Energieerhaltung | 56 |
| 1.3.2 Anwendung: Das mathematische Pendel | 59 |
| 2 Differentialgleichungen im Anwendungskontext | 68 |
| 2.1 Das SIR-Pandemiemodell | 68 |
| 2.1.1 Erste Eigenschaften des Modells | 73 |
| 2.1.2 Der Höhepunkt der Pandemie | 78 |
| 2.1.3 Exkurs: Basisreproduktionszahl und Herdenimmunität | 84 |
| 3 Existenz, Eindeutigkeit, Globalität | 89 |
| 3.1 Der Satz von Picard-Lindelöf | 89 |
| 3.1.1 Das assoziierte Fixpunktproblem | 97 |
| 3.1.2 Beweis der globalen Version | 103 |
| 3.1.3 Beweis der lokalen Version | 112 |
| 3.2 Globalität und stetige Abhängigkeit | 125 |
| 3.2.1 Das Escape-Lemma | 125 |
| 3.2.2 Das Lemma von Gronwall | 133 |
| 3.3 Wohlgestelltheit | 142 |
| 4 Ausblick: Autonome Differentialgleichungen | 145 |

Einführung

Dieses Vorlesungsmanuskript beschäftigt sich mit der Thematik der ‘Gewöhnlichen Differentialgleichungen’. *Differentialgleichungen* (kurz: DGL) sind Gleichungen der Form

$$y'(t) = f(t, y(t)) \quad (t \in I), \quad (1)$$

wobei $f \in C^0(I \times G; \mathbb{R}^n)$, $G \subset \mathbb{R}^n$ ein Gebiet ist und I ein offenes Intervall ist. Im Falle $n > 1$ spricht man auch von einem *Differentialgleichungssystem*.

Solche Gleichungen beschreiben — wie wir noch sehen werden — allerlei Zusammenhänge in Natur und Technik. Eine Funktion $y \in C^1(I; \mathbb{R}^n)$, die (1) für alle $t \in I$ erfüllt heißt *Lösung* der Differentialgleichung (1) auf I .

Beispiel 1.

$$y'(t) = t^3 \quad (t \in \mathbb{R}), \quad (2)$$

d.h. $n = 1$, $I = \mathbb{R}$, $G = \mathbb{R}$ und $f(t, z) = t^3$.

In diesem Fall können wir mithilfe der Grundlagenvorlesungen alle Lösungen bestimmen: Der Hauptsatz der Differential- und Integralrechnung besagt, dass jede Lösung eine Stammfunktion von t^3 sein muss. Alle solche Stammfunktionen sind gegeben durch

$$y(t) = \frac{1}{4}t^4 + C \quad (t \in \mathbb{R}) \quad (3)$$

wobei $C \in \mathbb{R}$ eine Konstante ist. Die Menge aller Lösungen von (2) heißt *allgemeine Lösung* von (2). Man könnte also sagen, dass die allgemeine Lösung durch alle Funktionen y aus Gleichung (3) gegeben ist.

In Kapitel 1 (‘Explizite Lösungstechniken’) wollen wir uns damit beschäftigen, wie man die allgemeine Lösung von komplizierteren Gleichungen explizit mit elementaren Funktionen ausdrücken kann.

Es ist allerdings so, dass man nicht alle Gleichungen explizit lösen kann. Gerade bei *Systemen von Gleichungen* (d.h. im Fall $n > 1$) ist an eine explizite Lösung nicht zu denken. Doch auch vergleichsweise einfache Gleichungen haben keine explizite Lösung.

Beispiel 2.

$$y'(t) = e^{-t^2} \quad (t \in \mathbb{R}). \quad (4)$$

Wiederum müsste jede Lösung eine Stammfunktion von e^{-t^2} sein. Allerdings lässt sich keine solche Stammfunktion mit elementaren Funktionen ausdrücken.

Wir müssen daher Methoden entwickeln, um Lösungen zu beschreiben, auch wenn wir sie nicht kennen. So können wir zum Beispiel sofort zeigen, dass jede Lösung von (4) monoton wachsend sein muss, denn es gilt $y'(t) > 0$ für alle $t \in \mathbb{R}$. Es ist erstaunlich, wie viele qualitative Eigenschaften ohne explizite Lösung diskutiert werden können.

In Kapitel 2 nähern wir uns der Diskussion zunächst aus dem Anwendungskontext. Aus aktuellem Anlass spezialisieren wir uns auf das RSI-Modell zur Beschreibung von Pandemien.

Wir werden lernen, Antworten auf anwendungsrelevante Fragen aus der Gleichung selbst zu erhalten: Wie viele Menschen werden maximal *gleichzeitig infiziert* sein? Ab wann besteht *Herdenimmunität*? Was bringt ein *Lockdown*?

In Kapitel 3 nähern wir uns der Beschreibung der Eigenschaften zunächst von einem theoretischen Standpunkt: Wir fragen uns, ob eine gegebene Differentialgleichung eine Lösung hat und – wenn ja – wie viele Lösungen es gibt. Außerdem wollen wir wissen, ob man stets eine *globale* Lösung finden kann ($I = \mathbb{R}$) oder ob es Lösungen gibt, die nach einer gewissen Zeit ‘aufhören’, zu existieren. Für all diese Fragestellungen betrachtet man hauptsächlich *Anfangswertprobleme* (kurz: AWP), d.h. man sucht Lösungen mit einem vorgeschriebenen Anfangswert, also von

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I), \\ y(t_0) = y_0, \end{cases} \quad (5)$$

für ein fixiertes $t_0 \in I$ und $y_0 \in G$. Der Anfangswert ist sehr wichtig, denn die Antwort auf die Frage nach Existenz, Eindeutigkeit und Globalität hängt — wie wir sehen werden — stark von dem Anfangswert ab! Stellen wir uns mal folgendes vor: Unsere DGL beschreibt die Bewegung eines Lavapartikels bei einem Vulkanausbruch. Wo unser Lavapartikel sich nach einer gewissen Zeit befindet hängt maßgeblich davon ab, wo es gestartet hat. Falls es tief in der Erde gestartet hat, befindet es sich vielleicht noch im Vulkaninneren. Wurde es aber bei der ersten Eruption bereits ausgespuckt, so wird es mittlerweile an der Oberfläche sein.

Diese Beobachtung zeigt, dass der Anfangswert als Teil des Problems wahrgenommen werden sollte. Wir werden ferner sehen, dass bei sehr vielen AWPs die Eindeutigkeitsfrage mit ‘Ja’ beantwortet werden kann. Ein gutes Beispiel dafür kennt man vielleicht noch aus den Analysis-Grundlagenvorlesungen:

Beispiel 3. Das Anfangswertproblem

$$\begin{cases} y'(t) = y(t) & (t \in \mathbb{R}) \\ y(0) = 1 \end{cases} \quad (6)$$

hat die Lösung $y(t) = e^t$. Dies könnte man durch Einsetzen einfach nachrechnen (HA). Vielleicht haben Sie die Exponentialfunktion sogar so definiert.

BEHAUPTUNG: Es gibt auch keine weiteren Lösungen.

BEGRÜNDUNG: Sei $\bar{y} \in C^1(\mathbb{R})$ eine weitere Lösung von (6). Definiere dann $g(t) := e^{-t}\bar{y}(t)$ für $t \in \mathbb{R}$. Man sieht mit der Produktregel

$$g'(t) = -e^{-t}\bar{y}(t) + e^{-t}\bar{y}'(t) \stackrel{(6)}{=} -e^{-t}\bar{y}(t) + e^{-t}\bar{y}(t) = 0. \quad (7)$$

Damit ist g konstant auf \mathbb{R} und somit gilt $g(t) = g(0) = 1$ für alle $t \in \mathbb{R}$. Mit der Definition von g schließen wir also $e^{-t}\bar{y}(t) = 1$, was umgestellt nichts anderes heißt als $\bar{y}(t) = e^t$. Wir sehen also, dass das AWP eine eindeutige, globale Lösung besitzt.

Die Existenz- und Eindeutigkeitsfrage ist auch im Anwendungskontext sehr wichtig. Es ist ja schließlich so, dass ein Anfangswertproblem stets einen Ablauf eines Naturphänomens

beschreibt. Falls es eine eindeutige Lösung gibt, so wissen wir: Unser Modell sagt eindeutig voraus, wie das Phänomen ablaufen wird. Gibt es hingegen mehrere Lösungen, so kann nur eine davon das beschreiben, was die Natur wirklich macht. In einem solchen Fall muss man sich fragen, ob das gegebene Modell das Naturphänomen wirklich vollständig beschreibt.

Der *Satz von Picard-Lindelöf* formuliert eine hinreichende Bedingung für die Existenz einer eindeutigen Lösung. Diese Bedingung ist jedoch nicht immer erfüllt:

Beispiel 4. Das Anfangswertproblem

$$\begin{cases} y'(t) = 2\sqrt{|y(t)|} & (t \in \mathbb{R}) \\ y(0) = 0 \end{cases} \quad (8)$$

hat mehrere Lösungen — sogar gleich unendlich viele:

BEHAUPTUNG: Für alle $a \geq 0$ ist

$$y_a(t) := \begin{cases} 0 & t \in (-\infty, a] \\ (t - a)^2 & t \in (a, \infty) \end{cases} \quad (9)$$

eine Lösung. Überzeugen Sie sich davon, dass y_a in $C^1(\mathbb{R})$ liegt und (8) erfüllt ist (HA).

Auch die Globalitätsfrage ist anwendungsrelevant: Im Anwendungskontext bedeutet Nicht-Globalität einer Lösung meistens, dass ein *singuläres Phänomen* auftritt. Zum Beispiel könnte man einen hängenden Wassertropfen an einem undichten Wasserhahn beschreiben, der erst immer runder wird und dann irgendwann heruntertropft. Das singuläre Phänomen ist hier dann das ‘Heruntertropfen’.

Wieder einmal ist es so, dass auch bei ‘einfach aussehenden’ Gleichungen die Lösungen nicht zwingend global sind. Wir werden aber auch hier Kriterien für die Globalität formulieren. Eine große Hilfe hierbei wird das (vielleicht schon bekannte) *Lemma von Gronwall* sein.

Beispiel 5. Für festes $c \in \mathbb{R}$ hat das Anfangswertproblem

$$\begin{cases} y'(t) = y(t)^2 & (t \in \mathbb{R}) \\ y(0) = c \end{cases} \quad (10)$$

nur im Fall $c = 0$ eine eindeutige, globale Lösung (nämlich $y(t) := 0$ für alle $t \in \mathbb{R}$). Für alle $c \neq 0$ gibt es eine eindeutige Lösung, die aber nicht auf ganz \mathbb{R} definiert ist — sie existiert nur auf einem kleineren Intervall. Daher ist die Bedingung ‘($t \in \mathbb{R}$)’ aus (11) meistens nicht erfüllt. Details werden wir im Laufe der Vorlesung noch ausarbeiten. Wir werden sehen, dass man stets eine *größtmögliche* Lösung finden kann, d.h. es gibt ein *maximales* offenes Intervall I_{max} , das den Anfangswert 0 enthält, sodass

$$\begin{cases} y'(t) = y(t)^2 & (t \in I_{max}) \\ y(0) = c \end{cases} \quad (11)$$

eine eindeutige Lösung besitzt. Wir werden auch sehen was am Rand von I_{max} mit der Lösung “schiefgehen” muss: Sie explodiert ins Unendliche.

Notationsverzeichnis

- $C^k(A; \mathbb{R}^n)$. Bezeichnet die Menge aller k -mal stetig differenzierbaren Funktionen von A nach \mathbb{R}^n . Hierbei ist $A \subset \mathbb{R}^m$ stets eine offene Menge.
- $\frac{d}{dt}y(t)$ oder $y'(t)$. Bezeichnet beides gleichermaßen die Ableitung von y nach t , d.h. $y'(t) = \lim_{h \rightarrow 0} \frac{y(t+h)-y(t)}{h}$.
- Für $f \in C^1(\mathbb{R}^m; \mathbb{R}^n)$ bezeichnet $Df(z) := (\partial_{z_j} f^i(z))_{i,j} \in \mathbb{R}^{m \times n}$ die *Jacobi-Matrix* von f . Manchmal schreiben wir auch $D_z f(z)$, gerade, wenn f noch von weiteren Variablen abhängt.
- $\text{dist}(y, A)$ bezeichnet die Distanz eines Vektors $y \in \mathbb{R}^n$ zu einer Menge $A \subset \mathbb{R}^n$, d.h.

$$\text{dist}(y, A) := \inf\{|x - y| : x \in A\}.$$

- e_1, \dots, e_n bezeichnet stets die kanonischen Basisvektoren des \mathbb{R}^n .
- G . Bezeichnet stets ein Teilgebiet von \mathbb{R}^n . Unter einem Gebiet verstehen wir eine offene, zusammenhängende Menge.
- (HA). Die so gekennzeichneten Aussagen sind den Lesenden als Übungsaufgabe überlassen.
- I . Bezeichnet stets ein offenes Intervall $I = (a, b)$. Ist $b < a$ so ist $I = \emptyset$. Die leere Menge ist in unserer Konvention ein Intervall.
- $\int^t f(s) ds$ bezeichnet das *unbestimmte Integral* über f . Das heißt: Man bildet eine Stammfunktion von f und wertet sie an t aus. Der Wert dieses Ausdrucks ist nur bis auf eine Konstante eindeutig bestimmt. Wenn diese Konstante explizit aufgeführt wird, so ist mit $\int^t f(s) ds$ stets eine beliebige aber feste Stammfunktion von f gemeint.
- $\ker(A)$ bezeichnet stets den Kern der Matrix (oder einer linearen Abbildung) A .
- $\mathbb{R}^{m \times n}/\mathbb{C}^{m \times n}$ bezeichnet die Menge aller (reellwertigen/komplexwertigen) Matrizen mit m Zeilen und n Spalten.
- Re/Im bezeichnet Realteil bzw. Imaginärteil einer komplexen Zahl. Wie werden es auch für Vektoren benutzen, in dem Fall ist Realteil und Imaginärteil komponentenweise zu verstehen.
- $\langle \cdot, \cdot \rangle$ bezeichnet stets das Euklidische Skalarprodukt zweier Vektoren in \mathbb{R}^n oder in \mathbb{C}^n , d.h. $\langle v, w \rangle = \sum_{i=1}^n v_i \overline{w_i}$, wobei der Querbalken das Konjugiertkomplexe notieren soll. $|\cdot|$ bezeichnet stets die Euklidische Norm eines Vektors in \mathbb{R}^n .
- $y^{(n)}(t)$ bezeichnet die n -te Ableitung der Funktion $y \in C^n(I; \mathbb{R})$.
- \mathbb{Z} . Bedeutet “zu zeigen”. Hiermit erläutern wir in einem Beweis, was als nächstes gezeigt werden soll. Dies unterteilt den Beweis dann in verschiedene Abschnitte.

Dankeschön!

Ich danke meiner Assistentin Saskia Glaffig für die ausführlichen und gut erklärten TeX-Lösungen der Übungen. Außerdem danke ich meiner Tutorin Simone Hermann für die aufmerksame Aufnahme der Vorlesungsvideos und das Engagement in den Tutoraten. Ferner danke ich Dominik Zimmermann für die vielen Korrekturen und Vorschläge zum Skript.

1 Explizite Lösungstechniken

Hier lernen wir Verfahren, um Gleichungen mit einer bestimmten Form explizit zu lösen, d.h. die allgemeine Lösung zu bestimmen. Wir gehen im Folgenden immer von der Form

$$y'(t) = f(t, y(t)) \quad (t \in I) \quad (12)$$

aus. Hierbei ist wie üblich $f \in C^0(I \times G; \mathbb{R}^n)$ für ein Gebiet $G \subset \mathbb{R}^n$ gegeben und $y \in C^1(I, \mathbb{R}^n)$ gesucht.

Wir machen spezielle Annahmen an f unter denen wir die Gleichung explizit auflösen können.

Wir stellen hierbei stets *instruktive Lösungswege* vor, d.h. Möglichkeiten, sich klarzumachen, wie man auf die Lösung kommt.

1.1 Separation der Variablen

Die folgende Lösungstechnik behandelt *skalare* Differentialgleichungen, d.h. $n = 1$. Zusätzlich nehmen wir an, dass die Funktion f *trennbar* ist. Genauer:

Definition 6. TRENNBARE VERÄNDERLICHEN.

Eine DGL heißt *Differentialgleichung mit trennbaren Veränderlichen*, falls sie die Gestalt

$$y'(t) = g(t)h(y(t)) \quad (t \in I) \quad (13)$$

hat, wobei $g \in C^0(I)$ und $h \in C^0(J)$ für ein offenes Intervall $J \subset \mathbb{R}$. Insbesondere heißt das, dass $f(t, z) = g(t)h(z)$ für alle $(t, z) \in I \times J$.

Man beachte, dass ein eindimensionales Gebiet G gerade ein Intervall ist, daher wurde hier die Notation J anstatt I verwendet.

Das folgende instruktive Beispiel soll zeigen, wie wir eine DGL mit trennbaren Veränderlichen lösen können.

Beispiel 7. EXPONENTIELLES WACHSTUM.

$$y'(t) = \alpha y(t) \quad (t \in \mathbb{R}). \quad (14)$$

Hier können wir $g(t) := \alpha$ und $h(z) := z$ nehmen und erhalten eine DGL mit trennbaren Veränderlichen. Wir wollen die allgemeine Lösung bestimmen.

INSTRUKTIVER LÖSUNGSWEG. Sicherlich ist die Nullfunktion eine Lösung, wie man leicht nachrechnet. OBdA müssen wir also nur nach Lösungen suchen, die nicht die Nullfunktion sind. Wir nehmen nun einfach mal an, dass $y(t) > 0$ für alle $t \in \mathbb{R}$. Dass diese Annahme gerechtfertigt ist, können wir zu dem gegebenen Zeitpunkt nicht begründen. Wenn wir aber am Ende auf eine Lösung kommen, die diese Annahme erfüllt, haben wir trotzdem eine Lösung gefunden und können uns freuen. Unter unserer Annahme können wir (14) umschreiben zu

$$\frac{y'(t)}{y(t)} = \alpha \quad (t \in \mathbb{R}). \quad (15)$$

Jetzt kann uns mit der ‘Kettenregel rückwärts’ folgendes auffallen: Es gilt

$$\frac{d}{dt} \log y(t) = \frac{y'(t)}{y(t)}. \quad (16)$$

Zusammen mit Gleichung (15) heißt das, dass

$$\frac{d}{dt} \log y(t) = \alpha \quad (t \in \mathbb{R}). \quad (17)$$

Über diese Gleichung können wir integrieren. Nach dem Hauptsatz der Differential- und Integralrechnung gilt daher

$$\log y(t) = \alpha t + \beta \quad (t \in \mathbb{R}) \quad (18)$$

für eine beliebige Konstante $\beta \in \mathbb{R}$. Wir können nach $y(t)$ auflösen und erhalten

$$y(t) = e^{\alpha t + \beta} = e^\beta e^{\alpha t} \quad (t \in \mathbb{R}). \quad (19)$$

Definieren wir $c := e^\beta > 0$, so haben wir also

$$y(t) = ce^{\alpha t} \quad (t \in \mathbb{R}). \quad (20)$$

Dass dies für alle beliebigen $c > 0$ wirklich eine Lösung ist, könnten wir durch Einsetzen nachrechnen (HA). Wir sehen: Obwohl wir auf dem Weg eine Annahme getroffen haben, könnten wir trotzdem noch Lösungen finden. Allerdings kann es durchaus sein, dass wir aufgrund der Annahme nur einen Teil aller Lösungen finden. So auch hier: Man sieht nämlich beim Einsetzen schnell, dass $c > 0$ gar nicht benötigt wird! In der Tat: Für alle $c \in \mathbb{R}$ definiert

$$y(t) = ce^{\alpha t} \quad (t \in \mathbb{R}) \quad (21)$$

eine Lösung. Haben wir jetzt alle Lösungen gefunden?

BEHAUPTUNG: Die allgemeine Lösung von (14) ist gegeben durch

$$\mathbb{L} := \{y \in C^1(\mathbb{R}) \mid \exists c \in \mathbb{R} : y(t) = ce^{\alpha t} \forall t \in \mathbb{R}\} \quad (22)$$

Dass jedes Element aus \mathbb{L} eine Lösung ist, rechnet man durch Einsetzen nach. Es sei nun $y_* \in C^1(\mathbb{R})$ eine beliebige Lösung von (14). Wir müssen zeigen, dass $y_* \in \mathbb{L}$. Dazu definieren wir $g(t) := e^{-\alpha t} y_*(t)$, $t \in \mathbb{R}$ und rechnen

$$g'(t) = -\alpha e^{-\alpha t} y_*(t) + e^{-\alpha t} y'_*(t) \stackrel{(14)}{=} -\alpha e^{-\alpha t} y_*(t) + e^{-\alpha t} \alpha y_*(t) = 0. \quad (23)$$

Damit ist $g \equiv \text{const.}$, d.h. es gibt $c_* \in \mathbb{R}$ so, dass $e^{-\alpha t} y_*(t) = c_*$ für alle $t \in \mathbb{R}$. Stellt man das um, so erhält man $y_*(t) = c_* e^{\alpha t}$.

Hat man einmal die allgemeine Lösung gefunden, so kann man auch sämtliche Anfangswertprobleme lösen.

Beispiel 8. Wir betrachten das Anfangswertproblem

$$\begin{cases} y'(t) = \alpha y(t) & (t \in \mathbb{R}) \\ y(1) = 2 \end{cases} \quad (24)$$

Da wir im vorigen Beispiel die allgemeine Lösung bestimmt haben wissen wir, dass es $c \in \mathbb{R}$ gibt, sodass $y(t) = ce^{\alpha t}$ für alle $t \in \mathbb{R}$. Nun bestimmen wir c so, dass der Anfangswert angenommen wird. Dazu setzen wir

$$2 \stackrel{!}{=} y(1) = ce^{\alpha t} \Big|_{t=1} = ce^\alpha. \quad (25)$$

Wir stellen um und finden $c = 2e^{-\alpha}$. Alles in allem gilt dann

$$y(t) = 2e^{-\alpha} e^{\alpha t} = 2e^{\alpha(t-1)}. \quad (26)$$

Anwendung 9. EXPONENTIELLES WACHSTUM IN DER PRAXIS. Wir beschreiben den Populationsbestand einer großen Hasenpopulation. Täglich werden neue Hasen geboren. Wie viele Hasen täglich geboren werden — so nehmen wir an — proportional dazu, wie viele Hasen schon da sind. Die Proportionalitätskonstante, die wir α nennen, heißt *Geburtenrate*. Bezeichnet also $y(t)$ die Anzahl der Hasen zum Zeitpunkt t , so werden zwischen den Zeitpunkten ‘ t ’ und ‘ $t + 1$ Tag’ also in etwa $\alpha y(t)$ Hasen geboren.

Wir können auch andere Zeitintervalle als ‘1 Tag’ betrachten. Gehen wir davon aus, dass der Geburtszeitpunkt zufällig verteilt ist, so würden wir erwarten, dass zwischen den Zeitpunkten ‘ t ’ und ‘ $t + h$ ’ ungefähr $\alpha y(t)h$ Hasen geboren werden. Hierbei muss h allerdings ziemlich klein sein: Ist h nicht klein, so kann man nicht annehmen, dass der Populationsbestand zwischen t und $t + h$ ungefähr durch $y(t)$ gegeben ist. In der Tat: bei großen Werten von h , z.B. viel größer als 1 Tag, würde sich die Population zwischen t und $t + h$ so stark verändern, dass $\alpha y(t)$ die täglichen Geburten in der Nähe von $t + h$ unterschätzt. In der Nähe von $t + h$ müsste man eigentlich eher mit $\alpha y(t + h)$ Geburten pro Tag rechnen. Für den gesamten Populationsbestand heißt das dann folgendes:

$$\underbrace{y(t+h)}_{\text{Populationsbestand bei } t+h} \approx \underbrace{y(t)}_{\text{Populationsbestand bei } t} + \underbrace{\alpha y(t)h}_{\text{neugeborene Hasen}}. \quad (27)$$

Die Approximation ist umso besser, wenn h klein ist, denn in kleinen Zeitintervallen verändert sich die Population kaum und $\alpha y(t)h$ ist ein guter Schätzer für die neugeborenen Hasen. Stellen wir (27) um, so erhalten wir für kleine h

$$\frac{y(t+h) - y(t)}{h} \approx \alpha y(t). \quad (28)$$

Wie oben bereits diskutiert können wir h gegen 0 schicken und im Grenzwert für $h \rightarrow 0$

Gleichheit erwarten. Wir erhalten

$$y'(t) = \alpha y(t). \quad (29)$$

Man könnte nun anmerken, dass es neben Geburten auch Sterbefälle unter den Hasen gibt. Man könnte wiederum annehmen, dass diese proportional zu der Anzahl der Hasen ist. Man spricht hierbei von einer *Sterberate* γ . Den Populationsbestand würde man dann wie oben modellieren. Für kleine Zeitdifferenzen h erhält man

$$\underbrace{y(t+h)}_{\text{Population bei } t+h} \approx \underbrace{y(t)}_{\text{Population bei } t} + \underbrace{\alpha y(t)h}_{\text{neugeborene Hasen}} - \underbrace{\gamma y(t)h}_{\text{verstorbene Hasen}}. \quad (30)$$

Verfährt man wie oben, kommt man auf die Differentialgleichung

$$y'(t) = (\alpha - \gamma)y(t). \quad (31)$$

Die allgemeine Lösung ist dann (siehe Beispiel 7 (p. 8) mit $\alpha - \gamma$ statt α) gegeben durch

$$y(t) = ce^{(\alpha-\gamma)t} \quad (c \in \mathbb{R}) \quad (32)$$

Man sieht: Ist $\gamma > \alpha$ so gilt stets

$$\lim_{t \rightarrow \infty} y(t) = 0. \quad (33)$$

Das zeigt, dass die Hasenpopulation in diesem Fall langfristig aussterben wird. Dies liegt auch auf der Hand, denn die Sterberate übersteigt hier die Geburtenrate. Die Differenz $\alpha - \gamma$ nennen wir auch *relative Geburtenrate*. Ist $\alpha - \gamma < 0$ so sprechen wir auch von einem *exponentiellen Zerfall* (anstatt von exponentiellem Wachstum).

Um die instruktive Lösungstechnik aus Beispiel 7 (p. 8) einzuüben betrachten wir noch ein anderes Beispiel, welches wir auch schon eingangs als Paradebeispiel für Nicht-Globalität von Lösungen diskutiert haben

Beispiel 10. NICHT-GLOBALE LÖSUNGEN. Wir betrachten wie in Beispiel 7 (p. 8)

$$y'(t) = y(t)^2 \quad (t \in I) \quad (34)$$

wobei wir hier I zunächst nicht spezifizieren wollen. Wir sehen, dass die Nulllösung $y(t) = 0$ eine globale Lösung ist. Man findet aber auch noch andere Lösungen:

INSTUKTIVER LÖSUNGSWEG: Nehmen wir wieder an, dass $y(t) \neq 0$ für alle $t \in I$ und formen um:

$$\frac{y'(t)}{y(t)^2} = 1 \quad (t \in I). \quad (35)$$

In Beispiel 7 (p. 8) haben wir dann mit einer ‘Kettenregel rückwärts’ die linke Seite als Ableitung eines anderen Ausdrucks umgeschrieben. Das wird immer so funktionieren. Allerdings ist es oft etwas schwer zu sehen, was der andere Ausdruck ist. Hier kann man

sich Abhilfe schaffen, indem man das unbestimmte Integral bildet

$$\int^t \frac{y'(s)}{y(s)^2} ds = t + C \quad (36)$$

für eine Integrationskonstante $C \in \mathbb{R}$. Mit ‘ \int^t ’ meinen wir stets, dass wir eine Stammfunktion bilden wollen und diese bei t auswerten wollen. Wir können nun $u = y(s)$ substituieren und erhalten (mit $du = y'(s)ds$)

$$\int^{y(t)} \frac{1}{u^2} du = t + C. \quad (37)$$

Man beachte: Da wir die Substitution durchgeführt haben, müssen wir die Stammfunktion auch bei der substituierten Variable $y(t)$ auswerten. Das entspricht der Tatsache, dass man bei einer Substitution die Grenzen ‘mitsubstituieren’ muss. Jetzt können wir die Stammfunktion bilden:

$$\left[-\frac{1}{u} \right]^{y(t)} = t + C. \quad (38)$$

Wir erhalten also

$$-\frac{1}{y(t)} = t + C \Leftrightarrow y(t) = -\frac{1}{t + C} \quad (39)$$

Somit hat unser Weg wiederum auf eine explizite Lösungsformel geführt. Man rechnet auch leicht nach, dass es sich für jedes $C \in \mathbb{R}$ wirklich um eine Lösung handelt (HA). Allerdings sehen wir: Jede Lösung, die wir gefunden haben, hat bei $t = -C$ eine Definitionslücke. Daher ist keine dieser Lösungen global definiert. Wir können nun das Anfangswertproblem

$$\begin{cases} y'(t) = y(t)^2 & (t \in I_{max}) \\ y(0) = 2 \end{cases} \quad (40)$$

betrachten. Hierbei heißt ‘ I_{max} ’ einfach, dass das größtmögliche Intervall um den Anfangswert 0 genommen werden soll, auf dem die Lösung definiert ist. Nehmen wir die Darstellung aus (39), so können wir berechnen

$$2 \stackrel{!}{=} y(0) = -\frac{1}{t + C} \Big|_{t=0} = -\frac{1}{C} \quad (41)$$

Dies formt sich um zu $C = -\frac{1}{2}$. Dadurch erhalten wir die explizite Lösungsformel

$$y(t) = -\frac{1}{t - \frac{1}{2}} = \frac{1}{\frac{1}{2} - t}. \quad (42)$$

Diese Lösung ist auf $(-\infty, \frac{1}{2})$ definiert — ein Intervall, was den Anfangswert 0 enthält. Bei $\frac{1}{2}$ lässt sich die Lösung nicht weiter fortsetzen, weil sie dort eine Definitionslücke hat. BEHAUPTUNG: Das Anfangswertproblem (40) hat die *maximale* Lösung $y(t) = \frac{1}{\frac{1}{2}-t}$, definiert für $t \in I_{max} := (-\infty, \frac{1}{2})$.

BEGRÜNDUNG: Wie bereits gesagt, rechnet man leicht nach, dass das oben gegebene y eine Lösung ist. Das Problem: Es könnte ja noch weitere Lösungen geben, denn wir haben beim Lösungsweg ja eine Annahme getroffen – nämlich dass $y(t) \neq 0$ für alle $t \in \mathbb{R}$. Diese weiteren Lösungen könnten auch theoretisch auf einem Intervall definiert sein, was größer ist als I_{max} . In dem Fall wäre die oben genannte Lösung nicht die *maximale* Lösung. Später werden wir ein einfaches Argument kennenlernen, welches solche weiteren Lösungen ausschließt. An dieser Stelle müssen wir uns allerdings mit einem komplizierterem Argument begnügen, was den Lesenden als die folgende Übungsaufgabe überlassen ist.

Übung 11. Es sei $y \in C^1(I)$ eine Lösung von

$$y'(t) = y(t)^2 \quad (t \in I). \quad (43)$$

Zeigen Sie: Gibt es ein $t_0 \in I$ mit $y(t_0) \neq 0$, so gilt $y(t) \neq 0$ für alle $t \in I$.

HINWEIS: In einer Umgebung von t_0 gilt $\frac{d}{dt} \left(\frac{1}{y(t)} \right) = -1$, also

$$\frac{1}{y(t)} - \frac{1}{y(t_0)} = t_0 - t. \quad (44)$$

Betrachten Sie $t_1 := \inf\{t \in I : t \geq t_0, y(t) = 0\}$. Nehmen Sie an, dass $t_1 \in I$ und zeigen Sie, dass (44) in (t_0, t_1) erfüllt ist. Betrachten Sie (44) für $t \rightarrow t_1-$ und folgern Sie einen Widerspruch. Insbesondere ist also t_1 am Rand von I und daher $y(t) \neq 0$ für alle $t \in I : t \geq t_0$. Verfahren Sie analog mit $t_2 := \sup\{t \in I : t \leq t_0, y(t) = 0\}$.

SINN DER ÜBUNG: Wir haben mit dieser Übung hergeleitet, dass alle Lösungen des AWP (40) aus dem vorigen Beispiel die Annahme ‘ $y(t) \neq 0$ für alle $t \in I'$ erfüllen müssen. Mit diesem zuätzlichen Wissen sind die Gleichungen (35)–(39) Äquivalenzumformungen. Es kommen also keine weiteren Lösungen infrage als diejenigen, die wir im vorigen Beispiel gefunden haben. Das zeigt die Behauptung, dass die von uns gefundene Lösung die *maximale* Lösung ist. Das vorige Beispiel ist damit abgeschlossen.

Die Beweistechnik aus dem Hinweis, wo wir uns geschickte Infima definieren, wird übrigens *Prinzip des kleinsten Verbrechers* genannt.

Übung 12. LOGISTISCHES WACHSTUM. Betrachten Sie die Anfangswertproblem

$$\begin{cases} y'(t) = y(t)(1 - y(t)) & (t \in \mathbb{R}), \\ y(0) = y_0, \end{cases} \quad (45)$$

wobei $y_0 \in (0, 1)$ beliebig.

- (a) Bestimmen Sie eine explizite Lösung. Fertigen Sie eine Zeichnung an.
- (b) Zeigen Sie: Jede Lösung erfüllt $y(t) \in (0, 1)$ für alle $t \in \mathbb{R}$.
- (c) Folgern Sie: Zu jedem $y_0 \in (0, 1)$ gibt es eine eindeutige globale Lösung.

- (d) Zeigen Sie, dass die Lösung streng monoton ist und einen eindeutigen Wendepunkt $t_0 \in \mathbb{R}$ mit $y(t_0) = \frac{1}{2}$ besitzt.

TIPP: Es gibt hier zwei Lösungsmöglichkeiten. Eine mit der expliziten Formel aus Teilaufgabe (c) und eine, in der man direkt aus der Gleichung abliest.

SINN DER ÜBUNG: Logistisches Wachstum (allgemeiner beschrieben durch die Gleichung $y'(t) = \alpha(\beta - y(t))y(t)$, $\alpha, \beta > 0$) beschreibt eine Population, die sich aber nur bis zu einem gewissen Grad vermehren kann, weil bei einer zu großen Population z.B. die Nahrung knapp wird. Falls $y(t) \approx 0$ verhält sich die Population beim Zeitpunkt t fast wie eine Population mit relativer Geburtenrate $\alpha\beta$. Wird $y(t)$ allerdings größer, so wird die Geburtenrate bis auf 0 gedrückt. Die obige Übung zeigt, dass in diesem Modell die Population nicht beliebig anwachsen kann.

Am Ende dieses Abschnittes fassen wir die wichtigste Erkenntnis, (nämlich den instruktiven Lösungsweg in den Beispielen 7 (p. 8) und 10 (p. 11)) in einem Satz zusammen.

Satz 13. Es seien $g \in C^0(I)$, $h \in C^0(\mathbb{R})$ so, dass $h(z) \neq 0$ für alle $z \in \mathbb{R}$. Ferner sei die Differentialgleichung

$$y'(t) = g(t)h(y(t)) \quad (t \in I) \quad (46)$$

gegeben. Sei G eine Stammfunktion von g und H eine Stammfunktion von $\frac{1}{h}$. Dann ist H invertierbar und falls $G(I) \subset H(\mathbb{R})$ ist

$$y(t) := H^{-1}(G(t)) \quad (t \in I) \quad (47)$$

eine Lösung von (46).

Beweis. Zunächst müssen wir zeigen, dass H invertierbar ist. Hierzu sehen wir zunächst, dass $\frac{1}{h(z)} \neq 0$ für alle $z \in \mathbb{R}$. Nach dem Zwischenwertsatz gilt dann entweder $\frac{1}{h(z)} > 0$ für alle $z \in \mathbb{R}$ oder $\frac{1}{h(z)} < 0$ für alle $z \in \mathbb{R}$. Falls $\frac{1}{h} > 0$, so gilt nach dem Haupsatz der Differential- und Integralrechnung

$$H'(z) = \frac{1}{h(z)} > 0 \quad \forall z \in \mathbb{R}. \quad (48)$$

Hieraus folgt, dass H streng monoton wachsend ist. Damit ist es invertierbar und außerdem gilt (nach einem Satz aus der Analysis) $H^{-1} \in C^1(H(\mathbb{R}))$. Analog folgert man die Invertierbarkeit auch wenn $\frac{1}{h} < 0$. In jedem Fall liegt die in (47) definierte Funktion y in $C^1(I)$. Ferner gilt für alle $t \in I$, dass

$$H(y(t)) = G(t). \quad (49)$$

Differenzieren wir beide Seiten nach t , so erhalten wir mit der Kettenregel $H'(y(t))y'(t) =$

$G'(t)$ für alle $t \in I$. Mit der Definition von G und H gilt

$$\frac{1}{h(y(t))}y'(t) = g(t) \quad \forall t \in I. \quad (50)$$

Stellen wir dies um, so erhalten wir $y'(t) = g(t)h(y(t))$, wie gewünscht.

Separation der Variablen lässt sich auch für Differential-Ungleichungen verwenden — In der Folgenden Übung sehen wir zum Beispiel, was *superexponentielles Wachstum* bedeutet.

Übung 14. DIFFERENTIALUNGLEICHUNGEN UND EXPONENTIELLES WACHSTUM.

Gegeben sei das Anfangswertproblem

$$\begin{cases} y'(t) = f(y(t)) & (t \in \mathbb{R}) \\ y(0) = y_0 \end{cases} \quad (51)$$

wobei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion mit $\frac{f(z)}{z} \geq c$ für alle $z \in \mathbb{R} \setminus \{0\}$ und $y_0 > 0$. Zeigen Sie, dass

$$y(t) \geq y_0 e^{ct} \quad \forall t \in (0, \infty). \quad (52)$$

1.2 Lineare Differentialgleichungen

Im Folgenden soll es um (Systeme von) *linearen* Differentialgleichungen gehen.

Definition 15. LINEARE SYSTEME.

Es sei $n \in \mathbb{N}$ beliebig. Eine DGL heißt *lineares homogenes System (erster Ordnung)*, falls sie die Form

$$y'(t) = A(t)y(t) \quad (t \in I) \quad (53)$$

hat, wobei $A \in C^0(I; \mathbb{R}^{n \times n})$ eine stetig von der Zeit abhängige Matrix ist, d.h.

$$f(t, z) := A(t)z \quad (t \in I, z \in \mathbb{R}^n) \quad (54)$$

ist für jedes feste t linear in der z -Variablen.

Bemerkenswert an linearen homogenen Systemen ist, dass die allgemeine Lösung

$$\mathbb{L} := \{y \in C^1(I; \mathbb{R}^n) : y'(t) = A(t)y(t) \forall t \in I\} \quad (55)$$

einen Untervektorraum von $C^1(I; \mathbb{R}^n)$ bildet (HA). Das heißt insbesondere, dass Summen/Differenzen und skalare Vielfache von Lösungen wieder Lösungen sind. Diese Beobachtung wird auch oft *Superpositionsprinzip* genannt.

Beispiel 16. SPEZIALFALL $n = 1$.

Im Falle $n = 1$ können wir bereits eine Lösung bestimmen. In dem Fall ist nämlich

$A \in C^0(I, \mathbb{R}^{1 \times 1}) = C^0(I, \mathbb{R})$ eine skalarwertige Abbildung, d.h. die DGL liest sich

$$y'(t) = a(t)y(t) \quad (t \in I), \quad (56)$$

wobei $a \in C^0(I, \mathbb{R})$. Wir betrachten die DGL als DGL mit getrennten Veränderlichen mit $g(t) := a(t)$ und $h(z) := z$. Unter der Annahme $y(t) > 0$ für alle $t \in I$ rechnen wir wieder

$$\frac{d}{dt} \log(y(t)) = \frac{y'(t)}{y(t)} = a(t), \quad (57)$$

also gilt für eine fest gewählte Stammfunktion $\int^t a(s) \, ds$ und eine Konstante $C \in \mathbb{R}$

$$\log(y(t)) = \int^t a(s) \, ds + C. \quad (58)$$

Es folgt

$$y(t) = \exp\left(\int^t a(s) \, ds + C\right). \quad (59)$$

Definiert man $c := e^C$, so erhält man

$$y(t) = c \exp\left(\int^t a(s) \, ds\right). \quad (60)$$

Da wir die Annahme $y(t) > 0$ für alle $t \in I$ getroffen haben muss hier $c > 0$ sein. Wir wissen aber: Das die Menge aller Lösungen ein Vektorraum ist, muss auch $y(t) = -c \exp\left(\int^t a(s) \, ds\right)$ eine Lösung sein. Daher ergibt sich die folgende

BEHAUPTUNG: Die allgemeine Lösung von (56) ist gegeben durch

$$\mathbb{L} := \left\{ y \in C^1(I) \mid \exists c \in \mathbb{R} : y(t) = c \exp\left(\int^t a(s) \, ds\right) \forall t \in I \right\} \quad (61)$$

BEGRÜNDUNG:

Schritt 1. \mathbb{Z} Jedes Element aus \mathbb{L} ist eine Lösung. Dies rechnet man leicht durch Einsetzen in die Gleichung nach. (HA)

Schritt 2. \mathbb{Z} Jede Lösung ist ein Element aus \mathbb{L} . Wir gehen vor wie in Beispiel 3 (p. 3): Es sei $y_* \in C^1(I)$ eine Lösung. Definiere nun für $t \in I$ die Hilfsfunktion $g(t) := \exp\left(-\int^t a(s) \, ds\right) y_*(t)$. Man rechnet nach, dass für alle $t \in I$ gilt

$$g'(t) = \exp\left(-\int^t a(s) \, ds\right) y'_*(t) + \exp\left(-\int^t a(s) \, ds\right) (-a(t)y_*(t)) \stackrel{(56)}{=} 0. \quad (62)$$

Das heißt also, dass es $c_* \in \mathbb{R}$ gibt mit $g(t) = c_*$ für alle $t \in I$. Setzt man die Definition von g ein und stellt die Gleichung um, so erhält man $y_*(t) = c_* \exp\left(-\int^t a(s) \, ds\right)$.

Wir werden in diesem Kapitel verschiedene Modifikationen vornehmen und Lösungsformeln

für ähnliche DGLs kennenlernen. Wir werden die DGLs stets auf die Situation in (56) zurückführen. Daher halten wir nochmal fest

Wichtige Merkformel 17. LÖSUNGSFORMEL FÜR LINEARE DGL.

Es sei $a \in C^0(I; \mathbb{R})$. Die allgemeine Lösung der DGL

$$y'(t) = a(t)y(t) \quad (t \in I) \quad (63)$$

ist gegeben durch

$$\mathbb{L} := \left\{ y \in C^1(I) \mid \exists c \in \mathbb{R} : y(t) = c \exp \left(\int^t a(s) \, ds \right) \forall t \in I \right\}. \quad (64)$$

Als Anwendung können wir uns hier eine Population mit zeitabhängiger relativer Geburtenrate vorstellen. So zum Beispiel könnte es sein, dass in einer bestimmten Jahreszeit die Sterberate größer ist als in anderen Jahreszeiten, z.B. durch Hitzewellen. Dies würde die relative Geburtenrate im Sommer kleiner machen als im Winter. Auch die Saisonalität von Viruspopulationen könnte mit einem solchen Modell erfasst werden. Dazu später mehr.

1.2.1 Lineare inhomogene Differentialgleichungen

Im Folgenden studieren wir die Situation, dass die Differentialgleichung eine *Inhomogenität* hat, d.h. zu dem linearen Term wird ein zeitabhängiger Term addiert. Für eine Anwendung einer solchen Gleichung müssen wir uns etwas gedulden — aber nicht lang!

Definition 18. INHOMOGENE LINEARE DGL. Eine Differentialgleichung der Form

$$y'(t) = a(t)y(t) + b(t) \quad (t \in I) \quad (65)$$

heißt *inhomogene lineare DGL*. Hierbei sind $a, b \in C^0(I; \mathbb{R})$. In diesem Fall gilt also $y'(t) = f(t, y(t))$ für

$$f(t, z) := a(t)z + b(t). \quad (66)$$

Wiederum können wir eine explizite Lösungsformel angeben. Hierfür betrachten wir eine geschickte *Hilfsfunktion*, die von Merkformel 17 (p. 17) inspiriert ist.

Satz 19. LÖSUNGSFORMEL MIT INHOMOGENITÄT.

Es seien $a, b \in C^0(I)$. Gegeben sei die DGL

$$y'(t) = a(t)y(t) + b(t) \quad (t \in I). \quad (67)$$

Man definiere

$$e(t) := \exp \left(\int^t a(s) \, ds \right) \quad (t \in I), \quad (68)$$

wobei $\int^t a(s) \, ds$ eine beliebige feste Stammfunktion von a ist. Dann ist die allgemeine

Lösung von (67) gegeben durch

$$\mathbb{L} := \left\{ y \in C^1(I) : y(t) = e(t) \left(c + \int^t \frac{b(u)}{e(u)} du \right) \text{ für ein } c \in \mathbb{R} \right\}. \quad (69)$$

Eigentlich müssten wir bei der obigen Definition von \mathbb{L} noch dazu schreiben, dass die Formel aus (69) für alle $t \in I$ gelten soll. Im Folgenden werden wir das manchmal unterdrücken, um die Notation schlank zu halten.

Beweis.

SCHRITT 1. \mathbb{Z}_L : Jedes Element aus \mathbb{L} ist eine Lösung. Dies rechnen wir an dieser Stelle einfach nach. Sicherlich ist das etwas unbefriedigend, weil man ja irgendwie auf die Formel gekommen sein muss. Die weiteren Ausführungen im Schritt 2 werden die Formel aber *konstruktiv* herleiten. Es sei nun für ein $c \in \mathbb{R}$

$$y(t) := e(t) \left(c + \int^t \frac{b(u)}{e(u)} du \right). \quad (70)$$

Man rechnet leicht nach, dass $e'(t) = a(t)e(t)$ für alle $t \in I$ und damit gilt nach der Produktregel

$$y'(t) = e'(t) \left(c + \int^t \frac{b(u)}{e(u)} du \right) + e(t) \frac{b(t)}{e(t)} \quad (71)$$

$$= a(t)e(t) \left(c + \int^t \frac{b(u)}{e(u)} du \right) + b(t) = a(t)y(t) + b(t). \quad (72)$$

SCHRITT 2. \mathbb{Z}_L : Jede Lösung ist ein Element aus \mathbb{L} . Es sei $y \in C^1(I)$ eine Lösung. Definiere die *Hilfsfunktion*

$$h(t) := \exp \left(- \int^t a(s) ds \right) y(t). \quad (73)$$

Wir berechnen mit der Produktregel

$$h'(t) = \exp \left(- \int^t a(s) ds \right) y'(t) + \exp \left(- \int^t a(s) ds \right) (-a(t)y(t)) \quad (74)$$

$$= \exp \left(- \int^t a(s) ds \right) (y'(t) - a(t)y(t)) = \frac{y'(t) - a(t)y(t)}{\exp \left(\int^t a(s) ds \right)} \quad (75)$$

$$\stackrel{(68)}{=} \frac{y'(t) - a(t)y(t)}{e(t)} \stackrel{(67)}{=} \frac{b(t)}{e(t)}. \quad (76)$$

Nach dem Hauptsatz der Differential- und Integralrechnung gilt dann also für ein $c \in \mathbb{R}$

$$h(t) = c + \int^t \frac{b(u)}{e(u)} du. \quad (77)$$

Nutzen wir nun die Definition von $h(t)$ in (73) so erhalten wir

$$y(t) = \exp\left(\int^t a(s) \, ds\right) \left(c + \int^t \frac{b(u)}{e(u)} \, du\right) = e(t) \left(c + \int^t \frac{b(u)}{e(u)} \, du\right). \quad (78)$$

Das heißt y ist nach Gleichung (69) ein Element aus \mathbb{L} .

Am Ende dieses Beweises sollten wir uns fragen, wie man auf die Hilfsfunktion h kommt. Hierzu zunächst folgende Beobachtung: Aus den Ausführungen in Beispiel 16 (p. 15) würde folgen, dass im Falle $b \equiv 0$ die Hilfsfunktion h eine Konstante wäre. Diese Lösungstechnik heißt daher *Variation der Konstanten* und kann auch in Form eines *Lösungsansatzes* formuliert werden, wie wir im Folgenden sehen werden. Dieser Lösungsansatz geht auf den Mathematiker Joseph-Louis Lagrange (1736-1813) (LINK ZUR ORIGINALVERÖFFENTLICHUNG 1766) zurück.

Beweisreflexion 20. DIE VARIATION DER KONSTANTEN. Wir betrachten mal wieder die DGL

$$y'(t) = a(t)y(t) + b(t) \quad (t \in I). \quad (79)$$

Wir wissen aus Beispiel 16 (p. 15), dass jede Lösung der *homogenen DGL* $\bar{y}'(t) = a(t)\bar{y}(t)$, ($t \in I$) die Gestalt

$$\bar{y}(t) = ce(t), \quad \text{wobei } e(t) := \exp\left(\int^t a(s) \, ds\right) \quad (80)$$

hat. Hierbei ist $c \in \mathbb{R}$ eine Konstante. Aus Satz 19 (p. 17) lernen wir, die Lösung der inhomogene Gleichung die folgende Gestalt hat:

$$y(t) = c(t)e(t), \quad (81)$$

wobei $c(t)$ ein zeitabhängiger Ausdruck ist. Wir wissen aus Satz 19 (p. 17) auch schon, wie $c(t)$ konkret aussieht. Allerdings ist die Formel dafür schwer zu merken. Man kann aber in jeder konkreten Situation die DGL durch einen *Ansatz* lösen. Das heißt: Wir nehmen an, dass unsere Lösung die Gestalt $y(t) = c(t)e(t)$ hat und bestimmen dann eine Funktion $c(t)$, sodass die Gleichung erfüllt ist. Mit anderen Worten: Wir suchen eine Funktion $c \in C^1(I; \mathbb{R})$, sodass

$$y(t) := c(t)e(t) \quad (t \in I), \quad (82)$$

eine Lösung von (79) ist. Dieser Ansatz heißt *Variation der Konstanten*, weil die Konstante c in (80) durch eine Funktion ersetzt wird.

In der Tat liefert dieser Ansatz stets eine Lösung. Differenzieren wir (82) so erhalten wir (mit der Produktregel und $e'(t) = a(t)e(t)$)

$$y'(t) = c(t)e'(t) + c'(t)e(t) = c(t)a(t)e(t) + c'(t)e(t) \stackrel{(82)}{=} a(t)y(t) + c'(t)e(t). \quad (83)$$

Andererseits soll aber nach (79) auch gelten

$$y'(t) = a(t)y(t) + b(t). \quad (84)$$

Setzen wir das mit (83) gleich so erhalten wir

$$a(t)y(t) + c'(t)e(t) = a(t)y(t) + b(t). \quad (85)$$

Subtrahieren wir nun $a(t)y(t)$ auf beiden Seiten, so bekommen wir

$$c'(t)e(t) = b(t) \quad (86)$$

und als Konsequenz

$$c'(t) = \frac{b(t)}{e(t)}. \quad (87)$$

Durch Integration erhalten wir

$$c(t) = c + \int^t \frac{b(u)}{e(u)} du. \quad (88)$$

Setzen wir dies in (82) ein, so erhalten wir exakt die Lösung aus Satz 19 (p. 17). Ein solcher Lösungsansatz kann auch ins Leere führen, z.B. wenn es keine Lösung gäbe, die die Gestalt in (82) hat. Wie Satz 19 (p. 17) und die obigen Ausführungen zeigen, geht der Ansatz in unserem Fall aber immer auf. Mehr noch: Mit einem solchen Lösungsansatz kann man auch im allgemeineren Kontext stets lineare DGLs mit Inhomogenitäten lösen — dasselbe funktioniert auch für Systeme! Voraussetzung ist aber, dass man die homogene Gleichung bereits gelöst hat, da man stets die Konstante aus der homogenen Lösung (80) variieren muss.

Beispiel 21. Wir bestimmen für $\alpha, \beta, \gamma \in \mathbb{R}$ die allgemeine Lösung von

$$y'(t) = \alpha y(t) + \gamma e^{\beta t} \quad (t \in \mathbb{R}) \quad (89)$$

Um eine Lösung zu bestimmen, können wir nun entweder die Lösungsformel verwenden oder die Variation der Konstanten. Wir zeigen hier mal letztere Methode. Die allgemeine Lösung der homogenen Gleichung $\bar{y}'(t) = \alpha \bar{y}(t)$ ist (siehe Merkformel 17, p. 17) gegeben durch

$$\bar{y}(t) = ce^{\alpha t} \quad (t \in \mathbb{R}). \quad (90)$$

Um nun eine Lösung der inhomogenen Gleichung zu finden machen wir folgenden Ansatz

$$y(t) \stackrel{!}{=} c(t)e^{\alpha t} \quad (91)$$

und berechnen

$$y'(t) = c'(t)e^{\alpha t} + \alpha c(t)e^{\alpha t} = c'(t)e^{\alpha t} + \alpha y(t). \quad (92)$$

Andererseits gilt laut (89) $y'(t) = \alpha y(t) + \gamma e^{\beta t}$. Wir setzen die mit dem obigen Ausdruck gleich und erhalten

$$c'(t)e^{\alpha t} + \alpha y(t) = \alpha y(t) + \gamma e^{\beta t}, \quad (93)$$

also

$$c'(t)e^{\alpha t} = \gamma e^{\beta t}. \quad (94)$$

Umgeformt heißt das

$$c'(t) = \gamma e^{(\beta-\alpha)t}. \quad (95)$$

Integrieren wir so erhalten wir

$$c(t) = \begin{cases} c + \frac{1}{\beta-\alpha} \gamma e^{(\beta-\alpha)t} & \alpha \neq \beta, \\ c + \gamma t & \alpha = \beta. \end{cases} \quad (96)$$

Damit gilt also für die allgemeine Lösung

$$y(t) = \begin{cases} e^{\alpha t} \left(c + \frac{\gamma}{\beta-\alpha} e^{(\beta-\alpha)t} \right) & \alpha \neq \beta \\ e^{\alpha t} (c + \gamma t) & \alpha = \beta. \end{cases} \quad (97)$$

Es ist bemerkenswert, dass die Funktion im Falle $\alpha = \beta$ ein anderes Verhalten zeigt, als im Fall $\alpha \neq \beta$. Wir werden dieses Phänomen später noch vertiefen.

Übung 22. BERNoulli'sCHE DIFFERENTIALGLEICHUNG. Es seien $a, b \in C^0(\mathbb{R}; \mathbb{R})$ und $\gamma > 0, \gamma \notin \mathbb{N}$. Betrachten Sie das AWP

$$\begin{cases} y'(t) = a(t)y(t) + b(t)y(t)^\gamma & (t \in \mathbb{R}) \\ y(t_0) = y_0 \end{cases} \quad (98)$$

wobei $t_0 \in \mathbb{R}$ und $y_0 \geq 0$. Beachten Sie, dass $y(t)^\gamma$ nur definiert ist, falls $y(t) \geq 0$ für alle $t \in \mathbb{R}$.

(a) Zeigen Sie: Ist $y_0 > 0$ und

$$|1 - \gamma| \int_{\mathbb{R}} |b(t)| \exp \left(- \int_{t_0}^t (1 - \gamma)a(s) \, ds \right) \, dt < y_0^{1-\gamma} \quad (99)$$

so gibt es eine eindeutige globale Lösung $y \in C^1(\mathbb{R}; \mathbb{R})$.

(b) Kann auf die Voraussetzungen ' $y_0 > 0$ ' und (99) verzichtet werden? Falls nein: Geben Sie Gegenbeispiele sowohl für die Eindeutigkeit als auch für die Globalität an.

HINWEIS: Betrachten Sie $z(t) := y(t)^{1-\gamma}$ und leiten Sie eine Differentialgleichung für z her.

SINN DER ÜBUNG: Man sieht, dass man nicht nur die standardmäßigen linearen DGLs lösen kann. Mit geschickten *Substitutionen* löst man auch kompliziertere DGLs, die vielfach in der Anwendung auftauchen. Auch die Bernoulli'sche Differentialgleichung wird

in der Populationsdynamik angewandt: vgl. [GIPLIN, AYALA (1973)]. Im Falle $a, b \equiv const.$, $a > 0, b < 0$ und “ $\gamma \approx 2$ ” erhält man schließlich das logistische Wachstum aus Übung 12 (p. 13).

1.2.2 Lineare Systeme mit konstanten Koeffizienten

Im Folgenden beschäftigen wir uns mit Differentialgleichungen der Form

$$y'(t) = Ay(t) \quad (t \in \mathbb{R}), \quad (100)$$

wobei hier $A \in \mathbb{R}^{n \times n}$ gegeben und $y \in C^1(\mathbb{R}; \mathbb{R}^n)$ gesucht ist. Dies ist ein Spezialfall von $y'(t) = A(t)y(t)$ in dem $A(t) \equiv A$ eine konstante Matrix ist. An dieser Stelle sei noch einmal gesagt, dass wir

$$y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} \quad (101)$$

stets als Spaltenvektor verstehen wollen, und demnach gilt — wie aus den Analysis-Grundlagenvorlesungen bekannt — für $y \in C^1(\mathbb{R}, \mathbb{R}^n)$

$$y'(t) = \begin{pmatrix} y'_1(t) \\ \vdots \\ y'_n(t) \end{pmatrix} \quad (t \in \mathbb{R}). \quad (102)$$

Die Einträge der Matrix A bezeichnen wir mit

$$A = (a_{ij})_{i,j=1,\dots,n} \quad (103)$$

Unsere Differentialgleichung hat daher — lang ausgeschrieben — die Form

$$\begin{pmatrix} y'_1(t) \\ \vdots \\ y'_n(t) \end{pmatrix} = A \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}y_j(t) \\ \vdots \\ \sum_{j=1}^n a_{nj}y_j(t) \end{pmatrix}. \quad (104)$$

Wir werden gleich sehen, dass die Globalität von Lösungen stets gewährleistet ist. Daher ist ‘ $t \in \mathbb{R}$ ’ an dieser Stelle keine Einschränkung.

In dieser Sektion ist es sinnvoll, die Situation zu *komplexifizieren*, d.h. zuzulassen, dass die betrachteten Größen \mathbb{C} -wertig sind. Der Grund ist, dass bei komplexwertigen Matrizen das *Diagonalisieren* leichter ist — und die Methode des Diagonalisierens werden wir uns stark zunutze machen!

Bevor wir aber damit loslegen, müssen wir zunächst Konventionen für das Differenzieren komplexerer Abbildungen treffen:

Definition 23. DIFFERENTIATION KOMPLEXWERTIGER ABBILDUNGEN.

Es sei $I \subset \mathbb{R}$ ein offenes Intervall in \mathbb{R} . Eine Abbildung $g : I \rightarrow \mathbb{C}$ heißt (*stetig*) *differenzierbar* auf I , falls $\operatorname{Re}(g)$ und $\operatorname{Im}(g)$ als Abbildungen von I nach \mathbb{R} (stetig) differenzierbar sind. Die *Ableitung* von g notieren wir mit

$$g'(t) := \operatorname{Re}(g)'(t) + i \operatorname{Im}(g)'(t). \quad (105)$$

Eine Abbildung $y : I \rightarrow \mathbb{C}^n$ heißt (stetig) differenzierbar, falls jede ihrer Komponenten (stetig) differenzierbar ist. Die Ableitung definieren wir dann mit der obigen Konvention exakt wie in Gleichung (102). Die Menge aller stetig differenzierbaren Abbildungen von I nach \mathbb{C}^n bezeichnen wir mit $C^1(I; \mathbb{C}^n)$.

Alternativ hätte man auch die äquivalente klassische Definition wählen können: Man nennt eine Funktion $g : I \rightarrow \mathbb{C}$ differenzierbar, falls für alle $t_0 \in I$

$$\lim_{t \rightarrow t_0} \frac{g(t) - g(t_0)}{t - t_0} =: g'(t_0) \quad (106)$$

in \mathbb{C} existiert. Dass die beiden Aussagen äquivalent sind folgt sofort aus der Beobachtung, dass für alle $t, t_0 \in I (\subset \mathbb{R})$

$$\frac{g(t) - g(t_0)}{t - t_0} = \frac{\operatorname{Re}(g)(t) - \operatorname{Re}(g)(t_0)}{t - t_0} + i \frac{\operatorname{Im}(g)(t) - \operatorname{Im}(g)(t_0)}{t - t_0}. \quad (107)$$

Die Details überlassen wir den Lesenden als Übungsaufgabe.

Die meisten Rechenregeln für das Differenzieren übertragen sich wie gewohnt ins Komplexe, sie bedürfen aber manchmal durchaus einer Herleitung! Um unsere Theorie zu retten werden wir die Kettenregel für die komplexe Exponentialfunktion benötigen, die wir als Übungsaufgabe nochmal klarstellen wollen:

Übung 24. RECHENREGELN UND DGLs IN \mathbb{C} .

- (a) Es sei $g : I \rightarrow \mathbb{C}$ stetig differenzierbar. Zeigen Sie: Dann definiert auch $e^g : t \mapsto e^{g(t)}, (t \in I)$ eine stetig differenzierbare Funktion und es gilt

$$\frac{d}{dt} e^{g(t)} = g'(t) e^{g(t)}. \quad (108)$$

- (b) Nun sei $a \in C^0(I; \mathbb{C})$. Bestimmen Sie die allgemeine Lösung (als Teilmenge von $C^1(I; \mathbb{C})$) der DGL

$$y'(t) = a(t)y(t) \quad (t \in I). \quad (109)$$

TIPP: Es gibt bei Übung 23 (a) mehrere Möglichkeiten: Entweder die Verwendung der Euler'schen Formel oder die Ausnutzung von Rechenregeln für Potenzreihen.

SINN DER ÜBUNG: Man sieht, dass die Theorie, die wir bisher hergeleitet haben auch im Komplexen ihre Gültigkeit hat.

Man überlegt sich leicht, dass der Hauptsatz der Differential- und Integralrechnung auch in \mathbb{C} gilt (HA), mit der folgenden Konvention: Für $g : I \rightarrow \mathbb{C}$ stetig und beschränkt gilt

$$\int_I g(s) \, ds := \int_I \operatorname{Re}(g(s)) \, ds + i \int_I \operatorname{Im}(g(s)) \, ds. \quad (110)$$

Da alles, was wir in Kapitel 1.2 gemacht haben auf dem Hauptsatz der Differential- und Integralrechnung und auf Gleichung (108) beruht hat, verallgemeinern sich Merkformel 17

(p. 17) und Satz 19 (p. 17) leicht auf komplexwertige Situationen. Insbesondere gilt Beispiel 21 (p. 20) auch für $\alpha, \beta \in \mathbb{C}$. Der Grundstein für die Betrachtung komplexwertiger DGLs ist somit gelegt.

Komplexifizierung kann sich übrigens auch in anderen Kontexten für nicht-lineare DGLs lohnen, so auch zum Beispiel hier:

***-Übung 25. KOMPLEXIFIZIERUNG VON DGLS.** Betrachten Sie das System

$$\begin{pmatrix} y'_1(t) \\ y'_2(t) \end{pmatrix} = \begin{pmatrix} y_1(t)^2 - y_2(t)^2 \\ 2y_1(t)y_2(t) \end{pmatrix} \quad (t \in \mathbb{R}) \quad (111)$$

Bestimmen Sie die allgemeine Lösung.

SINN DER ÜBUNG: Man sieht, dass sich kompliziert aussehende Differentialgleichungen durch Komplexifizierung substantiell vereinfachen können. Diesen konkreten Fall habe ich auch in meiner eigenen Forschung verwendet, siehe FORSCHUNGSARTIKEL, GLEICHUNG (2.6) UND (2.20) – ACHTUNG: LÖSUNGSSPOILER.

Sind die Differentialgleichungen, die wir betrachten, komplexwertig, so suchen wir auch komplexwertige Lösungen, d.h. für diesen Abschnitt sind Lösungen stets Elemente aus $C^1(I; \mathbb{C}^n)$.

Definition 26. DGL-SYSTEM MIT KONSTANTEN KOEFFIZIENTEN.

Es sei $A \in \mathbb{C}^{n \times n}$. Eine DGL der Form

$$y'(t) = Ay(t) \quad (t \in \mathbb{R}) \quad (112)$$

heißt *lineare homogene DGL (erster Ordnung) mit konstanten Koeffizienten*.

Im Falle $n = 1$ haben wir (112) bereits gelöst. Wir bewegen uns nun schrittweise auf eine Lösung für $n > 1$ zu.

Gedankenexperiment 27. DIAGONALMATRIZEN. Hier wollen wir einen Spezialfall studieren: Was passiert, wenn A eine *Diagonalmatrix* ist, d.h.

$$A = \text{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \quad \text{für } \lambda_1, \dots, \lambda_n \in \mathbb{C} ? \quad (113)$$

In diesem Fall sehen wir mit Gleichung (108), dass sich $y'(t) = Ay(t)$ vereinfacht zu

$$\begin{pmatrix} y'_1(t) \\ \vdots \\ y'_n(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 y_1(t) \\ \vdots \\ \lambda_n y_n(t) \end{pmatrix}. \quad (114)$$

Man beachte: zwei Vektoren sind genau dann gleich wenn alle ihre Komponenten gleich

sind. Wir schließen

$$y'_1(t) = \lambda_1 y_1(t), \quad y'_2(t) = \lambda_2 y_2(t), \quad \dots, \quad y'_n(t) = \lambda_n y_n(t). \quad (115)$$

Insbesondere können wir die Differentialgleichungen voneinander *entkoppeln*. Das heißt: wir können jede dieser einzelnen Komponentengleichungen separat lösen. Dies ist ein großer Vorteil gegenüber der allgemeinen Gestalt in (108). Wir lösen also nun für jede Komponente einzeln. Mit Merkformel 17 (p. 17), angewendet auf jede Komponente, sieht man: Es gibt $c_1, \dots, c_n \in \mathbb{C}$, so, dass

$$y_1(t) = c_1 e^{\lambda_1 t}, \quad y_2(t) = c_2 e^{\lambda_2 t}, \quad \dots, \quad y_n(t) = c_n e^{\lambda_n t}, \quad (116)$$

das heißt

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix} = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} = \begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}. \quad (117)$$

Die Umformung auf der rechten Seite (die man beim ersten Lesen ‘rückwärts’ lesen sollte) haben wir gemacht, um die allgemeine Lösung kompakt darstellen zu können. Es gilt nämlich

$$\mathbb{L} := \{y \in C^1(\mathbb{R}; \mathbb{C}^n) : y(t) = E(t)c \text{ für ein } c \in \mathbb{C}^n\}, \quad (118)$$

wobei

$$E(t) = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \quad (t \in \mathbb{R}). \quad (119)$$

Für Diagonalmatrizen ist das System also *entkoppelt*. Wir tasten uns nun weiter vor und sprechen über *diagonalisierbare Matrizen*. Eine Matrix A heißt *diagonalisierbar*, falls es eine invertierbare Matrix $S \in \mathbb{R}^{n \times n}$ gibt, sodass

$$S^{-1} A S = \text{diag}(\lambda_1, \dots, \lambda_n) \quad (120)$$

für gewisse $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. In der linearen Algebra haben Sie besprochen, wie bei konkret gegebener Matrix A die Matrix S und die *Eigenwerte* $\lambda_1, \dots, \lambda_n$ gefunden werden können. Für den weiteren Verlauf dieses Kapitels wäre es hilfreich, wenn Sie sich diese Konzepte noch einmal vergegenwärtigen (HA). Auch die *Jordan-Normalform* wird später noch benötigt.

Gedankenexperiment 28. DIAGONALISIERBARE MATRIZEN. Wieder wollen wir die allgemeine Lösung der Differentialgleichung

$$y'(t) = A y(t) \quad (t \in \mathbb{R}) \quad (121)$$

finden — diesmal unter der Annahme, dass A *diagonalisierbar* ist. Seien wiederum $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A . Wir definieren $D := \text{diag}(\lambda_1, \dots, \lambda_n)$ und schreiben

$$A = SDS^{-1}, \quad (122)$$

wobei S wie in (121) gewählt ist. Die Differentialgleichung schreibt sich dann

$$y'(t) = SDS^{-1}y(t) \quad (t \in \mathbb{R}) \quad (123)$$

Wir multiplizieren nun auf beiden Seiten mit S^{-1} und erhalten

$$S^{-1}y'(t) = DS^{-1}y(t) \quad (t \in \mathbb{R}) \quad (124)$$

Um die folgende Lösungsidee zu motivieren, wollen wir uns den Ausdruck mal anders geklammert anschauen. Es gilt

$$\frac{d}{dt}[S^{-1}y(t)] = D[S^{-1}y(t)] \quad (t \in \mathbb{R}) \quad (125)$$

Hierbei haben wir benutzt, dass wegen der Linearität der Ableitung $\frac{d}{dt}[S^{-1}y(t)] = S^{-1}y'(t)$ gilt (HA). Für eine Lösung $y \in C^1(\mathbb{R}; \mathbb{R}^n)$ können wir nun $z(t) := S^{-1}y(t)$, ($t \in \mathbb{R}$) definieren. Gleichung (125) liest sich dann wie folgt:

$$z'(t) = Dz(t) \quad (t \in \mathbb{R}). \quad (126)$$

Hiermit sind wir in derselben Situation wie in Gleichung (113): $z \in C^1(\mathbb{R}; \mathbb{R}^n)$ löst $z'(t) = Dz(t)$, wobei $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ eine Diagonalmatrix ist. Wir schließen aus dem vorangegangenen Gedankenexperiment, dass es ein $c \in \mathbb{R}^n$ gibt mit

$$z(t) = E(t)c, \quad \text{wobei } E(t) := \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \quad \forall t \in \mathbb{R}. \quad (127)$$

Wir erinnern uns nun, dass $z(t) = S^{-1}y(t)$ und folgern

$$y(t) = Sz(t) = SE(t)c \quad (t \in \mathbb{R}). \quad (128)$$

Damit erhalten wir die folgende

BEHAUPTUNG: Ist A diagonalisierbar mit $S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$, so ist die allgemeine Lösung von (121) gegeben durch

$$\mathbb{L} := \{y \in C^1(\mathbb{R}; \mathbb{C}^n) : y(t) = SE(t)c \text{ für ein } c \in \mathbb{C}^n\}, \quad (129)$$

wobei

$$E(t) = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}) \quad (t \in \mathbb{R}). \quad (130)$$

Übung 29. Beweisen Sie die obige Behauptung formal.

Wir haben im obigen Gedankenexperiment von der *Substitution* $z(t) := S^{-1}y(t)$ profitiert. Im Folgenden werden wir immer wieder Situationen kennenlernen, bei denen eine solche Substitution hilfreich ist. Sie hilft, Differentialgleichungen durch *Basiswechsel* auf eine einfachere Form zu bringen.

Aus der linearen Algebra kennen wir auch Kriterien für die Diagonalisierbarkeit von

Matrizen. Moralisch sind sehr viele Matrizen diagonalisierbar: Man kann sogar zeigen, dass die diagonalisierbaren Matrizen in $\mathbb{C}^{n \times n}$ dicht liegen (HA, schwer aber machbar).

Unser DGL-System lässt sich aber auch für nicht-diagonalisierbare Matrizen lösen — mithilfe der *Jordan-Normalform*! Diese existiert dann schlussendlich für jede Matrix $A \in \mathbb{C}^{n \times n}$, und somit ist man befähigt beliebige lineare Systeme mit konstanten Koeffizienten explizit zu lösen.

Betrachten wir also nun die DGL

$$y'(t) = Ay(t) \quad (t \in \mathbb{R}) \quad (131)$$

wobei A beliebig ist. Allgemein hat A eine Jordan-Normalform: Es gibt eine invertierbare Matrix $S \in \mathbb{C}^{n \times n}$ mit

$$S^{-1}AS = \begin{pmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_r \end{pmatrix}, \quad (132)$$

wobei $J_m \in \mathbb{C}^{n_m \times n_m}$, $n_m \geq 1$, $n_1 + \dots + n_r = n$ die einzelnen Jordanblöcke sind, d.h.

$$J_m = \begin{pmatrix} \lambda_m & 1 & \dots & 0 \\ \vdots & \ddots & 1 & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & \lambda_m \end{pmatrix} = \lambda_m I + N, \quad (133)$$

wobei $N = (n_{ij})_{i,j=1,\dots,n_m}$ die Matrix ist, die auf der ersten Nebendiagonalen nur Einser und sonst Nullen hat, d.h.

$$n_{ij} = \begin{cases} 1 & j = i + 1 \\ 0 & \text{sonst.} \end{cases} \quad (134)$$

Wir werden wieder in der Lage sein (131) durch Substitutionen so zu vereinfachen, dass wir nur die Gleichung für einen Jordanblock lösen müssen. In der Tat: Definiert man $\tilde{y}(t) := S^{-1}y(t)$, $(t \in \mathbb{R})$ so gilt für alle $t \in \mathbb{R}$

$$\tilde{y}'(t) = S^{-1}y'(t) = S^{-1}Ay(t) = S^{-1}AS(S^{-1}y(t)) = S^{-1}AS\tilde{y}(t). \quad (135)$$

Das bedeutet

$$\tilde{y}'(t) = \begin{pmatrix} J_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_r \end{pmatrix} \tilde{y}(t). \quad (136)$$

Dieses System lässt sich entkoppeln: \tilde{y} hat die Gestalt $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_r)^T$ wobei $\tilde{y}_m \in C^1(\mathbb{R}; \mathbb{C}^{n_m})$ Lösung von

$$\tilde{y}'_m(t) = J_m \tilde{y}_m(t) = (\lambda_m I + N) \tilde{y}_m(t). \quad m = 1, \dots, r. \quad (137)$$

Wir müssen für eine Lösung also nur die Funktionen \tilde{y}_m finden. Das heißt: Wir haben das Problem reduziert auf das Folgende:

$$\text{Finde } z \in C^1(\mathbb{R}, \mathbb{C}^d) \text{ so, dass } z'(t) = (\lambda I + N)z(t) \quad (t \in \mathbb{R}). \quad (138)$$

Hierbei ist $\lambda \in \mathbb{C}$ beliebig und $I, N \in \mathbb{C}^{d \times d}$ definiert wie in (134). Man beachte, dass wir hier in diesem Kontext unsere Dimension mit d notieren und nicht wie üblich mit n , weil wir ja bei dem obigen Entkopplungsschritt ein n -dimensionales System in viele Teilsysteme zerteilen. Beispielhaft wollen wir nun mal (138) für $d = 2$ lösen.

Gedankenexperiment 30. 2×2 -JORDANBLÖCKE. Es sei $d = 2$ und $\lambda \in \mathbb{C}$. Wir bestimmen die allgemeine Lösung für

$$z'(t) = (\lambda I + N)z(t) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} z(t) \quad (t \in \mathbb{R}). \quad (139)$$

Wir schreiben dazu die Gleichung in einzelnen Komponenten aus. Es gilt

$$\begin{pmatrix} z'_1(t) \\ z'_2(t) \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \begin{pmatrix} \lambda z_1(t) + z_2(t) \\ \lambda z_2(t) \end{pmatrix}. \quad (140)$$

Schauen wir nun auf die zweite Komponente, so erhalten wir die Gleichung $z'_2(t) = \lambda z_2(t)$. Aus dieser Gleichung schließen wir (siehe Formel 17 (p. 17)), dass es eine Konstante $c_2 \in \mathbb{C}$ geben muss mit der Eigenschaft, dass

$$z_2(t) = c_2 e^{\lambda t} \quad (t \in \mathbb{R}). \quad (141)$$

Damit ist z_2 schon (bis auf einen freien Parameter $c_2 \in \mathbb{C}$) komplett beschrieben. Setzen wir diese neue Erkenntnis in die erste Komponente von (140) ein, so erhalten wir

$$z'_1(t) = \lambda z_1(t) + z_2(t) = \lambda z_1(t) + c_2 e^{\lambda t}. \quad (142)$$

Diese Gleichung entspricht der Situation aus Beispiel 21 (p. 20) mit $\alpha = \beta = \lambda$ und $\gamma = c_2$. Mit (96) erhalten wir, dass es ein $c_1 \in \mathbb{C}$ gibt mit

$$z_1(t) = e^{\lambda t}(c_1 + c_2 t) = e^{\lambda t}c_1 + t e^{\lambda t}c_2 \quad (t \in \mathbb{R}). \quad (143)$$

Setzen wir dies und (141) zusammen, so erhalten wir für gewisse $c_1, c_2 \in \mathbb{C}$

$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \begin{pmatrix} e^{\lambda t}c_1 + t e^{\lambda t}c_2 \\ e^{\lambda t}c_2 \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & t e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \quad (144)$$

wobei wir im letzten Schritt eine Matrixmultiplikation ‘rückwärts’ durchgeführt haben, um den Ausdruck in eine angenehme Form zu bringen. Wir schließen — wie immer — dass die allgemeine Lösung gegeben ist durch

$$\mathbb{L} := \{z \in C^1(\mathbb{R}, \mathbb{C}^2) : z(t) = E(t)c, \text{ für ein } c \in \mathbb{C}^2\}, \quad (145)$$

wobei

$$E(t) = \begin{pmatrix} e^{\lambda t} & t e^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}. \quad (146)$$

Wir könnten an dieser Stelle noch in einem Satz formulieren, wie das Ganze in beliebigen

Dimensionen aussieht. Es ist aber vielleicht auch ohne einen solchen Satz einsichtig, dass man das obige Verfahren auch auf höherdimensionale Situationen übertragen kann. Außerdem wird man sich den Satz aufgrund der vielen Indizes kaum merken können. Wir sparen uns an dieser Stelle die Arbeit und formulieren einfach nochmal die Quintessenz:

1. Durch eine Substitution ist es stets möglich zu erreichen, dass die Matrix A in Jordan-Normalform angegeben ist. (Für Details siehe nochmal (135)).
2. Für Jordan-Blöcke lässt sich das System iterativ lösen — Die letzte Gleichung lässt sich stets entkoppeln und separat lösen, woraufhin die Lösung dann in die obigen Gleichungen eingesetzt werden kann.

Zusammenfassung 31. ZWEIDIMENSIONALE SYSTEME.

Zweidimensionale Systeme haben wir damit komplett “abgegrast”. Falls die DGL

$$y'(t) = Ay(t) \quad (t \in \mathbb{R}) \quad (147)$$

für eine Matrix $A \in \mathbb{R}^{2 \times 2}$ gegeben ist, so gibt es zwei Möglichkeiten.

Fall 1. A diagonalisierbar. Ist A diagonalisierbar, so gibt es $S \in \mathbb{C}^{2 \times 2}$ invertierbar und $\lambda_1, \lambda_2 \in \mathbb{C}$ mit

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (148)$$

In dem Fall ist, wie wir gesehen haben

$$\mathbb{L} = \{y \in C^1(\mathbb{R}; \mathbb{C}^2) : y(t) = SE(t)c \text{ für ein } c \in \mathbb{C}^2\}, \quad E(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}.$$

Fall 2. A nicht diagonalisierbar. Ist A hingegen nicht diagonalisierbar, so gibt es $S \in \mathbb{C}^{2 \times 2}$ invertierbar und $\lambda \in \mathbb{C}$ mit der Eigenschaft

$$S^{-1}AS = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}. \quad (149)$$

Substitutieren wir $z(t) = S^{-1}y(t)$, so erhalten wir

$$z'(t) = S^{-1}y'(t) \stackrel{(147)}{=} S^{-1}Ay(t) = S^{-1}ASz(t) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} z(t). \quad (150)$$

Nach dem vorigen Beispiel gilt dann

$$z(t) = E(t)c \text{ für ein } c \in \mathbb{C}^2, \quad \text{wobei } E(t) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}. \quad (151)$$

Da $y(t) = Sz(t)$ (siehe vorige Substitution) gilt $y(t) = SE(t)c$ und für ein $c \in \mathbb{C}^2$ und damit damit

$$\mathbb{L} = \{y \in C^1(\mathbb{R}; \mathbb{C}^2) : y(t) = SE(t)c \text{ für ein } c \in \mathbb{C}^2\}, \quad E(t) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Übung 32. LINEARE DGLs UND DIE JORDAN-NORMALFORM

(a) Finden Sie die allgemeine Lösung zu

$$z'(t) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} z(t). \quad (152)$$

(b) Finden Sie die allgemeine Lösung zu

$$y'(t) = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} y(t). \quad (153)$$

HINWEIS. Die Jordan-Normalform der unteren Matrix aus Teilaufgabe (b) ist exakt die Matrix aus Teilaufgabe (a).

SINN DER ÜBUNG. Diese Übung soll aufzeigen, dass man mit den entwickelten Methoden auch für höherdimensionale lineare DGLs mit konstanten Koeffizienten die Lösung explizit hinschreiben kann. Wir trainieren hierbei die in den vorherigen Gedankenexperimenten erlernten Lösungstechniken. Auch wiederholen wir in dieser Übung, wie man die Jordan-Normalform einer gegebenen Matrix findet.

***Übung 33. SYSTEME IN JORDAN-NORMALFORM** Finden und beweisen Sie für allgemeine $d \in \mathbb{N}$ eine Formel für die allgemeine Lösung von

$$z'(t) = (\lambda I + N)z(t) \quad (t \in \mathbb{R}). \quad (154)$$

wobei I die $d \times d$ -Einheitsmatrix ist und $N \in \mathbb{C}^{d \times d}$ wie in (134) ist.

SINN DER ÜBUNG: Dies ist der letzte Schritt, der uns zu einer allgemeinen Lösungsformel für allgemeine lineare Systeme mit konstanten Koeffizienten fehlen würde.

1.2.3 Die Fundamentalmatrix

Die Beobachtungen aus dem letzten Abschnitt legen nahe, dass die allgemeine Lösung von $y'(t) = Ay(t)$ ($A \in \mathbb{C}^{n \times n}$) stets von folgender Form ist:

$$\mathbb{L} := \{y \in C^1(\mathbb{R}, \mathbb{C}^n) : y(t) = M(t)c \text{ für ein } c \in \mathbb{C}^n\}, \quad (155)$$

wobei $M \in C^1(\mathbb{R}; \mathbb{C}^{n \times n})$ eine zeitabhängige Matrix ist. Eine weitere Beobachtung ist, dass $M(t)$ stets für alle $t \in \mathbb{R}$ invertierbar ist. Auch dies ist kein Zufall und gilt — wie wir später sehen werden — sogar allgemein für Gleichungen der Form $y'(t) = A(t)y(t)$, ($t \in I$)!

Die Formel in (155) zeigt nochmal explizit, dass die Menge aller Lösungen einen Untervektorraum von $C^1(\mathbb{R}; \mathbb{C}^n)$ bildet (HA). Wir lernen hier sogar mehr: Der Vektorraum in (155) ist (wegen der Invertierbarkeit von $M(t)$) exakt n -dimensional (HA)!

Aus (155) lernen wir auch, dass die allgemeine Lösung komplett durch die Abbildung $t \mapsto M(t)$ beschrieben werden kann. Weil diese Abbildung so wichtig ist, geben wir dieser einen speziellen Namen.

Definition 34. FUNDAMENTALMATRIX Es sei eine lineare DGL

$$y'(t) = A(t)y(t) \quad (t \in I) \quad (156)$$

für $A \in C^0(I; \mathbb{C}^{n \times n})$ gegeben.

Eine Abbildung $t \mapsto M(t)$, ($t \in I$) in $C^1(I; \mathbb{C}^{n \times n})$ heißt *Fundamentalmatrix* für die DGL falls

1. $M(t)$ invertierbar für alle $t \in I$.
2. Die allgemeine Lösung ist gegeben durch

$$\mathbb{L} := \{y \in C^1(I; \mathbb{C}^n) : y(t) = M(t)c \text{ für ein } c \in \mathbb{C}^n\}. \quad (157)$$

Für Systeme mit konstanten Koeffizienten können wir $t \mapsto M(t)$ nun bereits explizit angeben. Für Systeme mit nicht-konstanten Koeffizienten müssen wir uns noch ein wenig gedulden...

Haben wir eine Fundamentalmatrix gefunden, so lassen sich alle Anfangswertprobleme eindeutig lösen. In der Tat: Sei für ein $t_0 \in I$ und ein $y_0 \in \mathbb{C}^n$ das Anfangswertproblem

$$\begin{cases} y'(t) = A(t)y(t) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (158)$$

gegeben und sei $t \mapsto M(t)$ eine Fundamentalmatrix. Wir wissen, dass die Lösung die Gestalt $y(t) = M(t)c$, ($t \in I$) für ein $c \in \mathbb{C}^n$ hat. Um alle möglichen Werte für c zu bestimmen betrachten wir

$$y_0 \stackrel{!}{=} y(t_0) = M(t_0)c. \quad (159)$$

Wegen der Invertierbarkeit von $M(t_0)$ gibt es hierzu genau eine Lösung: $c = M(t_0)^{-1}y_0$. Daher ist

$$y(t) = M(t)M(t_0)^{-1}y_0 \quad (t \in I) \quad (160)$$

die einzige mögliche Lösung.

Man kann aus der Fundamentalmatrix noch weitere nützliche Dinge herauslesen.

Definition 35. FUNDAMENTALSYSTEM.

Es sei $t \mapsto M(t)$, ($t \in I$) eine Fundamentalmatrix für eine lineare homogene DGL $y'(t) =$

$A(t)y(t)$, $(t \in I)$. Die Menge von Abbildungen $\{t \mapsto M(t)e_1, \dots, t \mapsto M(t)e_n\} \subset C^1(I; \mathbb{C}^n)$, die jeweils t auf die Spaltenvektoren von $M(t)$ schickt, wird *Fundamentalsystem* genannt.

Man sieht: Jedes Element des Fundamentalsystems ist eine Lösung (HA). Außerdem lässt sich jede Lösung auf eindeutige Weise als Linearkombination von Elementen eines gegebenen Fundamentalsystems schreiben. In der Tat: Ist

$$c = c_1 e_1 + \dots + c_n e_n \in \mathbb{C}^n \quad (161)$$

beliebig vorgegeben, so ist die Lösung $y(t) = M(t)c = M(t)(c_1 e_1 + \dots + c_n e_n)$ gegeben durch

$$y(t) = c_1(M(t)e_1) + \dots + c_n(M(t)e_n). \quad (162)$$

Da jede Lösung $y \in \mathbb{L}$ die Gestalt $y(t) = M(t)c$ für ein eindeutiges $c \in \mathbb{C}^n$ hat, können wir y als eindeutige Linearkombination von Elementen des Fundamentalsystems schreiben.

Man beachte, dass wir für die Eindeutigkeit von c wiederum die Invertierbarkeit von $M(t)$ benötigen.

Zusammengefasst:

Merksatz 36.

Jedes Fundamentalsystem bildet eine *Basis* des Lösungsraumes \mathbb{L} . (163)

Insbesondere: Existiert eine Fundamentalmatrix $t \mapsto M(t)$, so ist \mathbb{L} ein Vektorraum der Dimension n (über \mathbb{C}).

Diese Eigenschaft rechtfertigt den Namen ‘Fundamentalsystem’.

Dankenswerterweise ist es nun so, dass wir zu jeder linearen DGL ein Fundamentalsystem finden können. Beweisen können wir das aber leider erst später.

Satz 37. EXISTENZ EINER FUNDAMENTALMATRIX.

Es sei $A \in C^0(I; \mathbb{C}^{n \times n})$. Zu jedem linearen DGL-System der Form $y'(t) = A(t)y(t)$, $(t \in I)$ gibt es eine Fundamentalmatrix $t \mapsto M(t)$ in $C^1(I; \mathbb{C}^{n \times n})$. Insbesondere bildet die Menge aller Lösungen einen ($\mathbb{C}-$)Vektorraum der Dimension n . Ferner hat für alle $y_0 \in \mathbb{C}^n$ und $t_0 \in I$ das Anfangswertproblem

$$\begin{cases} y'(t) = A(t)y(t) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (164)$$

eine eindeutige Lösung, gegeben durch

$$y(t) = M(t)M(t_0)^{-1}y_0 \quad (t \in I). \quad (165)$$

(Noch unvollständiger) Beweis. Die Existenz einer Fundamentalmatrix zeigen wir später. (Falls $A(t) \equiv A = \text{const.}$ haben wir es eigentlich schon bewiesen, da wir $t \mapsto M(t)$ stets explizit angeben können. Man müsste aber an dieser Stelle dann noch die Invertierbarkeit von $M(t)$ überprüfen (HA)). Die Existenz einer Fundamentalmatrix bedingt, dass die allgemeine Lösung ein Vektorraum der Dimension n ist, siehe Merksatz 36 (p. 32). Auch die eindeutige Existenz einer Lösung eines Anfangswertproblems haben wir bereits in der Diskussion vor (160) diskutiert. Die Kurzfassung nochmal hier: Jede Lösung der DGL ist von der Gestalt $y(t) = M(t)c$ für ein $c \in \mathbb{C}^n$. Damit der Anfangswert $y(t_0) = y_0$ angenommen wird muss gelten $M(t_0)c = y_0$. Da $M(t_0)$ invertierbar ist, besitzt dieses lineare Gleichungssystem genau eine Lösung $c = M(t_0)^{-1}y_0$. Daher ist die eindeutige Lösung von (164) gegeben durch

$$y(t) = M(t)c = M(t)M(t_0)^{-1}y_0. \quad (166)$$

Mit einer Fundamentalmatrix kann man also DGLs eindeutig lösen. Man beachte jedoch, dass Fundamentalmatrizen selbst nicht eindeutig durch die DGL bestimmt sind:

***-Beispiel 38. NICHTEINDEUTIGKEIT VON FUNDAMENTALMATRIZEN.**

Es sei eine lineare DGL gegeben, die eine Fundamentalmatrix $t \mapsto M(t)$ besitzt. Sei zusätzlich $B \in \mathbb{C}^{n \times n}$ eine beliebige invertierbare Matrix.

BEHAUPTUNG: Dann ist $t \mapsto M(t)B$ auch eine Fundamentalmatrix.

BEWEIS: Da $M(t)$ ein Fundamentalsystem ist gibt es zu jeder Lösung $y \in \mathbb{L}$ ein $c \in \mathbb{C}^n$ mit

$$y(t) = M(t)c \quad \forall t \in \mathbb{R}. \quad (167)$$

Setzen wir nun $\tilde{c} := B^{-1}c$, so erhalten wir

$$y(t) = M(t)B\tilde{c} \quad \forall t \in \mathbb{R}. \quad (168)$$

Dies zeigt, dass

$$\mathbb{L} \subset \{y \in C^1(\mathbb{R}, \mathbb{C}^n) : y(t) = (M(t)B)\tilde{c} \text{ für ein } \tilde{c} \in \mathbb{C}^n\}. \quad (169)$$

Man erhält in der obigen Mengeninklusion auch Gleichheit, indem man leicht nachrechnet, dass jedes Element $y(t) = M(t)B\tilde{c}$ auf der rechten Seite die Gleichung löst [denn schließlich hat es ja die Gestalt $y(t) = M(t)d$ für ein $d \in \mathbb{C}^n$, $d = B\tilde{c}$]. Zum Abschluss ist dann nur noch zu zeigen, dass $M(t)B$ für jedes $t \in \mathbb{R}$ invertierbar ist. Dies folgt aber auch leicht daraus, dass

$$\det(M(t)B) = \det(M(t))\det(B) \neq 0 \quad \forall t \in I. \quad (170)$$

Eine Fundamentalmatrix hat für uns einen sehr großen Nutzen: Wir können mit ihr sogar das *inhomogene Problem*

$$y'(t) = A(t)y(t) + b(t) \quad (t \in \mathbb{R}) \quad (171)$$

für ein $b \in C^0(\mathbb{R}; \mathbb{C}^n)$ lösen. Ist $t \mapsto M(t)$ eine Fundamentalmatrix für das homogene Problem, so können wir die *Variation der Konstanten* (siehe Beweisreflexion 20 (p. 19) und

Beispiel 21 (p. 20)) anwenden: Wieder einmal führt uns ein Ansatz der Form

$$y(t) = M(t)c(t) \quad (t \in \mathbb{R}) \quad (172)$$

für eine Funktion $c \in C^1(\mathbb{R}; \mathbb{C}^n)$ zum Ziel.

Übung 39. INHOMOGENE SYSTEME.

Es seien $\alpha, \omega, d \in \mathbb{R}$. Bestimmen Sie die allgemeine Lösung von

$$y'(t) = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} y(t) + \begin{pmatrix} 0 \\ d \sin(\alpha t) \end{pmatrix}. \quad (173)$$

Zeigen Sie: Falls $\alpha \neq \pm\omega$, so sind alle Lösungen auf $I = \mathbb{R}$ beschränkt. Falls $\alpha = \pm\omega$, so gibt es unbeschränkte Lösungen.

HINWEIS: Die Lösung des zugehörigen homogenen Problems ist in Anwendung 51 (p. 49) durchgeführt. Die Rechnungen dort dürfen benutzt werden um die Fundamentalmatrix $t \mapsto M(t)$ zu bestimmen. Damit sollte der Ansatz in (172) in diesem konkreten Fall gut durchführbar sein.

SINN DER ÜBUNG. Wir werden sehen, dass diese DGL auftaucht, wenn wir die Schwingung eines *getriebenen harmonischen Oszillators* beschreiben wollen. Der Fall $\alpha = \pm\omega$, der sogenannte *Resonanzfall*, nimmt wegen der unbeschränkten Lösungen eine besondere Bedeutung ein. Mit diesem Fall kann erklärt werden, warum an beim Schreiben an der Tafel manchmal die Kreide quietscht.

Methodisch ist die Aufgabe auch interessant: Wir üben hier auch nochmal die Variation der Konstanten in einer komplexeren Situation ein.

1.2.4 Dekomplexifizierung

Zum Abschluss der Sektion(en) über lineare Systeme machen wir noch eine Bemerkung zur *Komplexifizierung*. Bevor wir lineare Systeme studiert haben, haben wir immer *reellwertige* Lösungen in $y \in C^1(\mathbb{R}, \mathbb{R}^n)$ gesucht. Für die Gleichung

$$y'(t) = Ay(t) \quad (t \in \mathbb{R}) \quad (174)$$

erhalten wir aber mit unseren Methoden stets eine *komplexwertige* Lösung $y \in C^1(\mathbb{R}; \mathbb{C}^n)$, auch wenn $A \in \mathbb{R}^{n \times n}$ nur reelle Einträge hat! Komplexwertige Lösungen sind aber meistens physikalisch nicht sinnvoll. Daher müssen wir die Reellwertigkeit von Lösungen aktiv untersuchen! Dies werden wir hier tun, und zwar bereits für die allgemeinere Gleichung $y'(t) = A(t)y(t)$, $(t \in I)$ wobei $A \in C^0(I; \mathbb{R}^{n \times n})$ eine (zeitabhängige, reellwertige) Matrix ist.

Wir zeigen: Für jedes Anfangswertproblem mit reellwertigen Daten ist die Reellwertigkeit der Lösung garantiert! Diese Aussage ist uns wichtig, denn wir können damit zeigen, dass sich unsere Komplexifizierung *rückgängig machen* lässt.

Proposition 40. REELLWERTIGE LÖSUNGEN. Es sei $A \in C^0(I; \mathbb{R}^{n \times n})$ eine zeitabhängige Matrix mit nur reellwertigen Einträgen und die DGL

$$y'(t) = A(t)y(t) \quad (t \in I) \quad (175)$$

gegeben.

- (i) Ist $y \in C^1(I; \mathbb{C}^n)$ eine Lösung, so sind $t \mapsto \operatorname{Re}(y(t))$ und $t \mapsto \operatorname{Im}(y(t))$ auch Lösungen. (Hierbei sind Re und Im komponentenweise zu interpretieren)
- (ii) Für alle $y_0 \in \mathbb{R}^n$ hat jedes Anfangswertproblem

$$\begin{cases} z'(t) = A(t)z(t) & (t \in I) \\ z(t_0) = y_0 \end{cases} \quad (176)$$

eine eindeutige reellwertige Lösung $z \in C^1(I, \mathbb{R}^n)$.

Beweis. Punkt (i). Dies zeigen wir durch direktes Einsetzen. Es sei $\tilde{y}(t) := \operatorname{Re}(y(t))$ für $t \in \mathbb{R}$, d.h. $\tilde{y}_i(t) = \operatorname{Re}(y_i(t))$ für alle $i = 1, \dots, n$. Dann gilt

$$\tilde{y}'(t) = \operatorname{Re}(y'(t)) = \operatorname{Re}(A(t)y(t)) = \begin{pmatrix} \operatorname{Re}(\sum_{j=1}^n a_{1j}(t)y_j(t)) \\ \vdots \\ \operatorname{Re}(\sum_{j=1}^n a_{nj}(t)y_j(t)) \end{pmatrix} \quad (177)$$

$$\stackrel{a_{ij}(t) \in \mathbb{R} \forall i,j}{=} \begin{pmatrix} \sum_{j=1}^n a_{1j}(t)\operatorname{Re}(y_j(t)) \\ \vdots \\ \sum_{j=1}^n a_{nj}(t)\operatorname{Re}(y_j(t)) \end{pmatrix} \stackrel{\text{Def. von } \tilde{y}_i}{=} \begin{pmatrix} \sum_{j=1}^n a_{1j}(t)\tilde{y}_j(t) \\ \vdots \\ \sum_{j=1}^n a_{nj}(t)\tilde{y}_j(t) \end{pmatrix} = A\tilde{y}(t), \quad (178)$$

wobei der letzte Schritt einfach die Definition der Matrixmultiplikation (rückwärts gelesen) ist. Man beachte, dass wir am Anfang der zweiten Zeile die \mathbb{R} -Linearität des Realteils benutzt haben, d.h. $\operatorname{Re}(\alpha z) = \alpha \operatorname{Re}(z)$ für alle $z \in \mathbb{C}$ und $\alpha \in \mathbb{R}$. Wir haben gezeigt, dass $\tilde{y}'(t) = A\tilde{y}(t)$ für alle $t \in \mathbb{R}$ und somit löst $\tilde{y} = \operatorname{Re}(y)$ die Gleichung. Analog könnte man dieselbe Rechnung mit dem Imaginärteil machen. Wir überlassen sie an dieser Stelle den Lesenden als Übungsaufgabe.

Punkt (ii). Wir wissen aus Satz 37 (p. 32), dass eine eindeutige (nicht zwingend reellwertige) Lösung $z \in C^1(I; \mathbb{C}^n)$ existiert. Es muss noch gezeigt werden, dass $z \in C^1(I; \mathbb{R}^n)$ liegt, d.h. reellwertig ist. Dazu: Wir führen uns vor Augen, dass nach Punkt (i) auch $\tilde{z}(t) := \operatorname{Re}(z(t))$ eine Lösung der DGL ist. Ferner hat man

$$\tilde{z}(t_0) = \operatorname{Re}(z(t_0)) = \operatorname{Re}(y_0) = y_0, \quad (179)$$

da $y_0 \in \mathbb{R}^n$ reellwertig ist. Mit anderen Worten gilt dann

$$\begin{cases} \tilde{z}'(t) = A(t)\tilde{z}(t) & (t \in I) \\ \tilde{z}(t_0) = y_0. \end{cases} \quad (180)$$

Daher löst \tilde{z} dasselbe Anfangswertproblem wie z , nämlich (176). Nach Satz 37 (p. 32) sind Lösungen von linearen Anfangswertproblemen (176) aber eindeutig! Es folgt, dass

$z = \tilde{z}$, d.h. mit anderen Worten $z(t) = \operatorname{Re}(z(t))$ für alle $t \in I$. Wir schließen daraus, dass $z(t) \in \mathbb{R}^{n \times n}$ für alle $t \in I$. Daher gilt, dass die eindeutige Lösung reellwertig ist.

Eine berechtigte Frage ist nun: Wenn wir ein reellwertiges DGL-System gegeben haben, finden wir dann auch eine reellwertige Fundamentalmatrix? Die Antwort ist in der Tat “ja”!

***-Proposition 41.** Es sei $A \in C^0(I; \mathbb{R}^{n \times n})$ eine zeitabhängige reellwertige Matrix und die DGL

$$y'(t) = A(t)y(t) \quad (t \in I) \quad (181)$$

gegeben. Dann gibt es eine reellwertige Fundamentalmatrix $M \in C^1(I; \mathbb{R}^{n \times n})$.

Beweis. Satz 37 (p. 32) gibt uns die Existenz einer komplexwertigen Fundamentalmatrix $\tilde{M} \in C^1(I; \mathbb{C}^{n \times n})$. Definiere nun für ein beliebiges $t_0 \in I$

$$M(t) := \tilde{M}(t)\tilde{M}(t_0)^{-1} \quad (t \in I). \quad (182)$$

Wie wir in Beispiel 38 (p. 33) gesehen haben ist $t \mapsto M(t)$ auch eine Fundamentalmatrix. (Dort haben wir nämlich gezeigt, dass das Produkt einer Fundamentalmatrix mit einer beliebigen invertierbaren Matrix wieder eine Fundamentalmatrix ist).

BEHAUPTUNG: $M \in C^1(I; \mathbb{R}^{n \times n})$ hat nur reelle Einträge. Hierfür genügt es zu zeigen, dass

$$M(t)e_1, \dots, M(t)e_n \in \mathbb{R}^n \quad \forall t \in I. \quad (183)$$

In der Tat: Da $M(t)e_i$ gerade die i -te Spalte von $M(t)$ in einen Vektor geschrieben enthält, würde aus (183) folgen, dass $M(t)$ reelle Spaltenvektoren hat uns somit reellwertig ist. Beachte nun, dass für alle $i \in \{1, \dots, n\}$ gilt

$$M(t)e_i = \tilde{M}(t)\tilde{M}(t_0)^{-1}e_i. \quad (184)$$

Jetzt ist nach Satz 37 (p. 32), genauer nach 165, aber

$$y(t) := \tilde{M}(t)\tilde{M}(t_0)^{-1}e_i \quad (t \in I) \quad (185)$$

die eindeutige Lösung zum Anfangswertproblem

$$\begin{cases} y'(t) = A(t)y(t) & (t \in I) \\ y(t_0) = e_i. \end{cases} \quad (186)$$

Diese Lösung ist nach Proposition 40 (p. 34) aber \mathbb{R}^n -wertig, weshalb für alle $t \in I$ gilt, dass

$$M(t)e_i = \tilde{M}(t)\tilde{M}(t_0)^{-1}e_i = y(t) \in \mathbb{R}^n. \quad (187)$$

Es folgt wie oben beschrieben die Reellwertigkeit der Fundamentalmatrix.

Man beachte, dass eine reellwertige Fundamentalmatrix auch ein reellwertiges Fundamentalsystem liefert. Somit kann man also stets eine Basis des Lösungsraumes \mathbb{L} aus reellwertigen Lösungen finden.

Achtung: Nur weil es jetzt eine reellwertige Fundamentalmatrix gibt heißt das auf keinen Fall, dass jede Fundamentalmatrix reellwertig ist! Im Gegenteil: Mit Beispiel 38 (p. 33) kann man sich stets munter komplexwertige Fundamentalmatrizen konstruieren.

1.2.5 DGLs höherer Ordnung

Mittlerweile können wir schon ziemlich allgemeine lineare Differentialgleichungen lösen. Die Theorie lässt sich auch auf lineare Gleichungen *höherer Ordnung* anwenden.

“Höhere Ordnung” heißtt, dass höhere Ableitungen in der Gleichung vorkommen. Das heißtt, man hat z.B. eine Gleichung der Form

$$y''(t) = f(t, y(t), y'(t)) \quad (t \in I). \quad (188)$$

Aus historischen Gründen schreiben wir hier auch oft die gesamte DGL auf eine Seite, d.h.

$$y''(t) - f(t, y(t), y'(t)) = 0 \quad (t \in I). \quad (189)$$

Üblicherweise war die Regularität, die wir gefordert haben immer ‘ $y \in C^1(I)$ ’. Im Kontext von DGLs höherer Ordnung müssen wir natürlich höhere Regularitätsanforderungen stellen, denn schließlich müssen wir ja auch von höheren Ableitungen sprechen können. Daher ändert sich auch unsere Definition einer Lösungsbegriffs.

Definition 42. LINEARE DGLs HÖHERER ORDNUNG.

Es seien $a, b \in C^0(I; \mathbb{C})$ Eine Differentialgleichung der Form

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0 \quad (t \in I) \quad (190)$$

heißt *lineare (homogene) DGL zweiter Ordnung*. Eine *Lösung* ist eine Funktion $y \in C^2(I; \mathbb{C})$ die (190) für alle $t \in I$ erfüllt. Falls a, b konstante Funktionen sind, so heißtt die DGL *lineare DGL zweiter Ordnung mit konstanten Koeffizienten*. Allgemein heißtt eine DGL der Form

$$y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_1(t)y^{(1)}(t) + a_0(t)y_0(t) = 0 \quad (t \in I) \quad (191)$$

eine *lineare (homogene) DGL n-ter Ordnung*. Eine Lösung ist dann eine Funktion $y \in C^n(I; \mathbb{C})$ so, dass (191) erfüllt ist.

Man beachte, dass auch für solche Gleichungen das *Superpositionsprinzip* gilt, d.h. wieder einmal bildet die allgemeine Lösung

$$\mathbb{L} = \{y \in C^2(I; \mathbb{C}) : y''(t) + a(t)y'(t) + b(t)y(t) = 0 \forall t \in I\} \quad (192)$$

einen Vektorraum über \mathbb{C} .

Wir werden gleich sehen, dass sich DGLs höherer Ordnung lösen lassen, indem wir sie mithilfe einer geschickten Substitution als System von DGLs erster Ordnung schreiben. Haben die DGLs konstante Koeffizienten $a, b \equiv \text{const.}$, so lässt sich die Lösung sogar explizit bestimmen!

Gedankenexperiment 43. RÜCKFÜHRUNG AUF DGL-SYSTEME.

Wir beschränken uns hier mal (der Einfachheit halber) auf DGLs zweiter Ordnung mit konstanten Koeffizienten, d.h. auf

$$y''(t) + ay'(t) + by(t) = 0 \quad (t \in \mathbb{R}), \quad (193)$$

für $a, b \in \mathbb{C}$ zwei Konstanten. Für die Lösung verwenden wir einen einfachen Trick: Wir wandeln die DGL in ein System von DGLs erster Ordnung um. Sei hierzu y eine Lösung von (193). Wir setzen für $t \in \mathbb{R}$

$$\begin{aligned} z_0(t) &:= y(t), \\ z_1(t) &:= y'(t). \end{aligned} \quad (194)$$

Wir berechnen

$$z'_0(t) = y'(t) = z_1(t), \quad (195)$$

$$z'_1(t) = y''(t) \underset{(193)}{=} -ay'(t) - by(t) \underset{(194)}{=} -az_1(t) - bz_0(t). \quad (196)$$

Alles in Allem heißt das

$$\frac{d}{dt} \begin{pmatrix} z_0(t) \\ z_1(t) \end{pmatrix} = \begin{pmatrix} z'_0(t) \\ z'_1(t) \end{pmatrix} = \begin{pmatrix} z_1(t) \\ -bz_0(t) - az_1(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} z_0(t) \\ z_1(t) \end{pmatrix}. \quad (197)$$

Somit schließen wir, dass die Funktion

$$z(t) := \begin{pmatrix} z_0(t) \\ z_1(t) \end{pmatrix} \quad (t \in \mathbb{R}) \quad (198)$$

die folgende DGL erfüllt:

$$z'(t) = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} z(t). \quad (199)$$

Dies ist eine DGL, die wir bereits explizit lösen können! Es wird — wie im letzten Abschnitt erläutert — so sein, dass die allgemeine Lösung gegeben ist durch

$$\tilde{\mathbb{L}} = \{z \in C^1(\mathbb{R}; \mathbb{C}^2) : z(t) = M(t)c \text{ für ein } c \in \mathbb{C}^n\} \quad (200)$$

für eine Fundamentalmatrix $t \mapsto M(t)$, die wir explizit bestimmen können. Eigentlich wollen wir aber nicht die vektorwertige Lösung $z(t)$ haben, sondern die skalarwertige Gleichung (193) lösen. Wir müssen wieder resubstituieren. Dafür erinnern wir uns, dass $y(t) = z_0(t)$ die erste Komponente von $z(t)$ war. Das legt nahe, dass die allgemeine Lösung gegeben ist durch

$$\mathbb{L} = \{y \in C^1(\mathbb{R}; \mathbb{C}) : y \text{ ist erste Komponente eines Elementes aus } \tilde{\mathbb{L}}\}. \quad (201)$$

Mit (200) und der Tatsache, dass $\langle v, e_1 \rangle$ stets die erste Komponente eines Vektors $v \in \mathbb{C}^2$ liefert gilt

$$\mathbb{L} = \{y \in C^1(\mathbb{R}; \mathbb{C}) : y(t) = \langle M(t)c, e_1 \rangle\}. \quad (202)$$

Diese Formel zeigen wir im Laufe dieses Abschnittes formal und auch ohne die Annahme konstanter Koeffizienten (die wir bei Lichte betrachtet auch hier nicht sonderlich benutzt haben — sie hilft lediglich dabei die Lösung in (200) explizit zu bestimmen).

Wir haben also eine Korrespondenz zwischen linearen Systemen und linearen DGL's höherer Ordnung! Das heißt insbesondere, dass wir in den letzten Kapiteln bereits die Kenntnisse erworben haben, um DGLs höherer Ordnung zu lösen – und auch bereits gelöst haben:

Übung 44. EINE (INHOMOGENE) DGL HÖHERER ORDNUNG.

Das System aus Übung 39 (p. 34) kann verwendet werden um die allgemeine Lösung einer DGL zweiter Ordnung zu beschreiben. Geben Sie diese DGL an.

SINN DER ÜBUNG: Wir sehen an dieser Übung nochmal, dass wir auch inhomogene lineare DGLs höherer Ordnung bereits gut lösen können, weil wir ja schließlich auch inhomogene lineare Systeme lösen können. Die DGL, die hierbei auftaucht gehört zum *getriebenen harmonischen Oszillator*, welches im Ingenieurwesen sehr relevant ist und was wir später noch besprechen werden.

Das vorherige Gedankenexperiment zeigt, dass wir jede lineare DGL höherer Ordnung auf ein lineares System zurückführen können. Auch für nichtlineare DGLs höherer Ordnung ist ein Ansatz wie in (194) möglich. Wir diskutieren im Rest des Abschnittes, dass \mathbb{L} aus (201) wirklich die allgemeine Lösung ist.

Die Proposition, die wir als nächstes beweisen werden, klingt von der Formulierung her zunächst abstrakt. Sie besagt aber lediglich, dass jede Lösung einer DGL höherer Ordnung aus einer Lösung des zugehörigen Systems erhalten werden kann und umgekehrt. Mit anderen Worten: Es gibt eine 1-zu-1 Korrespondenz der beiden Mengen $\widetilde{\mathbb{L}}$ und \mathbb{L} aus Gedankenexperiment 43 (p. 38).

Proposition 45. 1-ZU-1 KORRESPONDENZ ZU LINEAREN SYSTEMEN.

Es sei für $a, b \in C^0(I; \mathbb{C})$ die DGL

$$y''(t) + a(t)y'(t) + b(t)y(t) = 0 \quad (t \in I) \quad (203)$$

gegeben. Es sei nun

$$\mathbb{L} := \{y \in C^2(I; \mathbb{C}) : y''(t) + a(t)y'(t) + b(t)y(t) = 0 \quad \forall t \in I\} \quad (204)$$

und

$$\widetilde{\mathbb{L}} := \left\{ z \in C^1(I; \mathbb{C}^2) : z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t) \quad \forall t \in I. \right\} \quad (205)$$

Dann definiert

$$\Phi : \mathbb{L} \rightarrow \widetilde{\mathbb{L}}, \quad \Phi(y) := \begin{pmatrix} y \\ y' \end{pmatrix} \quad (206)$$

einen Vektorraumisomorphismus (d.h. eine lineare bijektive Abbildung) zwischen \mathbb{L} und $\tilde{\mathbb{L}}$ (als Vektorräume über \mathbb{C}). Insbesondere gilt

$$\dim_{\mathbb{C}}(\mathbb{L}) = \dim_{\mathbb{C}}(\tilde{\mathbb{L}}) = 2. \quad (207)$$

Beweis. Es sei Φ definiert wie in der Aussage. Wir zeigen Wohldefiniertheit, Linearität, Surjektivität und Injektivität von Φ .

Schritt 1. \mathbb{Z} : Φ ist wohldefiniert, d.h. $\Phi(y) \in \tilde{\mathbb{L}}$ für alle $y \in \mathbb{L}$. Es sei $\Phi(y) = (y, y')^T$ für ein $y \in \mathbb{L}$ gegeben. Definiere $z = (z_0, z_1)^T$ mit $z_0 := y$ und $z_1 := y'$, d.h. $z = \Phi(y)$. Dann gilt

$$z'_0(t) = y'(t) = z_1(t) \quad \forall t \in I. \quad (208)$$

und (wiederum mit den Definitionen von z_0 und z_1)

$$z'_1(t) = y''(t) \stackrel{(203)}{=} -a(t)y'(t) - b(t)y(t) = -b(t)z_0(t) - a(t)z_1(t) \quad \forall t \in I. \quad (209)$$

Es ergibt sich aus den vorigen beiden Gleichungen, dass

$$\begin{pmatrix} z'_0(t) \\ z'_1(t) \end{pmatrix} = \begin{pmatrix} z_1(t) \\ -b(t)z_0(t) - a(t)z_1(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} z_0(t) \\ z_1(t) \end{pmatrix}, \quad (210)$$

wobei man für den letzten Rechenschritt am besten die Matrixmultiplikation rückwärts liest. Da $z = (z_0, z_1)^T$ schließen wir

$$z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t), \quad (211)$$

weshalb $z \in \tilde{\mathbb{L}}$ gilt. Erinnern wir uns zurück, dass $z = \Phi(y)$, so erhalten wir also $\Phi(y) \in \tilde{\mathbb{L}}$, was zu zeigen war.

Schritt 2: Φ ist linear. Dass Φ linear ist, überlassen wir den Lesenden als Übungsaufgabe (HA). Es folgt direkt aus der Linearität der Ableitung

$y \mapsto y'$ (als Abbildung von $C^2(I; \mathbb{C})$ nach $C^1(I; \mathbb{C})$).

Schritt 3: Φ ist surjektiv. Dazu sei $z = (z_0, z_1)^T \in \tilde{\mathbb{L}}$ beliebig vorgegeben. Es ist zu zeigen, dass $z = \Phi(y)$ für ein $y \in \mathbb{L}$. Da $z \in \tilde{\mathbb{L}}$ gilt

$$\begin{pmatrix} z'_0(t) \\ z'_1(t) \end{pmatrix} = z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \begin{pmatrix} z_0(t) \\ z_1(t) \end{pmatrix}. \quad (212)$$

Ausmultipliziert heißt das

$$\begin{pmatrix} z'_0(t) \\ z'_1(t) \end{pmatrix} = \begin{pmatrix} z_1(t) \\ -b(t)z_0(t) - a(t)z_1(t) \end{pmatrix}. \quad (213)$$

Vergleichen der ersten Komponenten ergibt

$$z'_0(t) = z_1(t) \quad \forall t \in I. \quad (214)$$

Dies zeigt (da $z_1 \in C^1(I; \mathbb{C})$) dass $z'_0 \in C^1(I; \mathbb{C})$, also $z_0 \in C^2(I; \mathbb{C})$. Mit diesem Wissen berechnen wir eine Gleichung für die zweite Ableitung von z_0 . Es gilt für alle $t \in I$

$$z''_0(t) = (z'_0(t))' \stackrel{(214)}{=} z_1(t)' = -b(t)z_0(t) - a(t)z_1(t) \quad (215)$$

$$\stackrel{(214)}{=} -b(t)z_0(t) - a(t)z'_0(t). \quad (216)$$

Bringen wir alle Terme auf eine Seite, so erhalten wir

$$z''_0(t) + a(t)z'_0(t) + b(t)z_0(t) = 0. \quad (217)$$

Insbesondere ist also $z_0 \in \mathbb{L}$, wie man direkt aus der Definition von \mathbb{L} (227) abliest. Man erhält, dass

$$z = \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} \stackrel{(214)}{=} \begin{pmatrix} z_0 \\ z'_0 \end{pmatrix} = \Phi(z_0), \quad (218)$$

wobei $z_0 \in \mathbb{L}$. Setzt man nun $y = z_0$, so erhält man, dass $z = \Phi(y)$ für ein $y \in \mathbb{L}$.

Schritt 4: Wir zeigen die Injektivität von Φ . Hierfür genügt es nach einem Satz aus der linearen Algebra zu zeigen, dass der *Kern*

$$\ker(\Phi) := \{y \in \mathbb{L} : \Phi(y) = 0\} \quad (219)$$

nur aus dem Nullvektor besteht. Es sei also $y \in \ker(\Phi)$, d.h. $y \in \mathbb{L}$ mit $\Phi(y) = 0$, wobei hier mit 0 der Nullvektor in $\tilde{\mathbb{L}} \subset C^1(\mathbb{R}, \mathbb{C}^2)$ gemeint ist, d.h. die Abbildung z definiert durch $z(t) := (0, 0)^T$ für alle $t \in I$. Dann gilt

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \Phi(y) = \begin{pmatrix} y \\ y' \end{pmatrix}, \quad (220)$$

woraus wir durch Vergleich der ersten Komponenten $y = 0$ erhalten. Es folgt, dass $\ker(\Phi) = \{0\}$ und somit die Injektivität von Φ . Es ist nun gezeigt, dass Φ ein Vektorraumisomorphismus ist. Die Dimensionsgleichung (207) folgt direkt aus einem Resultat aus der linearen Algebra: Zwei isomorphe Vektorräume haben dieselbe Dimension! Dass $\dim_{\mathbb{C}}(\tilde{\mathbb{L}}) = 2$ folgt aus Merksatz 36 (p. 32).

Der Vektorraumisomorphismus oben kann nun benutzt werden um *Anfangswertprobleme* richtig aufzustellen und zu lösen. Wir müssen uns allerdings erstmal im Klaren werden, was ein Anfangswertproblem im Kontext von DGLs höherer Ordnung überhaupt sein soll... Ganz intuitiv würde man vielleicht wie üblich Probleme wie

$$\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (221)$$

für ein $y_0 \in \mathbb{C}$ studieren wollen. Gibt es hierbei eine eindeutige Lösung? Leider nein! Dies

geht schon allein aus Dimensionsargumenten nicht: Laut der vorigen Proposition hat

$$\mathbb{L} := \{y \in C^2(I; \mathbb{C}) : y''(t) + a(t)y'(t) + b(t)y(t) = 0 \quad \forall t \in I\} \quad (222)$$

Dimension 2, es gibt also eine zweielementige Basis. Wähle nun $b_1, b_2 \in \mathbb{L}$ Funktionen, sodass $\{b_1, b_2\}$ eine Basis von \mathbb{L} ist. Es folgt, dass jede Lösung y von (221) eine Linearkombination aus b_1 und b_2 sein muss, d.h. es gibt $c_1, c_2 \in \mathbb{C}$ mit

$$y(t) = c_1 b_1(t) + c_2 b_2(t). \quad (223)$$

Falls (221) eine eindeutige Lösung hat, müssen sich die unbekannten Parameter c_1, c_2 eindeutig bestimmen lassen. Nun ist die einzige Bedingung, die in (221) gestellt wird, dass $y(t_0) = y_0$. Wir setzen

$$y_0 \stackrel{!}{=} y(t_0) = c_1 b_1(t_0) + c_2 b_2(t_0). \quad (224)$$

Dies ist eine lineare Gleichung für zwei Unbekannte $c_1, c_2 \in \mathbb{C}$. Aus der Linearen Algebra wissen wir: Gibt es bei einem linearen Gleichungssystem weniger Gleichungen als Unbekannte, so ist die Lösung niemals eindeutig. Ein herber Rückschlag für die Theorie der Anfangswertprobleme ...

Man sieht: Damit es eindeutige Lösungen gibt müssen wir mehr vorschreiben als nur einen Anfangswert! Die Frage ist nur: Was müssen wir noch vorschreiben? Hier können wir aus der Definition des Vektorraumisomorphismus Φ (siehe Proposition 45, p. 39) eine Antwort herauslesen: Wir benötigen auch die *Anfangsgeschwindigkeit*, also die erste Ableitung am Zeitpunkt t_0 !

Korollar 46. Es seien $a, b \in C^0(I; \mathbb{C})$, $y_0, y_1 \in \mathbb{C}$, $t_0 \in I$. Dann hat das Anfangswertproblem

$$\begin{cases} y''(t) + a(t)y'(t) + b(t)y(t) = 0 & (t \in I) \\ y(t_0) = y_0 \\ y'(t_0) = y_1 \end{cases} \quad (225)$$

eine eindeutige Lösung $y \in C^2(I; \mathbb{C})$. Diese ist gegeben durch $y = \Phi^{-1}(z)$, wobei Φ wie in Proposition 45 (p. 39) und $z \in C^1(I; \mathbb{C}^2)$ die eindeutige Lösung von

$$\begin{cases} z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t) \\ z(t_0) = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}. \end{cases} \quad (226)$$

Zusatz: $y = \Phi^{-1}(z)$ ist gerade die erste Komponente von z .

Beweisidee. Es seien \mathbb{L} und $\tilde{\mathbb{L}}$ definiert wie in Proposition 45 (p. 39), d.h.

$$\mathbb{L} := \{y \in C^2(I; \mathbb{C}) : y''(t) + a(t)y'(t) + b(t)y(t) = 0\} \quad (227)$$

und

$$\tilde{\mathbb{L}} := \left\{ z \in C^1(I; \mathbb{C}^2) : z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t) \right\}. \quad (228)$$

Es sei ferner $\Phi : \mathbb{L} \rightarrow \tilde{\mathbb{L}}$, gegeben durch

$$\Phi(y) = \begin{pmatrix} y \\ y' \end{pmatrix} \quad (229)$$

der Vektorraumisomorphismus aus Proposition 45 (p. 39). Es sei nun $y \in C^2(I; \mathbb{C})$ Lösung von (225). Dann gilt $y \in \mathbb{L}$, $y(t_0) = y_0$ und $y'(t_0) = y_1$. Man beachte: $z = \Phi(y)$ liegt in $\tilde{\mathbb{L}}$ und außerdem gilt

$$z(t_0) = (\Phi(y))(t_0) = \begin{pmatrix} y(t_0) \\ y'(t_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}. \quad (230)$$

Mit anderen Worten löst z das Anfangswertproblem

$$\begin{cases} z'(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} z(t) \\ z(t_0) = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}. \end{cases} \quad (231)$$

Dieses AWP hat nach Satz 37 (p. 32) eine eindeutige Lösung! Das zeigt: $z = \Phi(y)$ ist durch (225) eindeutig bestimmt. Da Φ bijektiv ist, ist aber auch $y = \Phi^{-1}(z)$ eindeutig bestimmt. Diese Aussage könnte man jetzt auch "schulbuchmäßig" beweisen, d.h. Existenz und Eindeutigkeit einzeln zeigen. Wir verzichten an dieser Stelle darauf und überlassen mit den oben angegebenen Ideen den Rest den Lesenden als Übungsaufgabe (HA).

NOCH ZUM ZUSATZ:

Dass $\Phi^{-1}(z)$ gerade die erste Komponente von $z = (z_0, z_1)^T$ ist, folgt direkt aus der Definition von Φ , d.h. aus $\Phi(y) = (y, y')^T$.

Übung 47. Wandeln Sie die obige Beweisidee in einen formal richtig aufgeschriebenen Existenz -und Eindeutigkeitsbeweis um.

Abschließend sei hier Folgendes gesagt: Die Beobachtungen und Techniken übertragen sich völlig analog auch auf DGLs n -ter Ordnung für $n \geq 2$. Die Notation ist hier nur wesentlich aufwändiger, denn die DGL wird stets in ein n -dimensionales System umgewandelt! Wir stellen hier die Aussagen der Vollständigkeit halber zur Verfügung, es lohnt aber nicht, die Aussagen im Detail zu besprechen.

***-Proposition 48. LINEARE DGLS HÖHERER ORDNUNG.**

Es sei für $a_0, \dots, a_{n-1} \in C^0(I; \mathbb{C})$ die DGL

$$y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_1(t)y^{(1)}(t) + a_0(t)y(t) = 0 \quad (t \in I) \quad (232)$$

gegeben. Es sei

$$A(t) := \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & & 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & \cdots & & \cdots & -a_{n-1}(t) \end{pmatrix}. \quad (233)$$

Falls $\tilde{\mathbb{L}}$ die allgemeine Lösung von $z'(t) = A(t)z(t), (t \in I)$ ist, so ist

$$\mathbb{L} := \{y \in C^n(I; \mathbb{C}) : y \text{ ist erste Komponente eines Elementes von } \tilde{\mathbb{L}}\} \quad (234)$$

die allgemeine Lösung von (232). Ferner ist

$$\Phi : \mathbb{L} \rightarrow \tilde{\mathbb{L}}, \quad \Phi : y \mapsto \begin{pmatrix} y \\ y' \\ \dots \\ y^{(n-1)} \end{pmatrix} \quad (235)$$

ein Vektorraumisomorphismus zwischen \mathbb{L} und $\tilde{\mathbb{L}}$.

Beweis. SCHRITT 1. $Z_{\mathbb{Z}}$: Jede Lösung $y \in C^n(\mathbb{R}; \mathbb{C})$ ist ein Element von \mathbb{L} (definiert wie in (234)). Setze für $k = 0, \dots, n-1$

$$z_k(t) := y^{(k)}(t). \quad (236)$$

Für

$$z(t) = \begin{pmatrix} z_0(t) \\ z_1(t) \\ \vdots \\ z_{n-1}(t) \end{pmatrix} \quad (237)$$

gilt

$$z'(t) = \begin{pmatrix} z'_0(t) \\ z'_1(t) \\ \vdots \\ z'_{n-1}(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ y''(t) \\ \vdots \\ y^{(n)}(t) \end{pmatrix} \quad (238)$$

$$\stackrel{(232)}{=} \begin{pmatrix} y'(t) \\ y''(t) \\ \vdots \\ -a_0(t)y(t) - a_1(t)y'(t) - \dots - a_{n-1}(t)y^{(n-1)}(t) \end{pmatrix} \quad (239)$$

$$\stackrel{(236)}{=} \begin{pmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ -a_0(t)z_0(t) - a_1(t)z_1(t) - \dots - a_{n-1}(t)z_{n-1}(t) \end{pmatrix}. \quad (240)$$

Zusätzlich rechnet man leicht nach, dass

$$A(t)z(t) = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & & 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & \dots & & \dots & -a_{n-1}(t) \end{pmatrix} \begin{pmatrix} z_0(t) \\ z_1(t) \\ \vdots \\ z_{n-3}(t) \\ z_{n-2}(t) \\ z_{n-1}(t) \end{pmatrix} \quad (241)$$

$$= \begin{pmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_{n-2}(t) \\ z_{n-1}(t) \\ -a_0(t)z_0(t) - a_1(t)z_1(t) - \dots - a_{n-1}(t)z_{n-1}(t) \end{pmatrix}. \quad (242)$$

Vergleicht man die letzten beiden Gleichungen, so erhält man $z'(t) = A(t)z(t)$ für alle $t \in \mathbb{R}$. Das heißt, dass $z \in \tilde{\mathbb{L}}$. Da $y(t) = z_0(t)$ ist also y die erste Komponente eines Elementes von $\tilde{\mathbb{L}}$. Wir folgern, dass $y \in \mathbb{L}$.

SCHRITT 2. \mathbb{Z}_2 : Jedes Element von \mathbb{L} ist eine Lösung von (232). Es sei nun $y \in \mathbb{L}$. Dann gibt es ein $z = (z_0, \dots, z_{n-1})^T \in \tilde{\mathbb{L}}$ mit $y = z_0$. Man beachte, dass wir unsere Indizierung hier mit Null beginnen, das heißt

$$z_k(t) = \langle z(t), e_{k+1} \rangle \quad \forall k = 0, \dots, n-1. \quad (243)$$

Wir zeigen nun mit Induktion nach k für $k \in 0, \dots, n-1$

$$y'(t) = z'_0(t) = \langle z'(t), e_1 \rangle = \langle A(t)z(t), e_1 \rangle, \quad (244)$$

wobei wir hier benutzt haben, dass die erste Komponente eines Vektors $v \in \mathbb{C}^n$ stets durch $\langle v, e_1 \rangle$ gegeben ist. Exakt wie in (242) folgern wir

$$A(t)z(t) = \begin{pmatrix} z_1(t) \\ z_2(t) \\ \vdots \\ z_{n-2}(t) \\ z_{n-1}(t) \\ -a_0(t)z_0(t) - a_1(t)z_1(t) - \dots - a_{n-1}(t)z_{n-1}(t) \end{pmatrix}. \quad (245)$$

Daher ist

$$y'(t) = \langle A(t)z(t), e_1 \rangle = z_1(t). \quad (246)$$

Als nächstes zeigen wir mit Induktion

$$y^{(k)}(t) = z_k(t) \quad \forall k = 0, \dots, n-1. \quad (247)$$

Für $k = 0, 1$ sind wir fertig. Es gelte die Behauptung nun für ein $k \in \{0, \dots, n-2\}$. Dann gilt auch

$$y^{(k+1)}(t) = \frac{d}{dt}y^{(k)}(t) \stackrel{\text{Ind.Hyp.}}{=} \frac{d}{dt}z_k(t) \stackrel{(243)}{=} \frac{d}{dt}\langle z(t), e_{k+1} \rangle \quad (248)$$

$$= \langle z'(t), e_{k+1} \rangle = \langle A(t)z(t), e_{k+1} \rangle \stackrel{(245)}{=} z_{k+1}(t). \quad (249)$$

Dadurch folgt (247). Mit (247) folgern wir nun

$$y^{(n)}(t) = \frac{d}{dt}y^{(n-1)}(t) \stackrel{(247)}{=} \frac{d}{dt}z_{n-1}(t) = \frac{d}{dt}\langle z(t), e_n \rangle = \langle A(t)z(t), e_n \rangle \quad (250)$$

$$= -a_0(t)z_0(t) - \dots - a_{n-1}(t)z_{n-1}(t) \quad (251)$$

$$= -a_0(t)y(t) - \dots - a_{n-1}(t)y^{(n-1)}(t). \quad (252)$$

Addiert man alles auf die linke Seite, so merkt man, dass y eine Lösung von (232) ist.

SCHRITT 3: Vektorraumisomorphismuseigenschaft von Φ . Studiert man das Vorgehen in den beiden vorangegangenen Beweisschritten, so sieht man bereits, dass Φ eine Bijektion ist. Dass Φ linear ist, überlassen wir den Lesenden als Übungsaufgabe.

Wieder mal können wir den gefundenen Vektorraumisomorphismus Φ verwenden, um Anfangswertprobleme eindeutig zu lösen:

***-Beobachtung 49.** Für vorgeschriebene $y_0, \dots, y_{n-1} \in \mathbb{C}$ hat das Anfangswertproblem

$$\begin{cases} y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \dots + a_0(t)y(t) = 0 & (t \in I) \\ y(t_0) = y_0 \\ y'(t_0) = y_1 \\ \vdots \\ y^{(n-1)}(t_0) = y_{n-1} \end{cases} \quad (253)$$

eine eindeutige Lösung. Diese ist gegeben durch $y = \Phi^{-1}(z)$, wobei Φ der Vektorraumisomorphismus aus der vorangegangenen Proposition ist und $z \in C^1(I; \mathbb{C}^n)$ die eindeutige Lösung von

$$\begin{cases} z'(t) = A(t)z(t) & (t \in I) \\ z(t_0) = \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix} \end{cases} \quad (254)$$

ist. Hierbei ist

$$A(t) := \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & & 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & \dots & & \dots & -a_{n-1}(t) \end{pmatrix}. \quad (255)$$

Übung 50. EINE STANDARDAUFGABE ÜBER DGLs HÖHERER ORDNUNG. Bestimmen Sie die allgemeine Lösung der DGL

$$y''(t) + 3y'(t) + 2y(t) = 0. \quad (256)$$

SINN DER ÜBUNG: Hier soll man eine Gelegenheit bekommen, die in dem Kapitel gelernten Techniken anzuwenden – Die Lösung lässt sich explizit bestimmen!

1.2.6 Anwendung: Der harmonische Oszillator

Im Folgenden studieren wir Schwingungen. Wir werden herleiten, wie Schwingungen mithilfe von Differentialgleichungen beschrieben werden können. Unser Ausgangspunkt hierbei ist die *Newton'sche Bewegungsgleichung*. Für ein Teilchen, welches sich in einem Kraftfeld befindet, d.h. auf das eine Kraft wirkt, gilt nach einem Axiom der Physik

$$\vec{F} = m\vec{a}, \quad (257)$$

wobei \vec{F} eine Kraft auf das Teilchen beschreibt, m dessen Masse und \vec{a} dessen Beschleunigung. Die Gleichung kann wie folgt als Differentialgleichung aufgefasst werden: Es sei $\vec{x} = \vec{x}(t)$ der Ortsvektor des Teilchens. Die Beschleunigung des Teilchens ist dann $\vec{a}(t) = \vec{x}''(t)$. Die Kraft \vec{F} hängt meist nur von dem Ort des Teilchens ab, d.h. $\vec{F} = \vec{F}(\vec{x}(t))$. Im Allgemeinen kann aber auch $\vec{F} = \vec{F}(t, \vec{x}(t), \vec{x}'(t))$ gelten, d.h. die Kraft kann auch von Zeit und Geschwindigkeit abhängen. Nach den Newton'schen Bewegungsgleichung gilt dann

$$m\vec{x}''(t) = \vec{F}(t, \vec{x}(t), \vec{x}'(t)). \quad (258)$$

Teilt man durch die Masse m so erhält man eine übliche DGL zweiter Ordnung.

Eine einfache Anwendung ist das *Federpendel*. Beim Federpendel schwingt ein Massestück

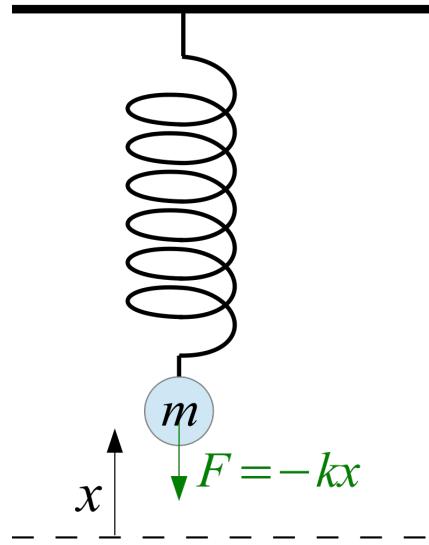


Abbildung 1: Das Federpendel

der Masse m , welches an einer Feder befestigt ist hin und her. Die Auslenkung des Pendels $x = x(t)$ kann durch eine reellwertige Funktion beschrieben werden. Wir wählen ein Koordinatensystem, sodass $x = 0$ gerade der *Ruhelage* des Pendels entspricht. D.h. der Koordinatenursprung befindet sich an dem Punkt, wo das Massestück hinge, wenn es nicht schwingen würde. Die Beschleunigung ist dann gegeben durch $x''(t)$. Die Kraft, die auf das Teilchen wirkt, ist gegeben durch die sogenannte *Rückstellkraft* der Feder. Nach dem *Hooke'schen Gesetz* ist diese gegeben durch $F(x) = -kx$ für eine *Federkonstante* $k > 0$, die aus den Materialeigenschaften der Feder berechnet werden kann. Die Newton'sche Bewegungsgleichung liefert damit also

$$mx''(t) = F(x(t)) = -kx(t), \quad (259)$$

mit anderen Worten

$$x''(t) = -\frac{k}{m}x(t). \quad (260)$$

Setzen wir $\omega := \sqrt{\frac{k}{m}}$, so erhalten wir

$$x''(t) + \omega^2 x(t) = 0. \quad (261)$$

Dieses Kapitel ist der Lösung dieser Gleichung gewidmet.

Anwendung 51. BEWEGUNGSGLEICHUNG FÜR DAS FEDERPENDEL. Es sei $\omega > 0$ beliebig vorgegeben. Die Gleichung

$$x''(t) + \omega^2 x(t) = 0 \quad (262)$$

ist eine DGL zweiter Ordnung, deren allgemeine Lösung wir bestimmen wollen. Da eine reellwertige physikalische Größe beschrieben wird, sind wir auch nur an reellwertigen Lösungen interessiert. Das heißt wir wollen Anfangswertprobleme der Form

$$\begin{cases} x''(t) + \omega^2 x(t) = 0 \\ x(0) = x_0 \\ x'(0) = v_0 \end{cases} \quad (263)$$

lösen, wobei $x_0 \in \mathbb{R}$ die Anfangsauslenkung und $v_0 \in \mathbb{R}$ die Anfangsgeschwindigkeit ist. Wir haben bereits im letzten Kapitel besprochen, warum wir Anfangsort *und* Anfangsgeschwindigkeit vorschreiben müssen. Vielleicht kann man sich vor dem hiesigen physikalischen Hintergrund auch gut vorstellen, dass es wirklich notwendig ist die Anfangsgeschwindigkeit zusätzlich vorzugeben.

SCHRITT 1: Bestimmen der allgemeinen Lösung. Die Allgemeine Lösung bestimmen wir wie folgt: Wir bilden das zugehörige lineare System aus Proposition 45 (p. 39), d.h.

$$z'(t) = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} z(t) =: Az(t), \quad (264)$$

wobei

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}. \quad (265)$$

Dies können wir lösen wie in Zusammenfassung 31 (p. 29). Um eine Lösung zu bestimmen, müssen wir nun die Matrix A diagonalisieren oder in Jordan-Normalform bringen. Hierzu berechnen wir zunächst die Eigenwerte.

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & 1 \\ -\omega^2 & -\lambda \end{pmatrix} = \lambda^2 + \omega^2 = (\lambda - i\omega)(\lambda + i\omega). \quad (266)$$

Damit sind die Eigenwerte gegeben durch $\lambda_{1,2} = \pm i\omega$. Nun müssen wir die Eigenvektoren finden. Dazu: Der Eigenraum zum Eigenwert $i\omega$ ist

$$\ker(A - i\omega I) = \ker \begin{pmatrix} -i\omega & 1 \\ -\omega^2 & -i\omega \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 1 \\ i\omega \end{pmatrix} \right\} \quad (267)$$

Analog berechnen wir den Eigenraum zum Eigenwert $-i\omega$

$$\ker(A + i\omega I) = \ker \begin{pmatrix} i\omega & 1 \\ -\omega^2 & i\omega \end{pmatrix} = \text{span} \left\{ \begin{pmatrix} 1 \\ -i\omega \end{pmatrix} \right\} \quad (268)$$

Damit gilt für $S := \begin{pmatrix} 1 & 1 \\ i\omega & -i\omega \end{pmatrix}$

$$S^{-1}AS = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix}. \quad (269)$$

Definieren wir nun

$$E(t) := \begin{pmatrix} e^{i\omega t} & 0 \\ 0 & e^{-i\omega t} \end{pmatrix} \quad (t \in \mathbb{R}) \quad (270)$$

so gilt nach Zusammenfassung 31 (p. 29), dass

$$M(t) = SE(t) = \begin{pmatrix} 1 & 1 \\ i\omega & -i\omega \end{pmatrix} E(t) = \begin{pmatrix} 1 & 1 \\ i\omega & -i\omega \end{pmatrix} \begin{pmatrix} e^{i\omega t} & 0 \\ 0 & e^{-i\omega t} \end{pmatrix} \quad (271)$$

die Fundamentalmatrix ist. Wir führen die Matrixmultiplikation durch und erhalten

$$M(t) = \begin{pmatrix} e^{i\omega t} & e^{-i\omega t} \\ i\omega e^{i\omega t} & -i\omega e^{-i\omega t} \end{pmatrix}. \quad (272)$$

Die allgemeine Lösung des Hilfsproblems ist also

$$\tilde{\mathbb{L}} = \left\{ M(t)c : c \in \mathbb{C}^2 \right\} = \left\{ c_1 \begin{pmatrix} e^{i\omega t} \\ i\omega e^{i\omega t} \end{pmatrix} + c_2 \begin{pmatrix} e^{-i\omega t} \\ i\omega e^{-i\omega t} \end{pmatrix} : c_1, c_2 \in \mathbb{C} \right\}. \quad (273)$$

Wie in Proposition 45 (p. 39) gezeigt, ist dann die allgemeine Lösung für (262) gegeben durch

$$\mathbb{L} = \Phi^{-1}(\tilde{\mathbb{L}}) \hat{=} \text{Menge aller ersten Komponenten von } \tilde{\mathbb{L}} \quad (274)$$

$$= \{c_1 e^{i\omega t} + c_2 e^{-i\omega t} : c_1, c_2 \in \mathbb{C}\}. \quad (275)$$

SCHRITT 2: REALISIERUNG DER ANFANGSWERTE. Es sei nun

$$x(t) = c_1 e^{i\omega t} + c_2 e^{-i\omega t} \quad (t \in \mathbb{R}) \quad (276)$$

ein Element aus \mathbb{L} . Wir möchten $c_1, c_2 \in \mathbb{C}$ so bestimmen, dass die Anfangswerte angenommen werden. Dazu setzen wir

$$x_0 \stackrel{!}{=} x(0) \stackrel{(276)}{=} c_1 e^{i\omega t} + c_2 e^{-i\omega t} \Big|_{t=0} = c_1 + c_2 \quad (277)$$

$$v_0 \stackrel{!}{=} x'(0) \stackrel{(276)}{=} i\omega c_1 e^{i\omega t} - i\omega c_2 e^{-i\omega t} \Big|_{t=0} = i\omega c_1 - i\omega c_2. \quad (278)$$

Die Gleichungen (277) – (278) definieren ein lineares Gleichungssystem

$$\begin{cases} c_1 + c_2 = x_0 \\ c_1 - c_2 = \frac{v_0}{i\omega} \end{cases} \quad (279)$$

was wir nach c_1, c_2 auflösen können. Wir erhalten

$$c_1 = \frac{1}{2} \left(x_0 + \frac{v_0}{i\omega} \right) \quad (280)$$

$$c_2 = \frac{1}{2} \left(x_0 - \frac{v_0}{i\omega} \right). \quad (281)$$

Damit gilt

$$x(t) = \frac{1}{2} \left(x_0 + \frac{v_0}{i\omega} \right) e^{i\omega t} + \frac{1}{2} \left(x_0 - \frac{v_0}{i\omega} \right) e^{-i\omega t}. \quad (282)$$

Dieser Lösung sieht man so überhaupt nicht an, dass sie reellwertig ist. Und das, obwohl Sie nach Proposition 40 (p. 34) auf jeden Fall reellwertig ist! Daher sollten wir sie so vereinfachen, dass sie als reellwertige Lösung erkennbar ist. Mithilfe der Formeln $\cos(z) = \frac{e^{iz} + e^{-iz}}{2}$ und $\sin(z) = \frac{e^{iz} - e^{-iz}}{2i}$ erhalten wir

$$x(t) = x_0 \left(\frac{e^{i\omega t} + e^{-i\omega t}}{2} \right) + \frac{v_0}{\omega} \left(\frac{e^{i\omega t} - e^{-i\omega t}}{2i} \right) \quad (283)$$

$$= x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t). \quad (284)$$

Wir haben die Bewegungsgleichung gefunden.

Wir haben nach (282) behauptet, dass aus Proposition 40 (p. 34) die Reellwertigkeit von $x(t)$ folgt. Dies ist eigentlich nicht zu 100 Prozent offensichtlich, weil diese Proposition 40 ja nur für lineare Systeme erster Ordnung gilt. Hier müssten wir eigentlich zunächst auf Korollar 46 (p. 42) referenzieren, welches das gegebene AWP auf ein (reellwertiges!) AWP mit einem linearen System erster Ordnung zurückführt. Dieses hat dann laut Proposition 40 eine reellwertige Lösung.

Die Trajektorie in (284) nennt man übrigens *harmonische Schwingung* mit Anfangsort x_0 und Anfangsauslenkung v_0 . Ein physikalisches System, was der Bewegungsgleichung (262) genügt, heißt im Allgemeinen auch *harmonischer Oszillator*. Der Parameter ω wird *Eigenfrequenz* genannt.

Übung 52. MAXIMALAMPLITUDE DES FEDERPENDELS.

Es sei $x \in C^2(\mathbb{R}; \mathbb{R})$ wie in Anwendung 51 (p. 49) die Auslenkung des Federpendels, (d.h. die Lösung von

$$\begin{cases} x''(t) + \omega^2 x(t) = 0 \\ x(0) = x_0 \\ x'(0) = v_0, \end{cases} \quad (285)$$

für $x_0 \in \mathbb{R}$ die Anfangsauslenkung und $v_0 \in \mathbb{R}$ die Anfangsgeschwindigkeit). Bestimmen

Sie die *maximale Auslenkung* des Federpendels, d.h. $\max_{t \in \mathbb{R}} x(t)$, in Abhängigkeit von x_0, v_0 .

HINWEIS: Entnehmen Sie die explizite Formel für x aus (284) und bestimmen Sie das Maximum.

SINN DER ÜBUNG: Dies ist ein Beispiel für eine physikalisch relevante Größe, die man aus der Bewegungsgleichung herauslesen kann. Ohne die explizite Lösung der Bewegungsgleichungen hätte man diese Größen nicht bestimmen können.

Das Modell des harmonischen Oszillators kennt noch einige Verallgemeinerungen. Mit den Methoden aus den vorigen Kapiteln wären Sie als Lesende befähigt, die Bewegungsgleichungen in allen Situationen zu bestimmen und sämtliche korrespondierenden Anfangswertprobleme zu lösen. Wir wollen das an dieser Stelle aber nicht vertiefen und weisen nur auf zwei wichtige Verallgemeinerungen hin

1. **GEDÄMPFTER HARMONISCHER OSZILLATOR.** Zusätzlich zu der Rückstellkraft der Feder $F_1(t) = -kx(t)$ können wir auch noch eine *Reibungskraft der Luft* berücksichtigen. Diese ist gegeben durch $F_2(t) = -\rho x'(t)$ für ein $\rho > 0$, welches man *Reibungskoeffizient* nennt. Die Gesamtkraft berechnet sich dann durch die Summe beider Kräfte, also $F(t) = F_1(t) + F_2(t)$. Die Bewegungsgleichung wird dann

$$x''(t) + 2\gamma x'(t) + \omega^2 x(t) = 0, \quad (286)$$

wobei $\gamma = \frac{\rho}{2m} > 0$ und $\omega = \sqrt{\frac{k}{m}} > 0$. Lösungen von (286) kann man wie im vorigen Beispiel, d.h. durch Rückführung auf ein System der Form $z'(t) = Az(t)$, bestimmen. Ein Sonderfall tritt hier auf, wenn $\gamma = \omega$ ist. Dann ist die Matrix A nämlich nicht diagonalisierbar und man muss mit der Jordan-Normalform arbeiten. In der Physik ist dieser Fall als *Kriechfall* bekannt. Die Reibung ist in dem Fall so stark, dass gar keine Schwingung zustande kommt und das Pendel nur in seine Ausgangslage “zurückkriecht”.

2. **GETRIEBENER HARMONISCHER OSZILLATOR.** Man könnte das Pendel auch mit einer *externen Kraft* anschubsen. Zum Beispiel könnte man die Schwingung mit einer Kraft $F_2(t) = B \sin(\alpha t)$ ($B \in \mathbb{R}$) unterstützen, bzw dagegen arbeiten. Die Gesamtkraft ist dann $F(t) = -kx(t) + B \sin(\alpha t)$ und die Bewegungsgleichung ist

$$x''(t) + \omega^2 x(t) = d \sin(\alpha t), \quad (287)$$

wobei $d = \frac{B}{m} \in \mathbb{R}$. Das zugrundelegende DGL-System dieser Bewegungsgleichung ist Gegenstand von Übung 39 (p. 34). Bei dieser Übung fällt eine Sache auf: Falls $\alpha \neq \pm\omega$ ist die Lösung beschränkt. Falls $\alpha = \pm\omega$, so ist die Lösung unbeschränkt. Und das, obwohl die externe Kraft beschränkt ist! Man spricht vom *Resonanzfall*. Die Quintessenz: Treibt man einen Oszillator auch nur ganz leicht mit seiner Eigenfrequenz an, so entstehen Schwingungen mit riesigen Auslenkungen. Dieses Phänomen wurde im Jahre 1940 der “Tacoma Narrows Bridge” im US-Bundesstaat Washington zum Verhängnis. Da leichte(!) Windböen in der Eigenfrequenz der Brücke wehten, stürzte sie auf dramatische Weise ein (siehe YOUTUBE VIDEO ‘TACOMA NARROWS BRIDGE COLLAPSE’). Seitdem gibt es strenge Regularien beim Brückenbau, die das Treffen der Eigenfrequenz verhindern. Auch beim Bau von Hochhäusern spielt das Phänomen eine Rolle.

1.3 Integrale der Bewegung und weitere Tricks

Es gibt einige weitere Tricks, um Differentialgleichungen zu lösen. Manchmal ist es zum Beispiel möglich, die DGL durch geschickte Substitutionen zu vereinfachen. Einen Teil davon haben wir bereits bei den linearen DGLs und auch in Übung 22 (p. 21) gesehen. Ein weiterer Substitutionstrick wird in der folgenden Übung zusammengefasst.

*-Übung 53. EULER-HOMOGENE GLEICHUNGEN.

Es sei eine *Euler-homogene Differentialgleichung*, d.h. eine DGL der Form

$$y'(t) = g\left(\frac{y(t)}{t}\right) \quad (t \in I) \quad (288)$$

gegeben, wobei $I \subset (0, \infty)$ und $g \in C^0(\mathbb{R}; \mathbb{R})$. Finden Sie eine geeignete Substitution, die die DGL vereinfacht. Wenden Sie diese an, um die allgemeine Lösung von

$$y'(t) = \frac{1}{t} \left(\sqrt{y^2(t) + t^2} + y(t) \right) \quad (t \in (0, \infty)) \quad (289)$$

zu finden.

SINN DER ÜBUNG: Dies ist nur ein kleiner Einblick in die Mathematik des 18. Jahrhunderts, in der man versucht hat, alle DGLs aus der Anwendung explizit zu lösen. Heute ist dank Computern der Fokus der Forschung anders. Nichtsdestotrotz kann es nicht schaden, wenn man DGLs durch geschickte Substitutionen vereinfachen kann.

Es gibt aber noch einen anderen hilfreichen Trick zum Berechnen von Lösungen: Das Finden von *Erhaltungsgrößen*. Ein alternativer Name für eine Erhaltungsgröße ist *Integral der Bewegung*.

Definition 54. INTEGRAL DER BEWEGUNG. Es sei für $f \in C^0(I \times \mathbb{R}^n, \mathbb{R}^n)$ die DGL

$$y'(t) = f(t, y(t)) \quad (t \in I) \quad (290)$$

gegeben. Ein Ausdruck $g \in C^1(I \times \mathbb{R}^n; \mathbb{R})$ heißt *Integral der Bewegung*, falls für jede Lösung $y \in C^1(I; \mathbb{R}^n)$ gilt

$$\frac{d}{dt}g(t, y(t)) = 0 \quad (t \in I). \quad (291)$$

Das Schöne an einem Integral der Bewegung ist, dass der Ausdruck $t \mapsto g(t, y(t))$ zeitlich konstant ist. Die Formel

$$g(t, y(t)) = c \quad \forall t \in I \quad (292)$$

(für ein $c \in \mathbb{R}$) erlaubt oft weitreichende Rückschlüsse auf Eigenschaften der Lösung y — vor allem wenn man c mithilfe von gegebenen Anfangswerten explizit bestimmen kann! Oft charakterisieren wir ein Integral der Bewegung daher durch die Notation

$$g(t, y(t)) \equiv \text{const.} \quad \forall t \in I \quad (293)$$

Das Konzept des Integrals der Bewegung ist uns nicht neu: Die meisten Lösungen im bisherigen Verlauf der Vorlesung haben wir konstruiert, indem wir ein Integral der Bewegung gefunden haben.

Beispiel 55. Für die DGL $y'(t) = \alpha y(t)$, $(t \in \mathbb{R})$ ist $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$g(t, z) := e^{-\alpha t} z \quad (t \in \mathbb{R}, z \in \mathbb{R}) \quad (294)$$

ein Integral der Bewegung. In der Tat

$$\frac{d}{dt} g(t, y(t)) = \frac{d}{dt} e^{-\alpha t} y(t) = -\alpha e^{-\alpha t} y(t) + e^{-\alpha t} y'(t) = 0, \quad (295)$$

wobei wir im letzten Schritt nochmal $y'(t) = \alpha y(t)$ eingesetzt haben.

Einige Typen von Differentialgleichungen können wir allein durch Kenntnis einer Erhaltungsgröße lösen, so wie die Klasse der *exakten Differentialgleichungen*.

Bevor wir diese einführen, benötigen wir das *Lemma von Poincare*

Lemma 56. LEMMA VON POINCARÉ. Es sei $\Omega \subset \mathbb{R}^2$ ein sternförmiges Gebiet und $F = (F_1, F_2)^T : \Omega \rightarrow \mathbb{R}^2$ eine stetig differenzierbare Funktion. Dann sind folgende Aussagen äquivalent

1. $\frac{\partial F_2}{\partial x_1} = \frac{\partial F_1}{\partial x_2}$ auf ganz Ω .
2. F besitzt eine Stammfunktion, d.h. es existiert $\phi \in C^1(\Omega; \mathbb{R})$ mit $F = \nabla \phi$.

Beweis. Siehe Analysis Grundlagen-Vorlesung (HA).

Beispiel 57. EXAKTE DIFFERENTIALGLEICHUNGEN.

Gegeben sei eine DGL der Form

$$p(t, y(t)) + q(t, y(t))y'(t) = 0 \quad \forall t \in I \quad (296)$$

wobei $p, q \in C^1(\mathbb{R}^2; \mathbb{R})$ zwei reellwertige stetig diffbare Funktionen sind. Wir bezeichnen die Parameter, von denen p und q abhängen mit $p = p(\tau, z)$ und $q = q(\tau, z)$. Wir treffen nun die folgende

$$\text{Annahme: } \frac{\partial p}{\partial z} = \frac{\partial q}{\partial \tau} \quad \text{auf } \mathbb{R}^2. \quad (297)$$

Ist diese Annahme erfüllt, so nennen wir eine DGL eine exakte Differentialgleichung. In diesem Fall gibt es (weil \mathbb{R}^2 sternförmig ist) nach dem *Lemma von Poincare* aus der Analysis eine *Stammfunktion* $\phi \in C^1(\mathbb{R}^2; \mathbb{R})$ mit

$$\nabla \phi(\tau, z) = \begin{pmatrix} p(\tau, z) \\ q(\tau, z) \end{pmatrix}. \quad (298)$$

BEHAUPTUNG: ϕ ist ein Integral der Bewegung, d.h. für jede Lösung gilt

$$\phi(t, y(t)) \equiv \text{const.} \quad (t \in \mathbb{R}). \quad (299)$$

BEWEIS: Nach der (mehrdimensionalen) Kettenregel gilt

$$\frac{d}{dt} \phi(t, y(t)) = D\phi(t, y(t)) \cdot \frac{d}{dt} \begin{pmatrix} t \\ y(t) \end{pmatrix} = \left\langle \nabla \phi(t, y(t)), \frac{d}{dt} \begin{pmatrix} t \\ y(t) \end{pmatrix} \right\rangle \quad (300)$$

$$\stackrel{(298)}{=} \left\langle \begin{pmatrix} p(t, y(t)) \\ q(t, y(t)) \end{pmatrix}, \begin{pmatrix} 1 \\ y'(t) \end{pmatrix} \right\rangle \quad (301)$$

$$= p(t, y(t)) + q(t, y(t))y'(t) = 0. \quad (302)$$

Wir haben ein Integral der Bewegung gefunden. In vielen Fällen lässt sich dieses Integral der Bewegung in (299) explizit nach $y(t)$ auflösen. Somit kann man eine Lösung explizit bestimmen. Ist die Annahme (305) nicht erfüllt, können wir leider keine Stammfunktion finden — denn die Annahme ist auch notwendig (siehe Analysis-Vorlesung, HA)! In diesem Fall erhalten wir also so kein Integral der Bewegung. Manchmal ist es aber möglich, die DGL (296) mit einer Funktion $\mu(t)$ durchzumultiplizieren, sodass sie exakt wird. Diesen Trick nennt man das Finden eines *integrierenden Faktors*.

Übung 58. LÖSEN EINER EXAKTEN DGL.

Gegeben sei die DGL

$$4t^3y(t) + (y(t)^2 + 1 + t^4)y'(t) = 0 \quad (t \in \mathbb{R}) \quad (303)$$

Finden Sie ein Integral der Bewegung und zeigen Sie, dass alle Lösungen auf $I = \mathbb{R}$ beschränkt sind. Können Sie die allgemeine Lösung explizit angeben?

HINWEIS. Für die explizite Lösung greifen Sie auf die Cardani'schen Formeln WIKIPEDIA "CARDANI'SCHE FORMELN" zurück.

SINN DER ÜBUNG: Man sieht: Es ist hier sehr kompliziert, die DGL explizit zu lösen. Ein Integral der Bewegung hinzuschreiben ist aber mit dem vorangegangenen Beispiel relativ einfach. Vielleicht kann man ja auch viele Eigenschaften der Lösung bereits aus dem Integral der Bewegung ablesen... In diesem Fall bräuchte man sie vielleicht gar nicht explizit zu kennen! Dies wollen wir im Laufe des nächsten Kapitels einüben. Dort studieren wir ein besonders anschauliches Integral der Bewegung — Die Energieerhaltung! Sie wird viele Rückschlüsse auf das Verhalten der DGL erlauben, ohne, dass man eine Lösung explizit ausrechnen muss.

*-Übung 59. DER INTEGRIERENDE FAKTOR.

Gegeben sei eine DGL der Form

$$p(t, y(t)) + q(t, y(t))y'(t) = 0 \quad \forall t \in I \quad (304)$$

wobei $p, q \in C^1(\mathbb{R}^2; \mathbb{R})$ mit $q(\tau, z) > 0$ für alle $(\tau, z) \in \mathbb{R}^2$. Zeigen Sie: Unter der folgenden

$$\text{Annahme: } \frac{1}{q(\tau, z)} \left(\frac{\partial p}{\partial z}(\tau, z) - \frac{\partial q}{\partial \tau}(\tau, z) \right) \text{ hängt nur von } \tau \text{ ab,} \quad (305)$$

gibt es ein Integral der Bewegung.

HINWEIS: Multiplizieren Sie (304) mit einem geeigneten Faktor $m(t)$, sodass die Situation in Beispiel 57 erreicht wird.

Der Begriff des Integrals der Bewegung lässt sich auch für eine DGL höherer Ordnung formulieren. Bei einer DGL der Form

$$y''(t) = f(t, y(t), y'(t)) \quad (306)$$

würde man ein Integral der Bewegung als eine Funktion $g \in C^1(I \times \mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ verstehen, für die gilt

$$\frac{d}{dt} g(t, y(t), y'(t)) = 0. \quad (307)$$

Es ist wichtig, dass das Integral der Bewegung hier auch von $y'(t)$ abhängen darf! Denn um (307) zu verifizieren, muss man stets die Gleichung (306) benutzen. Hängt der Ausdruck aber gar nicht von $y'(t)$ ab, so wird beim Ableiten in (307) auch niemals $y''(t)$ ins Spiel kommen. In dem Fall wird man also niemals in der Lage sein, (306) einzusetzen.

1.3.1 Energieerhaltung

Wir betrachten eine DGL der Form

$$y''(t) = F(y(t)) \quad (t \in I) \quad (308)$$

wobei $F \in C^0(\mathbb{R}^n; \mathbb{R}^n)$ eine Funktion ist. In der Tat ist dies einfach die Newton'sche Bewegungsgleichung (258), in dem Spezialfall, dass die Masse m auf $m = 1$ normiert ist und die Kraft F nur vom Ort (und nicht von der Geschwindigkeit) abhängt. Zusätzlich nehmen wir an, dass

$$F(x) = -\nabla U(x) \quad \forall x \in \mathbb{R}^n \quad (309)$$

wobei $U \in C^1(\mathbb{R}^n; \mathbb{R})$ eine Funktion ist und ∇U den *Gradienten* von U bezeichnet. Für $n = 1$ gibt es ein solches U im Allgemeinen immer, da F stets eine Stammfunktion besitzt. Im Fall $n = 2$ geht es nicht immer: Notwendig ist hierfür z.B. (wenn wir zusätzlich $F \in C^1$ annehmen), dass F die sogenannten *Integrabilitätsbedingungen* (siehe Analysis-Vorlesung) erfüllt.

Wir nennen U im Folgenden das *Potential*.

Um ein Integral der Bewegung zu erhalten, multiplizieren wir (308) im Skalarprodukt mit $y'(t)$ und erhalten

$$\langle y''(t), y'(t) \rangle = \langle F(y(t)), y'(t) \rangle = -\langle \nabla U(y(t)), y'(t) \rangle \quad (310)$$

Wir können nun auf beiden Seiten Ableitungen erkennen. In der Tat gilt nach der (mehrdimensionalen) Kettenregel

$$\frac{d}{dt} U(y(t)) = DU(y(t)) \cdot y'(t) = \langle \nabla U(y(t)), y'(t) \rangle \quad (311)$$

und mit der Produktregel

$$\frac{d}{dt} \frac{1}{2} |y'(t)|^2 = \frac{d}{dt} \frac{1}{2} \sum_{i=1}^n y'_i(t)^2 = \frac{1}{2} \sum_{i=1}^n 2y'_i(t)y''_i(t) = \sum_{i=1}^n y'_i(t)y''_i(t) = \langle y''(t), y'(t) \rangle. \quad (312)$$

Wir erhalten also aus (310)

$$\frac{d}{dt} \frac{1}{2} |y'(t)|^2 = -\frac{d}{dt} U(y(t)), \quad (313)$$

oder umgeformt

$$\frac{d}{dt} \left(\frac{1}{2} |y'(t)|^2 + U(y(t)) \right) = 0. \quad (314)$$

Satz 60. ENERGIEERHALTUNG. Es sei $y \in C^1(I; \mathbb{R}^n)$ eine Lösung von (308), d.h.

$$y''(t) = -\nabla U(y(t)) \quad (t \in I) \quad (315)$$

Dann gilt

$$\frac{1}{2} |y'(t)|^2 + U(y(t)) \equiv \text{const.} \quad (316)$$

Den ersten Summanden dieses Integrals der Bewegung nennen wir *kinetische Energie* und den zweiten Summanden nennen wir *potentielle Energie*. In diesem Sinne nennen wir das Integral der Bewegung (394) die *Gesamtenergie*.

Beweis. Hier machen wir eigentlich nur nochmal dieselbe Rechnung wie am Anfang dieses Abschnittes. Wir berechnen für $t \in I$

$$\frac{d}{dt} \frac{1}{2} |y'(t)|^2 = \frac{d}{dt} \frac{1}{2} \sum_{i=1}^n y'_i(t)^2 = \sum_{i=1}^n y'_i(t)y''_i(t) = \langle y''(t), y'(t) \rangle \quad (317)$$

und

$$\frac{d}{dt} U(y(t)) = D U(y(t)) \cdot y'(t) = \langle \nabla U(y(t)), y'(t) \rangle. \quad (318)$$

Daher erhalten wir mit der Summenregel des Differenzierens und der Linearität des Skalarproduktes

$$\frac{d}{dt} \left(\frac{1}{2} |y'(t)|^2 + U(y(t)) \right) = \langle y''(t), y'(t) \rangle + \langle \nabla U(y(t)), y'(t) \rangle \quad (319)$$

$$= \langle y''(t) + \nabla U(y(t)), y'(t) \rangle = \langle 0, y'(t) \rangle = 0. \quad (320)$$

Im Spezialfall $n = 1$ kann in den meisten Fällen durch (394) eine Lösung explizit bestimmt werden. In der Tat: Falls

$$\frac{1}{2} y'(t)^2 + U(y(t)) = C, \quad (321)$$

für ein festes $C \in \mathbb{R}$, so ist

$$y'(t) = \pm \sqrt{2(C - U(y(t)))}. \quad (322)$$

Die DGL

$$y'(t) = \sqrt{2(C - U(y(t)))} \quad (323)$$

lässt sich vortrefflich mithilfe der Separation der Variablen lösen. Ebenso lässt sich

$$y'(t) = -\sqrt{2(C - U(y(t)))} \quad (324)$$

für ein festes $C \in \mathbb{R}$ gut lösen. Die Konstante C taucht nun stets als freier Parameter bei der Lösung auf und kann durch die Anfangswerte bestimmt werden.

Ein Problem bleibt: wenn auf der rechten Seite von (322) das Nullniveau erreicht wird, so kann man vom Fall ‘+’ in den Fall ‘-’ wechseln. Dieses ‘wechseln’ kann theoretisch unendlich oft passieren. Wie oft es der Vorzeichenwechsel tatsächlich vorkommt, muss im Einzelfall untersucht werden.

Allgemein ist es ratsam, (322) nicht einfach auswendig zu lernen, sondern in jedem konkreten Fall neu herzuleiten. Der Trick ist hier ja immer derselbe: Man multipliziert die DGL $y''(t) = f(y(t))$ auf beiden Seiten mit $y'(t)$ und bildet danach die Stammfunktion auf beiden Seiten. Nochmal formal: ist U eine Stammfunktion von $-f \in C^0(\mathbb{R}; \mathbb{R})$, so gilt für $y \in C^2(I; \mathbb{R})$

$$y''(t) = f(y(t)) \quad \forall t \in I \quad (325)$$

$$\Rightarrow y''(t)y'(t) = f(y(t))y'(t) \quad \forall t \in I \quad (326)$$

$$\Leftrightarrow \frac{d}{dt} \frac{1}{2}y'(t)^2 = \frac{d}{dt}(-U)(y(t)) \quad \forall t \in I \quad (327)$$

$$\Leftrightarrow \frac{1}{2}y'(t)^2 + U(y(t)) = c. \quad \forall t \in I \quad (328)$$

$$\Leftrightarrow y'(t) = \pm \sqrt{2(C - U(y(t)))} \quad \forall t \in I \quad (329)$$

Diese Technik des *Multiplizierens und Integrierens* ist weitläufig in der gesamten Mathematik verbreitet, um Integrale der Bewegung zu finden. Man spricht hier auch von der Idee des *integrierenden Faktors*.

Dass diese Technik zum Ziel führt, geht in letzter Instanz auf eine Beobachtung der Mathematikerin Emmy Noether (1882-1935) zurück ([LINK ZUR ORIGINALVERÖFFENTLICHUNG 1918](#))

Zum Abschluss ein Wort zur Eindeutigkeit bei Anfangswertproblemen: Das Multiplizieren mit $y'(t)$ ist (im Allgemeinen) keine Äquivalenzumformung (siehe (326))! Daher kann (321) auch mehr Lösungen haben als die Differentialgleichung selbst. Dies tritt insbesondere auf, wenn $y'(t_0) = 0$ für ein $t_0 \in I$, d.h. wenn man in (322) mehrere Möglichkeiten für die Vorzeichenwahl hat. In vielen Fällen kann man das Vorzeichen von $y'(t)$ aber bereits aus der Gleichung ablesen, siehe Übung 61 (p. 58).

Übung 61. EXPLIZITE LÖSUNGEN DURCH ENERGIEERHALTUNG - I.

Bestimmen Sie eine Lösung des Anfangswertproblems

$$\begin{cases} y''(t) = e^{2y(t)} & (t \in \mathbb{R}) \\ y(0) = 0 \\ y'(0) = y_1 \end{cases} \quad (330)$$

für ein beliebiges $y_1 \geq 1$. Ist die Lösung eindeutig?

HINWEIS: Wenden Sie zuerst (durch Multiplizieren der Gleichung mit $y'(t)$ und Integrieren) den “Energieerhaltungstrick” (s. (325) ff.) an und zeigen Sie mit dessen Hilfe, dass $y'(t) > 0$ für alle $t \in \mathbb{R}$. Dies erspart Ihnen beim Wurzelziehen dann die Diskussion des Vorzeichens! Außerdem kann so auch die Eindeutigkeitsfrage direkt beantwortet werden!
SINN DER ÜBUNG: Dies ist eines der einfachsten Beispiele, die man in dem Kontext rechnen kann und soll vorbereiten auf das *Fadenpendel*, welches wir im nächsten Abschnitt durchrechnen werden.

Ein weiterer Trick hierzu: Falls $y''(t) > 0$ für alle $t \in I$ (was man ja oft direkt aus der Gleichung ablesen kann), so ist $y'(t)$ monoton und daher kann $y'(t)$ maximal einmal das Vorzeichen von ‘–’ nach ‘+’ wechseln. Dies hilft zum Beispiel beim Diskutieren der Voraussetzung $y_1 \geq 1$ in der vorigen Übung.

Übung 62. EXPLIZITE LÖSUNGEN DURCH ENERGIEERHALTUNG- II.

Es sei das Anfangswertproblem

$$\begin{cases} y''(t) = e^{2y(t)} & (t \in I) \\ y(0) = 0 \\ y'(0) = y_1 \end{cases} \quad (331)$$

für ein beliebiges $y_1 \in (0, 1)$ gegeben. Zeigen Sie, dass es keine globale Lösung auf $I = \mathbb{R}$ gibt.

HINWEIS. Angenommen wir hätten eine globale auf $I = \mathbb{R}$ definierte Lösung. Zeigen Sie, dass $y'(t) > 0$ für alle $t \in (0, \infty)$ gelten muss. Folgern Sie wie in der vorigen Übung

$$y'(t)^2 = e^{2y(t)} - 1 + y_1^2 \quad \forall t \in (0, \infty) \quad (332)$$

und führen Sie dies durch Separation der Variablen zu einem Widerspruch.

1.3.2 Anwendung: Das mathematische Pendel

Mithilfe des Energieerhaltungstricks werden wir die Gleichung

$$\varphi''(t) = -\omega^2 \sin \varphi(t) \quad (333)$$

untersuchen. Diese kommt zustande, wenn wir den Auslenkungswinkel eines schwingenden *Fadenpendels* betrachten. In der Schule haben wir hier für kleine Auslenkungswinkel φ angenommen, dass “ $\sin \varphi \approx \varphi$ ”, die sogenannte *Kleinwinkelnäherung*. Unter dieser Annahme ist

die Gleichung dieselbe wie für den harmonischen Oszillatoren in Sektion 1.2.6. Auf die Annahme, dass der Auslenkungswinkel klein ist, wollen wir aber jetzt verzichten. Denn wir wollen z.B. auch verstehen wann das Pendel einen *Überschlag* macht, d.h. einmal um den Ursprung herum schwingt. Dieses Phänomen können wir mit dem Energieerhaltungstrick begreifen, ohne, dass wir die Lösung der Gleichung explizit bestimmen müssen.

Im Folgenden wiederholen wir nochmal die Gleichung des Fadenpendels und wie sie hergeleitet wird.

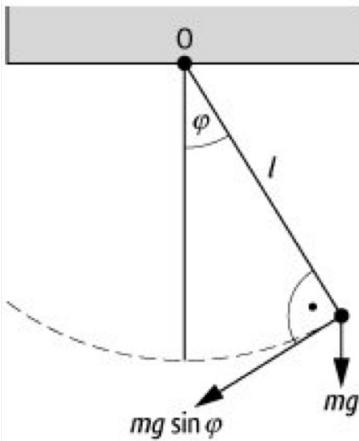


Abbildung 2: Das mathematische Pendel

Physikalische Herleitung 63. DAS FADENPENDEL.

Wir betrachten ein schwingendes Fadenpendel. An einem Faden der Länge l hängt ein Gewicht der Masse m und schwingt (siehe Abbildung 2). Auf die Masse wirkt nur die *Gewichtskraft*, welche die Masse immer wieder nach unten zieht. Wir möchten die Bewegung beschreiben. Wir sehen, dass wir den Ort nur über den *Auslenkungswinkel* φ beschreiben können, d.h. es gilt (wenn wir den Nullpunkt wie in Abbildung 2) legen.

$$\vec{x}(t) = l \begin{pmatrix} \sin \varphi(t) \\ -\cos \varphi(t) \end{pmatrix}. \quad (334)$$

Mithilfe dieser Formel können wir auch Geschwindigkeit und Beschleunigung in Abhängigkeit von φ berechnen, nämlich

$$\vec{x}'(t) = l\varphi'(t) \begin{pmatrix} \cos \varphi(t) \\ \sin \varphi(t) \end{pmatrix} \quad (335)$$

und

$$\vec{x}''(t) = l\varphi''(t) \begin{pmatrix} \cos \varphi(t) \\ \sin \varphi(t) \end{pmatrix} + l\varphi'(t) \begin{pmatrix} -\sin \varphi(t) \\ \cos \varphi(t) \end{pmatrix}. \quad (336)$$

Das Masseteilchen bewegt sich nun in der Gewichtskraft. Diese zeigt nach an jedem Punkt

proportional zur Erdbeschleunigung unten, d.h.

$$\vec{F}(x_1, x_2, x_3) = mg \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \quad (337)$$

Mit dem Newton'schen Bewegungsgesetz (258) erhalten wir

$$m\vec{x}''(t) = mg \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad (338)$$

d.h.

$$l\varphi''(t) \begin{pmatrix} \cos \varphi(t) \\ \sin \varphi(t) \end{pmatrix} + l\varphi'(t) \begin{pmatrix} -\sin \varphi(t) \\ \cos \varphi(t) \end{pmatrix} = g \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \quad (339)$$

Wir wollen eine Differentialgleichung für φ daraus herleiten. Hierzu verfahren wir wie folgt: Wir nehmen auf beiden Seiten von (339) das Skalarprodukt mit

$$v(t) := \begin{pmatrix} \cos \varphi(t) \\ \sin \varphi(t) \end{pmatrix}.$$

Für die linke Seite ergibt sich (HA)

$$\left\langle v(t), l\varphi''(t) \begin{pmatrix} \cos \varphi(t) \\ \sin \varphi(t) \end{pmatrix} + l\varphi'(t) \begin{pmatrix} -\sin \varphi(t) \\ \cos \varphi(t) \end{pmatrix} \right\rangle = l\varphi''(t). \quad (340)$$

(Hierbei haben wir die Linearität des Skalarproduktes und $\cos^2 \varphi + \sin^2 \varphi = 1$ benutzt). Die rechte Seite berechnet sich mit

$$\left\langle v(t), g \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\rangle = -g \sin \varphi(t). \quad (341)$$

Sind zwei Ausdrücke gleich, so müssen auch ihre Skalarprodukte mit $v(t)$ gleich sein. Daher folgt aus den vorangegangenen beiden Gleichungen und aus (339)

$$l\varphi''(t) = -g \sin \varphi(t), \quad (342)$$

oder auch

$$\varphi''(t) = -\frac{g}{l} \sin \varphi(t) = -\omega^2 \sin \varphi(t) \quad (343)$$

mit $\omega = \sqrt{\frac{g}{l}}$.

Wir untersuchen nun für ein $\alpha_0 > 0$ das Anfangswertproblem

$$\begin{cases} \varphi''(t) = -\omega^2 \sin \varphi(t) & (t \in \mathbb{R}) \\ \varphi(0) = 0, \\ \varphi'(0) = \alpha_0. \end{cases} \quad (344)$$

Das heißt, wir betrachten ein Pendel mit Pendeldaten $\omega = \sqrt{\frac{g}{l}}$, was aus der Ruhelage mit einer (Winkel-)geschwindigkeit $\alpha_0 > 0$ nach rechts schwingt. Wir fragen uns: Wie groß

müssen α_0 und ω sein, damit das Pendel einen Überschlag macht?

Anwendung 64. ENERGIEERHALTUNG ALS ÜBERSCHLAGSDETEKTOR. Es sei $\varphi \in C^2(\mathbb{R}; \mathbb{R})$ eine Lösung von (344) für $\alpha_0 > 0$. Wir stellen uns hier die folgende Frage

$$\text{Schwingt das Pendel oder überschlägt es sich?} \quad (345)$$

Die Antwort auf die Frage wird sicherlich mit der Anfangsgeschwindigkeit α_0 zusammenhängen: Ist sie klein, so sollte das Pendel schwingen. Ist sie groß, so sollte es sich überschlagen! Wir multiplizieren die Gleichung mit $\varphi'(t)$ und finden

$$\varphi''(t)\varphi'(t) = -\omega^2 \sin \varphi(t)\varphi'(t) \quad \forall t \in \mathbb{R}. \quad (346)$$

Mit der Kettenregel rückwärts (angewendet wie in der vorigen Sektion) heißt das

$$\frac{d}{dt} \frac{1}{2}\varphi'(t)^2 = \frac{d}{dt} \omega^2 \cos \varphi(t) \quad (347)$$

Integriert bedeutet dies, dass es ein $C \in \mathbb{R}$ gibt mit

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \cos \varphi(t) + C. \quad (348)$$

Setzen wir $t = 0$ und die Anfangswerte aus (344) ein, so erhalten wir (mit $\cos 0 = 1$)

$$\frac{1}{2}\alpha_0^2 = \omega^2 + C, \quad (349)$$

d.h. $C = \frac{1}{2}\alpha_0^2 - \omega^2$. Insgesamt gilt also

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \cos \varphi(t) + \frac{1}{2}\alpha_0^2 - \omega^2. \quad (350)$$

Rein von der Anschauung her könnte man sich zwei plausible Verhaltensweisen des Pendels vorstellen:

FALL 1. (Das Pendel überschlägt sich) Der Auslenkungswinkel wächst mit einer strikt positiven Winkelgeschwindigkeit $\varphi'(t) > \delta > 0$ an (d.h. das Pendel macht irgendwann einen Überschlag).

FALL 2. (Das Pendel schwingt) Der Auslenkungswinkel erreicht irgendwann eine maximale Auslenkung $\varphi_{max} \in (0, \pi)$ und kehrt dann um, d.h. schwingt zurück.

Es könnte natürlich aber auch sein, dass keiner der beiden Fälle zutrifft, so z.B. könnte der Auslenkungswinkel monoton anwachsen und irgendwo stehen bleiben.

Tun wir mal zuerst so, als würde Fall 2 eintreten: Falls das Pendel bei einer Auslenkung $\varphi_{max} \in (0, \pi)$ umkehrt, so muss es einen Umkehrzeitpunkt $t^* > 0$ geben mit $\varphi'(t^*) = 0$ und $\varphi(t^*) = \varphi_{max} \in (0, \pi)$. Falls es ein solches t^* gibt, so ergibt Einsetzen von $t = t^*$ in (350)

$$0 = \omega^2 \cos \varphi(t^*) + \frac{1}{2}\alpha_0^2 - \omega^2, \quad (351)$$

das heißt

$$\cos \varphi_{max} = \frac{1}{\omega^2} \left(\omega^2 - \frac{1}{2} \alpha_0^2 \right) = 1 - \frac{1}{2} \frac{\alpha_0^2}{\omega^2}. \quad (352)$$

Wann gibt es ein solches $\varphi_{max} \in (0, \pi)$? Nun ja, da $\cos : (0, \pi) \rightarrow (1, -1)$ bijektiv ist, ist das genau dann der Fall wenn

$$1 - \frac{1}{2} \frac{\alpha_0^2}{\omega^2} > -1, \quad (353)$$

mit anderen Worten: Einen solchen Punkt t^* mit Nullableitung kann es nur dann geben wenn

$$\alpha_0^2 < 4\omega^2. \quad (354)$$

VERMUTUNG: Ist $\alpha_0^2 < 4\omega^2$, so erwarten wir, dass das Pendel schwingt. Das müssen wir aber noch formal beweisen, denn wir haben einfach so getan, als hätten wir schon einen Punkt t^* wo $\varphi'(t^*) = 0$ ist. Und wir haben auch einfach so angenommen, dass das Pendel an diesem Punkt auch wirklich umkehrt.

Das muss nicht zwingend so sein: Die Funktion $f(x) = x^3$ erfüllt bei $x = 0$ auch, dass $f'(x) = 0$ ist. Jedoch kehrt sich f bei $x = 0$ nicht um, sondern wächst weiter an! Um den Sachverhalt zu klären benötigen wir ein *Konvexitätsargument*, welches wir in der nächsten Übung genauer besprechen wollen.

Auch haben wir noch nicht diskutiert, dass das Pendel sich tatsächlich überschlägt, wenn $\alpha_0^2 > 4\omega^2$.

Ferner ist auch interessant, was im Grenzfall $\alpha_0^2 = 4\omega^2$ passiert. Schlägt das Pendel über, schwingt es zurück oder kommt es gar zum Stillstand?

Bevor wir den kommenden Satz besprechen können, benötigen wir die folgende Charakterisierung konvexer Funktionen.

Übung 65. GEOMETRISCHE BEDEUTUNG VON KONVEXITÄT.

Es sei $f \in C^2((a, b); \mathbb{R})$ konvex.

1. Zeigen Sie: Dann gilt für alle $x, x_0 \in (a, b)$, dass

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0). \quad (355)$$

2. Folgern Sie: Ist $f'(x_0) = 0$ für ein $x_0 \in (a, b)$, so ist x_0 ein globales Minimum von f auf (a, b) .
3. Formulieren Sie analoge Versionen von Teil 1 und 2 für konkave Funktionen.
4. Zeigen Sie: f ist genau dann konvex wenn für alle $x, x_0 \in (a, b)$ die Ungleichung (355) gilt.

SINN DER ÜBUNG: Wir werden diese Aussage im nächsten Beweis benötigen. Des Weiteren schadet es nie ein Paar Aussagen über konvexe Funktionen zu kennen. Die geometrische Interpretation der Aussage (355) ist, dass der Graph einer konvexen Funktion stets oberhalb von allen Tangenten an dem Graphen liegt.

Satz 66. DAS FADENPENDEL Es sei $\varphi \in C^2(\mathbb{R}; \mathbb{R})$ eine Lösung von (344) mit $\omega, \alpha_0 > 0$. Dann tritt genau einer der folgenden drei Fälle ein.

1. $\alpha_0^2 > 4\omega^2$. Dann gilt $\varphi(t) \geq t\sqrt{\alpha_0^2 - 4\omega^2}$ für alle $t \geq 0$. Insbesondere ist die Lösung unbeschränkt.
2. $\alpha_0^2 < 4\omega^2$. Dann gibt es $\varphi_{max} \in (0, \pi)$ so, dass $\varphi(t) \in [-\varphi_{max}, \varphi_{max}]$ für alle $t \in [0, \infty)$. Genauer ist φ_{max} gegeben durch

$$\varphi_{max} = \arccos\left(1 - \frac{1}{2}\frac{\alpha_0^2}{\omega^2}\right). \quad (356)$$

3. $\alpha_0^2 = 4\omega^2$. Dann gilt

$$\varphi(t) = 2 \arctan\left(\sinh \frac{\alpha_0}{2}t\right) \quad \forall t \in [0, \infty). \quad (357)$$

Insbesondere ist φ monoton und konvergiert gegen π für $t \rightarrow \infty$.

Beweis. Wie in Anwendung 64 (p. 62) leiten wir her

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \cos \varphi(t) + \frac{1}{2}\alpha_0^2 - \omega^2. \quad (358)$$

Nun zu Fall 1. Angenommen, dass $\alpha_0^2 > 4\omega^2$. Dann gilt

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \cos \varphi(t) + \frac{1}{2}\alpha_0^2 - \omega^2 \geq \frac{1}{2}\alpha_0^2 - 2\omega^2 \quad \forall t \in \mathbb{R}. \quad (359)$$

Wir folgern, dass

$$|\varphi'(t)| \geq \sqrt{\alpha_0^2 - 4\omega^2} \quad \forall t \geq 0. \quad (360)$$

Insbesondere kann φ' nicht mehr das Vorzeichen wechseln, da es kein $t_0 \in \mathbb{R}$ mit $\varphi'(t_0) = 0$ geben kann. (Würde φ' nämlich das Vorzeichen wechseln, so ergäbe sich hier ein Widerspruch zum Zwischenwertsatz). Da $\varphi'(0) = \alpha_0 > 0$ gilt also $\varphi'(t) > 0$ für alle $t \geq 0$. Mit (360) schließen wir, dass

$$\varphi'(t) \geq \sqrt{\alpha_0^2 - 4\omega^2} \quad \forall t \geq 0. \quad (361)$$

Mit dem Hauptsatz der Differential -und Integralrechnung erhalten wir

$$\varphi(t) = \varphi(t) - \varphi(0) = \int_0^t \varphi'(s) \, ds \geq \int_0^t \sqrt{\alpha_0^2 - 4\omega^2} \, ds = t\sqrt{\alpha_0^2 - 4\omega^2}. \quad (362)$$

Der erste Fall ist damit geklärt.

Nun zu Fall 2. Wir definieren

$$L := \arccos\left(1 - \frac{1}{2} \frac{\alpha_0^2}{\omega^2}\right) \in (0, \pi). \quad (363)$$

Es sei also $\alpha_0^2 < 4\omega^2$. Wir nehmen nun an, dass es ein $\bar{t} > 0$ gibt mit $|\varphi(\bar{t})| > L$. O.B.d.A können wir annehmen, dass $|\varphi(s)| < \pi$ für alle $s \in [0, \bar{t}]$, ansonsten verkleinern wir \bar{t} geeignet. (Hierfür ist wichtig, dass $L < \pi$, d.h. wenn $|\varphi|$ das Niveau π überschreitet, dann muss es nach dem Zwischenwertsatz das Niveau L bereits vorher überschritten haben). Wir unterscheiden die Fälle $\varphi(\bar{t}) > 0$ und $\varphi(\bar{t}) < 0$. Falls $\varphi(\bar{t}) > 0$, so können wir ein $\underline{t} \in [0, \bar{t}]$ wählen mit $0 < \varphi(s) < \pi$ für alle $s \in (\underline{t}, \bar{t})$ und $\varphi(\underline{t}) < L$. (Man wähle \underline{t} einfach kurz vor dem Zeitpunkt, wo L das erste Mal überschritten wird). Wir behaupten nun, dass φ auf (\underline{t}, \bar{t}) konkav ist. In der Tat:

$$\varphi''(t) = -\omega^2 \sin \varphi(t) < 0 \quad \forall t \in (\underline{t}, \bar{t}), \quad (364)$$

da $\sin(z) > 0$ für alle $z \in (0, \pi)$. Es sei nun $t^* \in (\underline{t}, \bar{t})$ mit $\varphi(t^*) = L$. Nun gilt nach (358)

$$\frac{1}{2} \varphi'(t^*)^2 = \omega^2 \cos(L) + \frac{1}{2} \alpha_0^2 - \omega^2 = \omega^2 \left(1 - \frac{1}{2} \frac{\alpha_0^2}{\omega^2}\right) + \frac{1}{2} \alpha_0^2 - \omega^2 = 0. \quad (365)$$

Insbesondere gilt also $\varphi'(t^*) = 0$. Nun ist also φ auf (\underline{t}, \bar{t}) konkav und φ' bei t^* eine Nullstelle hat. Nach Übung 65 (p. 63) hat φ dann aber bei t^* ein globales Maximum auf (\underline{t}, \bar{t}) , also

$$\varphi(t) \leq \varphi(t^*) = L \quad \forall t \in (\underline{t}, \bar{t}). \quad (366)$$

Hieraus würde aus Stetigkeitsgründen auch folgen, dass $\varphi(\bar{t}) \leq L$. Das ist aber ein Widerspruch, da nach unserer Annahme $\varphi(\bar{t}) > L$. Es ergibt sich also ein Widerspruch zur Grundannahme, dass das Niveau L überschritten wird. Im Fall, dass $\varphi(\bar{t}) < 0$ kann man analog vorgehen: Man wählt sich $\underline{t} \in [0, \bar{t}]$ so, dass $\varphi(\underline{t}) > -L$ und $-\pi < \varphi(s) < 0$ für alle $s \in (\underline{t}, \bar{t})$. Aus der Differentialgleichung lernt man dann wieder, dass

$$\varphi''(t) = -\omega^2 \sin \varphi(t) > 0 \quad \forall t \in (\underline{t}, \bar{t}), \quad (367)$$

d.h. φ ist konvex auf (\underline{t}, \bar{t}) . Nun gibt es auch wieder nun $t^* \in (\underline{t}, \bar{t})$ mit $\varphi(t^*) = -L$. Aus (358) erhält man wiederum

$$\frac{1}{2} \varphi'(t^*)^2 = \omega^2 \cos(-L) + \frac{1}{2} \alpha_0^2 - \omega^2 = 0, \quad (368)$$

da \cos ungerade ist. Wiederum würde aus der Konvexität (siehe Übung 65 (p. 63)) folgen, dass t^* eine globale Minimumsstelle von φ in $(-\underline{t}, \bar{t})$ ist. Ein Widerspruch, da dann das Niveau L wiederum nicht unterschritten werden kann. Wir haben also gezeigt

$$|\varphi(t)| \leq L \quad \forall t \geq 0 \quad (369)$$

und folgern die Behauptung.

Nun zu Fall 3. Hier, im Fall $\alpha_0^2 = 4\omega^2$, lösen wir (358) explizit. Nach (358) gilt

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \cos \varphi(t) + \frac{1}{2}\alpha_0^2 - \omega^2 = \omega^2(1 + \cos \varphi(t)) \quad (370)$$

Nun benutzen wir, dass $\cos z = 1 - 2\sin^2(\frac{z}{2})$ für alle $z \in \mathbb{R}$ und erhalten

$$\frac{1}{2}\varphi'(t)^2 = \omega^2 \left(2 - 2\sin^2 \frac{\varphi(t)}{2}\right) = 2\omega^2 \cos^2 \frac{\varphi(t)}{2}. \quad (371)$$

Wir kommen auf die DGL

$$\varphi'(t)^2 = 4\omega^2 \cos^2 \frac{\varphi(t)}{2} = \alpha_0^2 \cos^2 \frac{\varphi(t)}{2}. \quad (372)$$

Da $\varphi'(0) > 0$ und $\cos \frac{\varphi(0)}{2} = 1 > 0$ kann man die Wurzel in einer Umgebung von 0 mit ‘+’-Vorzeichen ziehen und erhält

$$\varphi'(t) = \alpha_0 \cos \frac{\varphi(t)}{2} \quad \forall t \in [0, M]. \quad (373)$$

Hierbei ist

$$M := \sup\{m \in [0, \infty) : \varphi'(t) > 0 \text{ und } \varphi(t) \in [0, \pi] \quad \forall t \in [0, m]\} > 0. \quad (374)$$

Wir lösen nun (373) auf $[0, M]$. Da $\cos \frac{z}{2} = \frac{1}{\sqrt{1+\tan^2 \frac{z}{2}}}$ (was direkt aus $\tan(w) = \frac{\sin(w)}{\cos(w)}$ folgt, HA) erhalten wir

$$\varphi'(t) = \alpha_0 \frac{1}{\sqrt{1 + \tan^2 \frac{\varphi(t)}{2}}}, \quad (375)$$

d.h.

$$\varphi'(t) \sqrt{1 + \tan^2 \frac{\varphi(t)}{2}} = \alpha_0. \quad (376)$$

Wir integrieren und erhalten

$$\int_0^t \sqrt{1 + \tan^2 \frac{\varphi(s)}{2}} \varphi'(s) \, ds = \alpha_0 t. \quad (377)$$

Wir substituieren $u = \varphi(s)$ und erhalten (mit der Tatsache, dass $\varphi(0) = 0$)

$$\int_0^{\varphi(t)} \sqrt{1 + \tan^2 \frac{u}{2}} \, du = \alpha_0 t. \quad (378)$$

Jetzt müssen wir nur noch das Integral auf der linken Seite auswerten. Dazu formen wir um

$$\int_0^{\varphi(t)} \sqrt{1 + \tan^2 \frac{u}{2}} du = \int_0^{\varphi(t)} \frac{1 + \tan^2 \frac{u}{2}}{\sqrt{1 + \tan^2 \frac{u}{2}}} du \quad (379)$$

Nun substituieren wir $v = \tan \frac{u}{2}$. Beachte, dass dann gilt $\frac{dv}{du} = \frac{1}{2} (1 + \tan^2 \frac{u}{2})$. Dadurch erhalten wir

$$\int_0^{\varphi(t)} \sqrt{1 + \tan^2 \frac{u}{2}} du = 2 \int_0^{\tan \frac{\varphi(t)}{2}} \frac{1}{\sqrt{1 + v^2}} dv = 2 \text{Arsinh} \left(\tan \frac{\varphi(t)}{2} \right), \quad (380)$$

wobei $\text{Arsinh}(y) := \log(y + \sqrt{1 + y^2})$ die Umkehrfunktion von \sinh ist (siehe Analysis-Vorlesung). Mit (378) erhalten wir

$$2 \text{Arsinh} \left(\tan \frac{\varphi(t)}{2} \right) = \alpha_0 t. \quad (381)$$

Wir lösen nun nach $\varphi(t)$ auf und erhalten

$$\varphi(t) = 2 \arctan \left(\sinh \frac{\alpha_0}{2} t \right) \quad \forall t \in [0, M]. \quad (382)$$

Wir zeigen nun $M = \infty$. In der Tat: Angenommen es wäre $M \in (0, \infty)$ so wäre (wegen der Definition von M in (374)) entweder $\varphi(M) \in \{0, \pi\}$ oder $\varphi'(M) = 0$. Aus (382) folgt allerdings

$$\varphi(M) = \lim_{t \rightarrow M} \varphi(t) = \lim_{t \rightarrow M} 2 \arctan \left(\sinh \frac{\alpha_0}{2} t \right) = 2 \arctan \left(\sinh \frac{\alpha_0}{2} M \right) \in (0, \pi) \quad (383)$$

und (wieder aus (382))

$$\varphi'(M) = \lim_{t \rightarrow M} \varphi'(t) = \lim_{t \rightarrow M} \frac{\alpha_0 \cosh \frac{\alpha_0}{2} t}{1 + \sinh^2 \frac{\alpha_0}{2} t} = \frac{\alpha_0 \cosh \frac{\alpha_0}{2} M}{1 + \sinh^2 \frac{\alpha_0}{2} M} > 0. \quad (384)$$

Wir folgern einen Widerspruch zur Annahme und schließen $M = \infty$. Wiederum aus (382) folgt

$$\varphi(t) = 2 \arctan \left(\sinh \frac{\alpha_0}{2} t \right) \quad \forall t \in [0, \infty). \quad (385)$$

Monotonie und Grenzverhalten für φ folgen sofort aus der Monotonie von \arctan und \sinh .

Das explizite Ausrechnen des dritten Falles war ziemlich mühsam. Später werden wir einen Weg sehen, wie man ohne die explizite Lösung zeigen kann, dass das Pendel nicht überschwingt. Eigentlich müssen wir aber auch erstmal überhaupt zeigen, dass die Gleichung $\varphi''(t) = -\omega^2 \varphi(t)$ überhaupt eine Lösung hat, die auf $[0, \infty)$ definiert ist...

2 Differentialgleichungen im Anwendungskontext

In diesem Abschnitt wollen wir über Differentialgleichungen in der Anwendung sprechen. Wir diskutieren anhand des Beispiels eines Pandemiemodells, wie Vorgänge mit Differentialgleichungen modelliert werden können und was für Vorhersagen man mit mathematischen Methoden aus dem Modell herauslesen kann.

Im Unterschied zum letzten Kapitel besprechen wir hier Modelle, bei denen wir keine Chance haben, eine Lösung explizit hinzuschreiben. Dennoch gibt es Informationen, die man durch das mathematische Studium des Modells erhalten kann.

Modelle sind stets eine Vereinfachung der Wirklichkeit. Das Modell, was wir hier besprechen ist in der Vergangenheit sehr häufig erweitert worden, weil seine Aussagekraft in gewissen Situationen begrenzt ist. In der Anwendung ist es daher auch besonders wichtig, die Modellannahmen zu diskutieren.

Ein weiteres Problem bei Modellen ist oft die Wahl der *Modellparameter*. Für das *Pandemiemodell*, was wir in diesem Kapitel besprechen werden, hängen die Modellparameter davon ab, wie ansteckend die Krankheit ist, wie viele Kontakte in der Bevölkerung stattfinden, wie lange eine Person infektiös ist, etc. Wir werden in diesem Skript auch einen Fokus auf die epidemiologische Bedeutung der Parameter legen — dieser Teil des Skriptes wird dementsprechend interdisziplinär gestaltet sein. Hierbei geht das Skript auch u.U. detaillierter vor als die grundständige mathematische Literatur.

Nichtdestotrotz ist die genaue Bestimmung der Parameter ein schwieriges Problem. Aber auch hier ist die Theorie der Differentialgleichungen von Nutzen: Je mehr theoretische Aussagen man zur Verfügung hat, desto besser kann man die Parameter an die verfügbare Datenlage anpassen. Auch das werden wir sehen.

2.1 Das SIR-Pandemiemodell

Wir begeben uns gedanklich in eine Pandemiesituation. Täglich infizieren sich Menschen mit — beispielsweise — einem Virus. Die infizierten Menschen können wiederum weitere Personen anstecken. In unserem Modell teilen wir die Bevölkerung in drei Gruppen auf.

Bezeichnung 67. MODELLIERTE BEVÖLKERUNGSGRUPPEN.

Die Bevölkerung wird in die Gruppen S, I und R aufgeteilt. Die Gruppe I bezeichnet die Gruppe der *Infizierten*. Die Gruppe R bezeichnet die Gruppe der *Genesenen oder Verstorbenen*. Die Gruppe S bezeichnet die Gruppe der *Anfälligen*, also diejenigen, die weder infiziert, noch genesen/verstorben sind.

Wir bezeichnen außerdem mit $I(t)$ den Bevölkerungsanteil der Infizierten zu einem Zeitpunkt t . Gleichsam bezeichnen wir mit $R(t)$ den Anteil der Genesenen oder Verstorbenen in einer Bevölkerung und mit $S(t)$ den Anteil der Anfälligen zum Zeitpunkt $t \geq 0$.

Diese Bezeichnungen sind in der Literatur üblich. Die Buchstaben sind aufgrund englischer Bezeichnungen gewählt ($I \hat{=} infectious$, $R \hat{=} recovered/removed$, $S \hat{=} susceptible$).

Wegen den drei verwendeten Buchstaben heißt das Modell *SIR-Modell*. Man beachte, dass wir stets *Bevölkerungsanteile* in Prozent der Gesamtbevölkerung modellieren. Das heißt

auch

$$S(t) + I(t) + R(t) = 1 \quad \forall t \in \mathbb{R}. \quad (386)$$

Dass wir die Genesenen und die Verstorbenen in einen Topf werfen können, liegt an der folgenden Annahme.

Modellannahme 68. EINMAL IMMUN, IMMER IMMUN.

Wir nehmen hier (zunächst) an, dass Personen, die einmal eine Infektion durchgemacht haben, nach einer Infektion komplett immun sind. Dadurch können sie aus dem Pandemiegeschehen entfernt werden. Wir können die Genesenen also genauso aus dem Geschehen entfernen wie die Verstorbenen.

Diese Annahme ist nicht in jeder Pandemie zutreffend. Später werden wir das Modell erweitern und eine *Reinfektion* ermöglichen.

Wir bezeichnen ab hier der Einfachheit halber die Gruppe R nur als Gruppe der *Genesenen*.

Als nächstes diskutieren wir, wie sich die Bevölkerungsanteile mit der Zeit verändern. Wir starten zunächst mit den Genesenen.

Modellierung 69. ÄNDERUNGSRATE VON $R(t)$. Wir fragen uns hier, wie viele Menschen in einer Zeiteinheit zwischen t und $t + h$ neu zu den Genesenen dazukommen. Wie schon in Anwendung 9 (p. 10) stellen wir uns zunächst vor, dass das Zeitintervall h ein Tag ist.

Um zu den Genesenen dazuzukommen, müssen infizierte Personen eine Infektion *auskurieren*. Das heißt man kann nur von Gruppe I nach Gruppe R wechseln und nicht von Gruppe S nach Gruppe R . Wie viele an einem Tag neu genesen hängt vor allem davon ab, wie viele Personen derzeit infiziert sind — Sind viele infiziert, so werden auch viele genesen.

Es ist sogar sinnvoll, anzunehmen, dass der Zusammenhang *proportional* ist: Verdoppeln sich die Infektionszahlen so werden sich (mit leichter Verzögerung) auch die Zahl der neu genesenen verdoppeln. Man kann daher modellieren

$$R(t + 1 \text{ Tag}) - R(t) \approx \beta I(t) \quad (387)$$

Hierbei ist $\beta > 0$ die Rate der Infizierten, die an einem Tag von einer Infektion genesen.

Der Parameter β ist der erste Modellparameter, der von dem modellierten Virus abhängt. Zu dessen Bestimmung benötigt man konkrete Daten über die Krankheit. Sinnvoll ist hier die folgende Wahl:

$$\beta = \frac{1}{\text{durchschnittliche Dauer einer Infektion (in Tagen)}}. \quad (388)$$

Das sieht man z.B. so: Dauert die Infektion immer exakt N Tage, so können wir die Infizierten in N Gruppen aufteilen — Hierbei ist die k -te Gruppe die Gruppe derjeniger, die schon k Tage lang infiziert sind ($k = 0, \dots, N - 1$). Falls wir über einige Tage ungefähr gleichviele Neuinfektionen haben, so können wir annehmen, dass jede Gruppe gleichgroß ist. Damit hat jede Gruppe anteilmäßig $\frac{1}{N}I(t) = \beta I(t)$ Mitglieder. Da nur eine Gruppe — nämlich die $N - 1$ -ste Gruppe — am Folgetag genesen wird, ist die Anzahl der täglich

neu Genesenen eben durch $\beta I(t)$ gegeben.

Verallgemeinern wir die Diskussion in auf allgemeine Zeitschritte $h > 0$ so erhalten wir

$$\frac{R(t+h) - R(t)}{h} \approx \beta I(t), \quad (389)$$

mit Gleichheit für sehr kleine h , denn bei großen h müsste man eigentlich eher $I(t+h)$ als $I(t)$ berücksichtigen. Der Grenzwert für $h \rightarrow 0$ liefert

$$R'(t) = \beta I(t). \quad (390)$$

Die Herleitung der Änderungsrate und die Einführung des Modellparameters β geschah unter der folgenden

Modellannahme 70. KURZE INFektION, KLEINE VERZÖGERUNG. Die Wahl von β in (388) haben wir unter der Annahme gerechtfertigt, dass jede Infektion gleichlang dauert. Häufig verhält es sich aber anders: Es gibt zumeist nur eine *Mindestdauer* einer Infektion. Jeden Tag nach dem Verstreichen dieser Mindestdauer gibt es dann eine gewisse Wahrscheinlichkeit, dass man sich an diesem Tag vollends erholt. Dauert eine Infektion z.B. mindestens 4 Tage, so müssten wir eigentlich die neu genesenen zufällig aus dem Pool der Infizierten beim Zeitpunkt ‘t-4 Tage’ wählen. In der Tat gibt es Modellansätze mit *verzögerten Differentialgleichungen*, d.h. man würde ansetzen

$$R'(t) = \tilde{\beta} I(t - \tau), \quad (391)$$

wobei τ die Mindestdauer der Infektion ist. Die Rate $\tilde{\beta}$ muss hier dann aber anders gewählt werden als in (388). Sie ist in dem Fall die Wahrscheinlichkeit, dass sich eine fixierte Person, deren Infektionsmindestdauer bereits abgelaufen ist, an einem Tag nach von der Infektion erholt.

Die Modellannahme der kleinen Verzögerung erweist sich als unrealistisch wenn entweder die Infektion sehr lange andauert oder das Infektionsgeschehen sehr unkonstant ist. In diesen Fällen kann nicht davon ausgegangen werden, dass $R'(t)$ proportional zu $I(t)$ ist. Da wir uns im Laufe dieser Vorlesung nur mit einer ‘Standardpandemie’ beschäftigen können, wollen wir diesen Sachverhalt hier nicht vertiefen.

Verzögerte Differentialgleichungen sind allerdings ein schönes Thema für ein Seminar oder eine Bachelor/Masterarbeit.

Als nächstes wollen wir die Änderung der Anfälligen, also der Gruppe S , besprechen.

Modellierung 71. ÄNDERUNGSRATE VON $S(t)$. Wir stellen uns mal experimentell ein Großraumbüro mit M Personen vor, in dem sich normalerweise nur Mitglieder der Gruppe S und der Gruppe R aufhalten. Heute betrachten wir aber einen Tag, an dem zusätzlich noch eine infizierte Person in dem Großraumbüro arbeitet und mit allen Personen dort Kontakt hat. Es sei p die Wahrscheinlichkeit, dass ein Infizierter einen Anfälligen bei einem eintägigen Kontakt ansteckt. Es sei $\tilde{S}(t)$ die (absolute, nicht anteilmäßige(!)) Zahl

der Anfälligen im Büro. Dann wird dieser Tag im Großraumbüro einen Anteil von $p\tilde{S}(t)$ Neuinfektionen hervorbringen. Das heißt auch: Die Zahl der Anfälligen sinkt um $p\tilde{S}(t)$. In Formeln

$$\tilde{S}(t + 1 \text{ Tag}) - \tilde{S}(t) \approx -p\tilde{S}(t). \quad (392)$$

Was passiert nun, wenn nicht nur eine infizierte Person, sondern gleich mehrere in dem Büro anwesend sind? Sagen wir mal es sind $\tilde{I}(t)$ Personen anwesend. In dem Fall würde wohl jede infizierte Person gleichermaßen $p\tilde{S}(t)$ Personen anstecken und man hätte

$$\tilde{S}(t + 1 \text{ Tag}) - \tilde{S}(t) \approx -p\tilde{S}(t)\tilde{I}(t). \quad (393)$$

Aus diesem Gesetz könnte man sicherlich wieder eine Differentialgleichung herleiten. Aber wir übersehen hier noch etwas: Die Gesamtbevölkerung ist kein Großraumbüro — Es haben nämlich an einem Tag nicht alle mit Allen Kontakt. Hat ein Infizierter täglich nur mit k von M Personen aus dem Großraumbüro Kontakt, so würde man nicht $p\tilde{S}(t)\tilde{I}(t)$ Neuinfektionen erwarten, sondern nur $\frac{k}{M}p\tilde{S}(t)\tilde{I}(t)$, d.h.

$$\tilde{S}(t + 1 \text{ Tag}) - \tilde{S}(t) \approx -p\frac{k}{M}\tilde{S}(t)\tilde{I}(t). \quad (394)$$

Da wir hier nun berücksichtigt haben, dass nicht alle mit Allen Kontakt haben, ist dieses Gesetz dann nicht nur für ein Großraumbüro, sondern auch für die Gesamtbevölkerung anwendbar. Wir definieren nun also $S(t) := \frac{\tilde{S}(t)}{M}$ den Anteil der Anfälligen und $I(t) := \frac{\tilde{I}(t)}{M}$ den Anteil der Infizierten. Wir erhalten mit (394) (die wir nochmal durch M dividieren)

$$S(t + 1 \text{ Tag}) - S(t) = -pkS(t)I(t). \quad (395)$$

Den Lesenden ist nun als Übungsaufgabe überlassen, mit dieser Erkenntnis die Differentialgleichung

$$S'(t) = -pkS(t)I(t) \quad (396)$$

zu begründen. Für spätere Rechnungen definieren wir den Modellparameter

$$\alpha := pk > 0, \quad (397)$$

wobei

$$p = \text{Wahrscheinlichkeit, sich bei einem eintägigen Kontakt zu infizieren} \quad (398)$$

und

$$k = \text{durchschnittliche Anzahl der Kontakte eines Infizierten an einem Tag.} \quad (399)$$

Die beiden Parameter zu einem zusammenzufassen, ist etwas gewagt: Der Parameter p ist etwas rein medizinisches, was damit zusammenhängt, wie ansteckend das betrachtete Virus ist. Der Parameter k hängt gar nicht mit dem Virus zusammen, sondern nur mit dem Kontaktverhalten der Bevölkerung. Das heißt aber insbesondere, dass man k mithilfe von Maßnahmen steuern kann. So zum Beispiel könnte man Infizierte in eine Quarantäne schicken, sodass sie keine Kontakte mehr haben.

Man könnte hier als Kritik anführen, dass bei der Berechnung von p die Kontaktlänge auf einen Tag gesetzt wurde. Bei manchen Viren reicht eine viel kürzere Zeit für eine Infektion bereits aus. Für das SARS-CoV-2 Virus zum Beispiel haben Kontakte mit einer Länge von 15 Minuten bereits ein gewisses Risiko.

Dieser Sachverhalt lässt sich aber mathematisch klären, da man stets festlegen darf, wie schnell die Zeit t vergehen soll. Man muss die Zeit also nicht in der Einheit ‘Tage’ messen sondern kann die Zeit auch in 15-Minütigen Zeiteinheiten unterteilen, wenn das sinnvoller ist. In dem Fall ist k die Anzahl der Kontakte in einer Zeiteinheit.

Dennoch gibt es eine diskutable Annahme:

Modellannahme 72. GLEICHVERTEILTHEIT DER KONTAKTE. Wir haben hier wieder mal *Proportionalitätsannahmen* gemacht. Diese sind immer nur sinnvoll, wenn der modellierte Sachverhalt *homogen verteilt* ist. So zum Beispiel haben wir angenommen, dass alle Infizierten gleich viele tägliche Kontakte k haben.

In der Praxis bemüht man sich, solche Annahmen loszuwerden. Wir könnten zum Beispiel im Vorfeld die Infizierten in zwei weitere Gruppen aufteilen: ‘Infizierte mit vielen Kontakten’ und ‘Infizierte mit wenigen Kontakten’. Diesen Gruppen würde man dann verschiedene Kontaktzahlen k_1 oder k_2 zuweisen. Man kann auch noch mehr Abstufungen machen. Auch Modelle, bei denen die Anzahl der Abstufungen gegen unendlich geht, werden untersucht. Dies sind dann *stochastische Modelle*, da der Kontaktanzahl eine Wahrscheinlichkeitsverteilung zugewiesen wird.

Es gilt nun also, wie in Modellierung 69 und 71 diskutiert,

$$R'(t) = \beta I(t), \quad (400)$$

$$S'(t) = -\alpha S(t)I(t). \quad (401)$$

Wie sich $I(t)$ verändert, müssen wir nicht mehr mithilfe einer Modellierung diskutieren, denn

$$R(t) + S(t) + I(t) = 1 \quad \forall t \in \mathbb{R}, \quad (402)$$

weswegen wir durch Ableiten beider Seiten der Gleichung feststellen

$$R'(t) + S'(t) + I'(t) = 0 \quad \forall t \in \mathbb{R}. \quad (403)$$

Wir folgern

$$I'(t) = -S'(t) - R'(t) = \alpha S(t)I(t) - \beta I(t). \quad (404)$$

Damit haben wir unser (nichtlineares) Differentialgleichungssystem

$$\begin{aligned} S'(t) &= -\alpha S(t)I(t) \\ I'(t) &= \alpha S(t)I(t) - \beta I(t) \\ R'(t) &= \beta I(t) \end{aligned} \quad (405)$$

Definition 73. SIR-PANDEMIE.

Die Gleichung (405) heißt *SIR*-Gleichung. Für ein offenes Intervall $I \supset [0, \infty)$ heißt eine Lösung $y \in C^1(I; \mathbb{R}^3)$ mit

$y(t) = (S(t), I(t), R(t))^T$, ($t \in I$) von (405) *SIR-Pandemie* mit Modellparametern $\alpha, \beta > 0$.

Von nun an studieren wir also Lösungen von (405), also von

Merkformel 74. SIR PANDEMIE-GLEICHUNG.

Für $\alpha, \beta > 0$ und $I \supset [0, \infty)$ betrachten wir

$$\begin{cases} S'(t) = -\alpha S(t)I(t) \\ I'(t) = \alpha S(t)I(t) - \beta I(t) \\ R'(t) = \beta I(t) \end{cases} \quad (t \in I) \quad (406)$$

Leider wissen wir zu diesem Zeitpunkt noch nicht ob eine Lösung von (405) auf einem so großen Intervall $I \supset [0, \infty)$ existiert. Wir werden aber später noch zeigen, dass dem so ist.

Gedankenexperiment 75. WACHSTUM AM ANFANG DER PANDEMIE. Eine interessante Situation ergibt sich am Anfang einer Pandemie, d.h. in einer Situation wo noch niemand genesen ist. Hier kann man $R(t) \approx 0$ annehmen. Damit gilt $S(t) + I(t) \approx 1$. Setzen wir also $S = 1 - I$ in die zweite Gleichung von (405) so erhalten wir

$$I'(t) = \alpha(1 - I(t))I(t) - \beta I(t) = I(t)(\alpha - \beta - \alpha I(t)), \quad (407)$$

eine Differentiagleichung nur für I ! Dieses Wachstumsverhalten haben wir am Anfang in Übung 12 (p. 13) als *logistisches Wachstum* kennengelernt. Am Anfang der Pandemie kann das logistische Wachstum also als eine gute Näherung für das Pandemiegeschehen aufgefasst werden. Wir haben in Übung 12 (p. 13) auch gesehen, dass logistisches Wachstum am Anfang der Pandemie stets aussieht wie exponentielles Wachstum. Das erklärt, warum am Anfang einer Pandemie von einem exponentiellen Wachstum der Infizierten auszugehen ist.

2.1.1 Erste Eigenschaften des Modells

Es folgen einige *Plausibilitätseigenschaften* des Modells. Wichtig hierbei ist, dass wir die Pandemie mit einem realistischen Anfangswert starten, d.h. wir müssen annehmen, dass

$$S(0) \in (0, 1), \quad I(0) \in (0, 1), \quad R(0) \in (0, 1) \quad (408)$$

und

$$S(0) + I(0) + R(0) = 1. \quad (409)$$

Wir bezeichnen in diesem Abschnitt

$$Z := \{(s, i, r) \in (0, 1)^3 : s + i + r = 1\}, \quad (410)$$

die Menge aller zulässigen Anfangswerte. Dass $R(0) > 0$ ist (d.h., dass am Anfang ein gewisser Teil der Bevölkerung immun ist) wird hier der Einfachheit halber angenommen, ist aber eigentlich nicht notwendig.

Wir sollten uns zunächst vergewissern, dass die Pandemie zu jedem Zeitpunkt sinnvolle Werte annimmt, d.h. dass es in dem Modell zum Beispiel nach einiger Zeit keine negativen Infiziertenzahlen gibt. Das könnte ja schließlich in der Theorie sein — es würde aber sicherlich kein gutes Licht auf das Modell werfen.

Satz 76. Es sei $y = (S, I, R)^T \in C^1(I, \mathbb{R}^3)$ eine SIR-Pandemie mit $y(0) \in Z$. Dann gilt $y(t) \in Z$ für alle $t \in [0, \infty)$.

Beweis. Dass $S(t) + I(t) + R(t) = 1$ für alle $t \in \mathbb{R}$ gilt ist allerdings nicht schwer zu zeigen. In der Tat

$$\frac{d}{dt}(S(t) + I(t) + R(t)) = S'(t) + I'(t) + R'(t) \quad (411)$$

$$\stackrel{(405)}{=} -\alpha S(t)I(t) + \alpha S(t)I(t) - \beta I(t) + \beta I(t) = 0. \quad (412)$$

Damit haben wir, dass $S + I + R$ konstant sein muss. Deswegen gilt

$$S(t) + I(t) + R(t) = S(0) + I(0) + R(0) = 1 \quad \forall t \in I. \quad (413)$$

Als nächstes zeigen wir, dass $S(t), I(t), R(t) > 0$ für alle $t \in [0, \infty)$. Hierzu sei

$$T := \sup\{t \geq 0 : S(\tau), I(\tau), R(\tau) > 0 \quad \forall \tau \in [0, t]\} \in (0, \infty]. \quad (414)$$

Sicherlich gilt $T > 0$, da $S(0), I(0), R(0) > 0$ und S, I und R stetige Funktionen sind. Wir behaupten nun, dass $T = \infty$. Um dies zu zeigen, nehmen wir an, dass $T < \infty$. Wir werden zeigen, dass

$$R(T) > 0, \quad S(T) > 0, \quad I(T) > 0. \quad (415)$$

Dies steht im Widerspruch zur Wahl von T , denn wenn $S(T), I(T), R(T) > 0$ wäre so gäbe es ein $\epsilon > 0$ mit $S(\tau), I(\tau), R(\tau) > 0$ für alle $\tau \in [T - \epsilon, T + \epsilon]$. Da aber (nach Definition von T in (414)) auch $S(\tau), I(\tau), R(\tau) > 0$ für alle $\tau \in [0, T - \epsilon]$, folgt insgesamt $S(\tau), I(\tau), R(\tau) > 0$ für alle $\tau \in [0, T + \epsilon]$. Es ergibt sich dadurch ein Widerspruch zur Wahl von T in (414). Wir arbeiten nun der Reihe nach die Behauptungen in (415) ab.

Schritt 1. Z: $R(T) > 0$. Für $t \in (0, T)$ gilt

$$R'(t) \stackrel{(405)}{=} \beta I(t) > 0, \quad (416)$$

wobei wir hier benutzt haben, dass wegen (414) gilt $I(t) > 0$ für alle $t \in (0, T)$. Daher ist R auf $(0, T)$ monoton steigend und es gilt

$$R(T) = \lim_{t \rightarrow T, t < T} R(t) \geq \lim_{t \rightarrow T, t < T} R(0) = R(0) > 0. \quad (417)$$

Schritt 2. \mathbb{Z} : $I(T) > 0$. Wir berechnen für $t \in (0, T)$

$$\frac{d}{dt} \log I(t) = \frac{I'(t)}{I(t)} \stackrel{(405)}{=} \frac{\alpha S(t)I(t) - \beta I(t)}{I(t)} = \alpha S(t) - \beta. \quad (418)$$

Da $S(t) > 0$ für alle $t \in (0, T)$ erhalten wir, dass

$$\frac{d}{dt} \log I(t) \geq -\beta \quad \forall t \in (0, T), \quad (419)$$

weswegen (nach dem Hauptsatz der Differential - und Integralrechnung und der Monotonie des Integrals) gilt

$$\log I(t) = \log I(0) + \int_0^t \frac{d}{ds} \log I(s) \, ds \geq \log I(0) + \int_0^t (-\beta) \, ds \quad \forall t \in (0, T). \quad (420)$$

Bilden wir nun auf der rechten Seite die Stammfunktion, so erhalten wir

$$\log I(t) \geq \log I(0) - \beta t \quad \forall t \in (0, T). \quad (421)$$

Man erhält durch Anwenden der (monotonen!) Exponentialfunktion und Verwenden des Exponentialgesetzes

$$I(t) = \exp(\log I(t)) \stackrel{(421)}{\geq} \exp(\log I(0) - \beta t) = I(0)e^{-\beta t} \quad \forall t \in (0, T) \quad (422)$$

und durch Betrachtung des Grenzwertes $t \rightarrow T$, ($t < T$) folgern wir

$$I(T) \geq I(0)e^{-\beta T} > 0. \quad (423)$$

Schritt 3. \mathbb{Z} : $S(T) > 0$. Wir berechnen mit (405) auf dem Intervall $(0, T)$

$$S'(t) = -\alpha S(t)I(t) \leq 0 \quad \forall t \in (0, T), \quad (424)$$

weswegen wir $S(t) \leq S(0)$ für alle $t \in (0, T)$ folgern. Benutzen wir dies, so erhalten wir folgende Abschätzung

$$\frac{d}{dt} \log I(t) = \frac{I'(t)}{I(t)} \stackrel{(405)}{=} \frac{\alpha S(t)I(t) - \beta I(t)}{I(t)} = \alpha S(t) - \beta \leq \alpha S(0) - \beta, \quad (425)$$

für $t \in (0, T)$. Wir folgern, dass für alle $t \in (0, T)$ gilt

$$\log I(t) - \log I(0) = \int_0^t \frac{d}{ds} \log I(s) \, ds \stackrel{(425)}{\leq} \int_0^t (\alpha S(0) - \beta) \, ds = (\alpha S(0) - \beta)t, \quad (426)$$

d.h. umgeformt

$$\log I(t) \leq \log I(0) + (\alpha S(0) - \beta)t \quad \forall t \in (0, T). \quad (427)$$

Wenden wir nun die (streng monotone!) Exponentialfunktion auf beiden Seiten an, so erhalten wir

$$I(t) \leq I(0)e^{(\alpha S(0) - \beta)t} \quad \forall t \in (0, T). \quad (428)$$

Als nächstes beobachten wir, dass für $t \in (0, T)$ gilt

$$\frac{d}{dt} \log S(t) = \frac{S'(t)}{S(t)} \stackrel{(405)}{=} \frac{-\alpha S(t)I(t)}{S(t)} = -\alpha I(t) \stackrel{(428)}{\geq} -\alpha I(0)e^{(\alpha S(0) - \beta)t}. \quad (429)$$

Man beachte nun, dass $(\alpha S(0) - \beta)t \leq |\alpha S(0) - \beta|t \leq |\alpha S(0) - \beta|T$ für alle $t \in (0, T)$. Als Folgerung daraus erhalten wir

$$\frac{d}{dt} \log S(t) \geq -\alpha I(0)e^{|\alpha S(0) - \beta|T}, \quad (430)$$

weswegen wir wieder durch Integration folgern

$$\log S(t) - \log S(0) \geq -\alpha I(0)e^{|\alpha S(0) - \beta|T}t \quad \forall t \in (0, T), \quad (431)$$

d.h.

$$\log S(t) \geq \log S(0) - \alpha I(0)e^{|\alpha S(0) - \beta|T}t \quad \forall t \in (0, T). \quad (432)$$

Wiederum durch Anwenden der Exponentialfunktion erhalten wir

$$S(t) \geq S(0) \exp(-\alpha I(0)e^{|\alpha S(0) - \beta|T}t) \quad \forall t \in (0, T). \quad (433)$$

Betrachten wir den Grenzwert $t \rightarrow T$, ($t < T$) so erhalten wir wegen der Stetigkeit von S

$$S(T) \geq S(0) \exp(-\alpha I(0)e^{|\alpha S(0) - \beta|T}T) > 0. \quad (434)$$

Da wir nun (415) gezeigt haben erhalten wir (wie direkt unter (415) erklärt) einen Widerspruch und folgern $T = \infty$. Insbesondere gilt also

$$S(t), I(t), R(t) > 0 \quad \forall t \in [0, \infty). \quad (435)$$

Zugleich folgt, dass $S(t), I(t), R(t) \in (0, 1)$ für alle $t \in [0, \infty)$, denn wenn $S(t), I(t), R(t) < 0$ und $S(t) + I(t) + R(t) = 1$, so gilt

$$1 = S(t) + I(t) + R(t) > S(t), \quad (436)$$

$$1 = S(t) + I(t) + R(t) > I(t), \quad (437)$$

$$1 = S(t) + I(t) + R(t) > R(t). \quad (438)$$

Die Behauptung ist gezeigt.

Ein solcher Satz ist nicht nur wichtig, um die Sinnhaftigkeit des Modells zu überprüfen, sondern auch für das Studium qualitativer Eigenschaften. Eine Konsequenz von elementarer Wichtigkeit ist zum Beispiel

Korollar 77. DISKUSSION DER MONOTONIE. Es sei y eine SIR-Pandemie mit $y(0) \in Z$. Es gilt

1. S ist streng monoton fallend auf $[0, \infty)$.
2. R ist streng monoton wachsend auf $[0, \infty)$.

Beweis. Da $y(t) \in Z$, gilt $S(t), I(t), R(t) \in (0, 1)$ für alle $t \in \mathbb{R}$. Mit der Tatsache, dass $\alpha > 0$ ist, folgern wir aus (405)

$$S'(t) = -\alpha S(t)I(t) < 0 \quad \forall t \in [0, \infty). \quad (439)$$

Außerdem gilt

$$R'(t) = \beta I(t) > 0 \quad \forall t \in [0, \infty). \quad (440)$$

Die Behauptung ist gezeigt.

Bei jedem DGL-System aus der Anwendung ist es sinnvoll, nach solchen Monotonierelationen zu suchen. Für I ist es schwierig, eine klare Monotonierelation zu zeigen. Es gilt

$$I'(t) = I(t)(\alpha S(t) - \beta) \quad \begin{cases} > 0 & S(t) > \frac{\beta}{\alpha}, \\ < 0 & S(t) < \frac{\beta}{\alpha}. \end{cases} \quad (441)$$

Wir halten fest: Falls die Anzahl der Anfälligen unter ein bestimmtes Niveau — nämlich unter $\frac{\beta}{\alpha}$ — fällt, so sinken die Infektionszahlen. Davor steigen sie an. Wir wissen allerdings zu diesem Zeitpunkt noch nicht, ob $S(t)$ das Niveau $\frac{\beta}{\alpha}$ jemals erreichen wird.

Falls $S(0) < \frac{\beta}{\alpha}$, so wird wegen der Monotonie von S das Niveau $\frac{\beta}{\alpha}$ niemals erreicht. Gleichung (441) zeigt dann, dass die Zahl der Infektion stetig kleiner wird. Wir werden noch sehen, dass die Pandemie in einem solchen Fall auch ausstirbt. Falls $S(0) > \frac{\beta}{\alpha}$, so können wir keine Aussage treffen. S könnte ja monoton fallend sein, aber die gesamte Zeit über dem Niveau $\frac{\beta}{\alpha}$ verharren. Allerdings werden wir im kommenden Abschnitt zeigen, dass das Niveau $\frac{\beta}{\alpha}$ in jedem Fall erreicht wird.

Man beachte, dass wegen der strengen Monotonie, $S(t) = \frac{\beta}{\alpha}$ maximal für einen Zeitpunkt $t = t^*$ erreicht wird. Diesen Zeitpunkt nennen wir — sofern er existiert — den *Höhepunkt der Pandemie*.

Aus (441) und der Monotonie von S lernt man dann

$$I'(t) \underset{(405)}{=} I(t)(\alpha S(t) - \beta) \quad \begin{cases} > 0 & t < t^*, \\ < 0 & t > t^*. \end{cases} \quad (442)$$

Man sieht mit dem *Vorzeichenwechselkriterium* für Extremstellen, dass I bei t^* ein lokales Maximum annimmt. Man muss hier aber beachten, dass das Vorzeichenwechselkriterium für Extremstellen kein klassisches Resultat aus der Analysis-Grundvorlesung, sondern zunächst nur aus der Schule ist. Eine formal richtige Argumentation funktioniert so:

Übung 78. DER VERLAUF EINER PANDEMIE. Es sei y eine SIR-Pandemie mit $y(0) \in Z$.

- (a) Zeigen Sie, dass $t \mapsto \log I(t)$ strikt konkav ist.
- (b) Folgern Sie: Falls es $t^* \geq 0$ gibt mit $S(t^*) = \frac{\beta}{\alpha}$, so hat $t \mapsto I(t)$ ein eindeutiges globales Maximum bei $t = t^*$.

SINN DER ÜBUNG: Man sieht, dass der Zeitpunkt t^* wirklich den absoluten Höhepunkt der Pandemie darstellt. Im nächsten Abschnitt wollen wir uns damit beschäftigen, wie viele Menschen an diesem Höhepunkt wirklich infiziert sind.

Diese Diskussion legt Folgendes nahe

Vermutung 79. Es sei y eine SIR-Pandemie mit $y(0) \in Z$. Dann tritt genau einer der folgenden zwei Fälle ein

- Falls $S(0) < \frac{\beta}{\alpha}$ so gilt $I' < 0$ und die Pandemie stirbt schließlich aus, d.h. $\lim_{t \rightarrow \infty} I(t) = 0$.
- Falls $S(0) > \frac{\beta}{\alpha}$ so steigt das Infektionsgeschehen bis zu einem Pandemie-Höhepunkt $t_* > 0$ wo $I(t^*)$ sein Maximum annimmt und $S(t_*) = \frac{\beta}{\alpha}$. Danach sinken die Infektionszahlen und die Pandemie stirbt aus, d.h. $\lim_{t \rightarrow \infty} I(t) = 0$.

Zwei Punkte in dieser Vermutung sind noch nicht geklärt: Die Existenz des Höhepunktes t_* im zweiten Fall und die Tatsache, dass die Pandemie schließlich ausstirbt.

Die Existenz des Pandemiehöhepunktes klären wir gleich am Anfang des nächsten Abschnittes in Satz 80. Das Aussterben der Pandemie folgt mit Übung 84.

2.1.2 Der Höhepunkt der Pandemie

In diesem Abschnitt wollen wir zeigen, dass die Pandemie im Falle $S(0) > \frac{\beta}{\alpha}$ ihren Höhepunkt t^* auch wirklich erreicht. Außerdem wollen wir bestimmen, wie viele Leute auf diesem Höhepunkt gleichzeitig infiziert sind.

Ehe wir anfangen wiederholen wir nochmal das DGL-System (405), was wir studieren

$$\begin{aligned} S'(t) &= -\alpha S(t)I(t), \\ I'(t) &= \alpha S(t)I(t) - \beta I(t), \\ R'(t) &= \beta I(t). \end{aligned} \quad (443)$$

Hierbei sind stets $\alpha, \beta > 0$ zu wählen wie in (388) und (397).

Satz 80. Es sei $y \in C^1(I; \mathbb{R}^3)$ eine SIR-Pandemie mit $y(0) \in Z$ und $S(0) > \frac{\beta}{\alpha}$. Dann gibt es genau ein $t^* \in [0, \infty)$ mit $S(t^*) = \frac{\beta}{\alpha}$.

Beweis. Zur Existenz: Angenommen ein solches $t^* \geq 0$ wie in der Aussage gibt es nicht.
ZWISCHENBEHAUPTUNG. Unter der obigen Annahme gilt

$$S(t) > \frac{\beta}{\alpha} \quad \forall t \in [0, \infty). \quad (444)$$

BEWEIS DER ZWISCHENBEHAUPTUNG. Angenommen es gibt ein $t_0 \in [0, \infty)$ mit $S(t_0) \leq \frac{\beta}{\alpha}$. Wegen der Annahme, dass der Wert $\frac{\beta}{\alpha}$ nicht angenommen wird, gilt sogar $S(t_0) < \frac{\beta}{\alpha}$. Dies führt also auf folgende Situation

$$S(0) > \frac{\beta}{\alpha}, \quad S(t_0) < \frac{\beta}{\alpha}. \quad (445)$$

Nach dem Zwischenwertsatz gibt es dann ein t_1 zwischen 0 und t_0 mit $S(t_1) = \frac{\beta}{\alpha}$, ein Widerspruch. Die Zwischenbehauptung ist bewiesen.

Mit (441) berechnen wir nun, dass $I'(t) > 0$ für alle $t \in I$ ist. Damit gilt auch $I(t) \geq I(0) > 0$ für alle $t > 0$. Dadurch hat man mit der dritten Gleichung aus (405), dass

$$R'(t) = \beta I(t) \geq \beta I(0). \quad (446)$$

Dann gilt aber

$$R(t) - R(0) = \int_0^t R'(s) \, ds \geq \int_0^t (\beta I(0)) \, ds = \beta I(0)t. \quad (447)$$

Schauen wir uns nun die linke und rechte Seite mal vergleichend an. Die rechte Seite der Gleichung ist unbeschränkt für große t , denn da $I(0) > 0$ gilt

$$\lim_{t \rightarrow \infty} \beta I(0)t = \infty. \quad (448)$$

Die linke Seite hingegen ist beschränkt, da

$$|R(t) - R(0)| \leq |R(t)| + |R(0)| \leq 2 \quad \forall t \in [0, \infty), \quad (449)$$

weil $R(0), R(t) \in (0, 1)$ für alle $t \in [0, \infty)$. Die Gleichungen (448) und (449) widersprechen sich. Daher muss die Annahme falsch sein und wir folgern die Existenz von t^* . Die Eindeutigkeit folgt aus der strengen Monotonie, siehe Korollar 77.

Wir haben also gesehen: Ist $S(0) > \frac{\beta}{\alpha}$, so erreicht die Pandemie stets einen eindeutigen Höhepunkt, an dem (siehe Übung 78 (p. 78)) die Infektionszahlen auch ein globales Maximum annehmen.

Wir können leider nicht mit elementaren Funktionen, charakterisieren, wann dieser Zeitpunkt t^* erreicht ist. Wir wissen nur, dass $S(t^*) = \frac{\beta}{\alpha}$.

Gerne würden wir aber eine Frage beantworten: Wie viele Personen sind auf dem Höhepunkt einer Pandemie gleichzeitig infiziert? Eine Antwort würde uns helfen, die maximale Auslastung der Gesundheitssysteme während einer Pandemie beziffern zu können.

Für diese Frage wäre es sicherlich hilfreich, den Zeitpunkt t^* zu kennen. Man benötigt ihn aber gar nicht explizit! Da $t \mapsto S(t)$ streng monoton (fallend) ist haben wir eine *Umkehrabbildung* zur Verfügung.

Die Lehre daraus ist, dass man das Pandemigeschehen nicht nur in Abhängigkeit der Zeit ausdrücken kann, sondern auch (äquivalent) in Abhängigkeit der Zahl der Anfälligen.

Wenn wir die Infektionszahlen in Abhängigkeit der Anfälligen ausdrücken können, wissen wir genau, was am Höhepunkt passiert — denn wir kennen ja die Zahl der Anfälligen am Höhepunkt, siehe (441)!

Im nächsten Satz werden wir sehen, dass es tatsächlich möglich ist, das Pandemigeschehen in Abhängigkeit der Anfälligen zu beschreiben.

Für den weiteren Verlauf des Abschnittes definieren wir

$$S_{\min} := \inf_{t \in [0, \infty)} S(t) = \lim_{t \rightarrow \infty} S(t) \in [0, S(0)]. \quad (450)$$

Die beiden Größen stimmen überein, weil S monoton fallend ist.

Satz 81. Es sei $y \in C^1(I; \mathbb{R}^3)$ eine SIR-Pandemie. Dann gibt es eine stetig differenzierbare Funktion $f : (S_{\min}, S(0)] \rightarrow [0, \infty)$ mit $S \circ f = id$, $f \circ S = id$. Ferner erfüllt

$$\widehat{I} := I \circ f : (S_{\min}, S(0)] \rightarrow (0, 1) \quad (451)$$

die Gleichung

$$\widehat{I}(x) = I(0) + S(0) - x + \frac{\beta}{\alpha} (\log x - \log S(0)). \quad (452)$$

Beweis. Die Existenz von $f = S^{-1}$ folgt direkt aus der strengen Monotonie von S . Außerdem gilt

$$S'(t) \stackrel{(405)}{=} -\alpha I(t)S(t) \neq 0 \quad \forall t \in [0, \infty), \quad (453)$$

weswegen die Umkehrabbildung $f = S^{-1}$ differenzierbar ist und es gilt nach einem Resul-

tat aus der Analysis-Vorlesung zur Ableitung der Umkehrfunktion

$$f'(x) = (S^{-1})'(x) = \frac{1}{S'(S^{-1}(x))} = \frac{1}{S'(f(x))} \quad (454)$$

$$\stackrel{(405)}{=} \frac{1}{-\alpha I(f(x))S(f(x))} = \frac{-1}{\alpha x I(f(x))} \quad \forall x \in (S_{min}, S(0)], \quad (455)$$

wobei wir im letzten Schritt benutzt haben, dass $f = S^{-1}$. Den aufmerksamen Beobachter:innen dieses Beweises wird nicht entgangen sein, dass das Resultat zur Ableitung der Umkehrfunktion in der Analysis eigentlich nur für Funktionen auf offenen Intervallen gezeigt wird. Hier haben wir es aber mit einem halboffenen Intervall zu tun. Diese Situation lässt sich leicht auflösen: Wir wissen, dass die *SIR*-Pandemie auf einer offenen Obermenge des Intervalls $[0, \infty)$ existiert. Damit muss es aus Stetigkeitsgründen auch ein $\epsilon > 0$ geben mit $S'(t) \neq 0$ für alle $t \in (-\epsilon, \infty)$. Die Umkehrung der Funktion kann also genausogut auf diesem offenen Oberintervall erfolgen.

Nun muss nur noch die Gleichung (452) nachgerechnet werden. Dazu differenzieren wir für $x \in (S_{min}, S(0)]$

$$\widehat{I}'(x) = (I \circ f)'(x) = I'(f(x))f'(x) \quad (456)$$

$$\stackrel{(452)}{=} (\alpha I(f(x))S(f(x)) - \beta I(f(x))) \frac{-1}{\alpha x I(f(x))} \quad (457)$$

$$= (\alpha I(f(x))x - \beta I(f(x))) \frac{-1}{\alpha x I(f(x))} = -1 + \frac{\beta}{\alpha x}. \quad (458)$$

Sei nun $x_0 \in (S_{min}, S(0)]$ beliebig aber fest. Wir integrieren den obigen Ausdruck von x_0 bis $S(0)$ und erhalten

$$\begin{aligned} \widehat{I}(S(0)) - \widehat{I}(x_0) &= \int_{x_0}^{S(0)} \widehat{I}'(x) \, dx = \int_{x_0}^{S(0)} \left(-1 + \frac{\beta}{\alpha x} \right) \, dx \\ &= x_0 - S(0) + \frac{\beta}{\alpha} (\log S(0) - \log x_0). \end{aligned} \quad (459)$$

Man beachte nun, dass

$$\widehat{I}(S(0)) = I(f(S(0))) = I(0). \quad (460)$$

Setzt man diese Information in (459) ein und multipliziert auf beiden Seiten mit (-1) , so erhält man

$$\widehat{I}(x_0) - I(0) = S(0) - x_0 + \frac{\beta}{\alpha} (\log x_0 - \log S(0)). \quad (461)$$

Addiert man noch $I(0)$ auf beiden Seiten und benennt die Variable x_0 in x um, so folgt (452).

Man beachte, dass wir in dem vorigen Satz nicht vorraussetzen mussten, dass $S(0) > \frac{\beta}{\alpha}$.

Gleichung (452) ist der Zusammenhang, den wir gesucht haben: Wir können die Infektionszahlen in Abhängigkeit der Anfälligen ausdrücken. In der Tat: sind x Leute anfällig, so sind $\widehat{I}(x) = I(S^{-1}(x))$ Leute infiziert. In der Anwendung wird hier oft eine Doppelbezeich-

nung verwendet: Die neue Variable x wird auch mit S bezeichnet. Außerdem wird zwischen \widehat{I} und I notationsmäßig nicht unterschieden. Falls man diese Unterscheidung nicht macht, sollte man aber den Anfangswert mit $(I_0, S_0, R(0))$ anstatt $(I(0), S(0), R(0))$ notieren, um nicht durcheinanderzukommen. (Aus Gründen, die wir später sehen werden, notieren wir jedoch $R(0)$ nicht mit R_0). Mit dieser Konvention liest sich (452) wie

$$I(S) = I_0 + S_0 - S + \frac{\beta}{\alpha}(\log S - \log S_0). \quad (462)$$

Diesen Zusammenhang findet man häufig so aufgeschrieben in der epidemiologischen Literatur. Wir werden aber im Folgenden nur Gleichung (452) verwenden.

Gedankenexperiment 82. EIN INTEGRAL DER BEWEGUNG FÜR DAS SIR-PANDEMIEMODELL.
Setzen wir in Gleichung (452) $x = S(t)$ für ein $t \in [0, \infty)$ ein so erhalten wir

$$I(t) \underset{(451)}{=} \widehat{I}(S(t)) = I(0) + S(0) - S(t) + \frac{\beta}{\alpha}(\log S(t) - \log S(0)). \quad \forall t \in [0, \infty). \quad (463)$$

Umgestellt merkt man

$$I(t) + S(t) - \frac{\beta}{\alpha} \log S(t) = I(0) + S(0) - \frac{\beta}{\alpha} \log S(0). \quad (464)$$

Diese Gleichung lässt sich auch als eine *Erhaltungsgröße* sehen:

$$I(t) + S(t) - \frac{\beta}{\alpha} \log S(t) \equiv \text{const.} \quad (465)$$

Mit anderen Worten haben wir ein *Integral der Bewegung* gefunden.

Wir erhalten das folgende

Korollar 83. Es sei $y \in C^1(I; \mathbb{R}^3)$ eine SIR-Pandemie mit $y(0) \in Z$, $S(0) > \frac{\beta}{\alpha}$. Dann gilt

$$\max_{t \in [0, \infty)} I(t) = I(0) + S(0) - \frac{\beta}{\alpha} + \frac{\beta}{\alpha} \left(\log \frac{\beta}{\alpha} - \log S(0) \right). \quad (466)$$

Beweis. Wir wissen aus Gleichung (441) oder aus Übung 78 (p. 78), dass

$$\max_{t \in [0, \infty)} I(t) = I(t^*), \quad (467)$$

wobei t^* der (nach Satz 80 (p. 79) existente und eindeutige) Zeitpunkt ist mit $S(t^*) = \frac{\beta}{\alpha}$. Setzt man nun t^* in (463) ein, so erhält man

$$I(t^*) = I(0) + S(0) - \frac{\beta}{\alpha} + \frac{\beta}{\alpha} \left(\log \frac{\beta}{\alpha} - \log S(0) \right), \quad (468)$$

und die Behauptung folgt zusammen mit (467).

Nimmt man an, dass $I(0) \approx 0$ und $S(0) \approx 1$, d.h. dass am Anfang der Betrachtung der Pandemie fast niemand immun ist und die Infektionszahlen sehr klein sind, so erhält man

$$\max_{t \in [0, \infty)} I(t) \approx 1 + \frac{\beta}{\alpha} \left(\log \frac{\beta}{\alpha} - 1 \right). \quad (469)$$

Mit dieser Formel kann man also bereits am Anfang der Pandemie die maximale Auslastung der Krankenhäuser vorraussagen – vorausgesetzt man kennt α und β .

Eine interessante Beobachtung: Ist $\frac{\beta}{\alpha} \approx 0$, so ist $\max_{t \in [0, \infty)} I(t) \approx 1$, d.h. die Pandemie verbreitet sich so schnell, dass fast alle gleichzeitig infiziert sind. Falls $\frac{\beta}{\alpha} \approx 1$, so ist $\max_{t \in [0, \infty)} I(t) \approx 0$, d.h. die Pandemie kann die Gesundheitssysteme in diesem Fall nicht belasten.

Wir haben in dieser Sektion gesehen, dass es manchmal sinnvoll sein kann, die Zeit durch andere (zeitmonotone) Größen auszudrücken. Dies kann Integrale der Bewegung liefern und wird in der Praxis häufig angewendet.

Eine Größe, die in unserer Diskussion immer wieder auftaucht, ist $\frac{\beta}{\alpha}$. Sie hängt von den Eigenschaften des Virus ab und bestimmt u.a., wo der Höhepunkt der Pandemie ist und wie viele Menschen gleichzeitig infiziert sind.

Ist $\frac{\beta}{\alpha} > 1$, so ist Zahl der Infizierten auf ganz $[0, \infty)$ rückläufig, siehe (441). Dies macht die (aussterbende) Pandemie gänzlich uninteressant. Ist $\frac{\beta}{\alpha} < 1$ so kann sich (auch nach (441)) die Pandemie nur weiter ausbreiten falls $S(0) > \frac{\beta}{\alpha}$, d.h. mehr als $\frac{\beta}{\alpha}$ Prozent der Bevölkerung sind anfällig.

Im nächsten Abschnitt widmen wir uns der virologischen Bedeutung von $\frac{\beta}{\alpha}$ und entdecken in der DGL das Phänomen der *Herdenimmunität*, d.h. wir finden einen Zeitpunkt, ab dem das Infektionsgeschehen exponentiell abfällt.

Übung 84. WIE VIELE PERSONEN INFIZIEREN SICH IM LAUFE DER PANDEMIE INSGESAMT?

Es sei y eine *SIR*-Pandemie mit $\frac{\beta}{\alpha} < 1$.

- (a) Zeigen Sie: $\lim_{t \rightarrow \infty} I(t) = 0$.
- (b) Beweisen Sie: $S_{min} > 0$ und es gilt folgende Gleichung

$$S_{min} - \frac{\beta}{\alpha} \log S_{min} = I(0) + S(0) - \frac{\beta}{\alpha} \log S(0). \quad (470)$$

- (c) Wir definieren die Funktion $W : (0, \frac{\beta}{\alpha}) \rightarrow (\frac{\beta}{\alpha} - \frac{\beta}{\alpha} \log \frac{\beta}{\alpha}, \infty)$ durch $W(x) := x - \frac{\beta}{\alpha} \log x$. Zeigen Sie, dass W bijektiv ist.
- (d) Wie viele Personen werden sich während des Verlaufes Pandemie insgesamt infizieren? Geben Sie eine explizite Formel an.

HINWEIS: Für Teilaufgabe (a) benötigen Sie nur (442) und das DGL-System! Es gibt hier auch andere Lösungswege, die mehr Eigenschaften benutzen, aber diese sind oft komplizierter als es sein müsste.

SINN DER ÜBUNG: Sicherlich ist auch hilfreich zu wissen, wie viele Personen während des gesamten Pandemieverlaufs infiziert werden. Wieder einmal purzelt diese Größe aus dem Integral der Bewegung heraus. Teilaufgabe (a) klärt das *Aussterben der Pandemie* und ist damit eine der ersten Aussagen, die wir in dieser Vorlesung über das *asymptotische Verhalten* einer Lösung, d.h. über das Verhalten für große Zeiten, treffen können. Dies steht zurzeit aber noch auf wackeligen Beinen, denn wir können ja noch nicht einmal beweisen, dass die Lösung überhaupt für alle Zeiten existiert. Dies soll uns im nächsten Kapitel bevorstehen.

2.1.3 Exkurs: Basisreproduktionszahl und Herdenimmunität

In diesem Abschnitt möchten wir — mal mehr und mal weniger mathematisch — ein Paar grundlegende epidemiologische Phänomene besprechen und mit den Grundlagen der letzten Abschnitte untersuchen.

Die Basisreproduktionszahl. Zu Anfang der Diskussion wiederholen wir mal, wie α und β in unserem Modell zu wählen waren. Es galt

$$\beta = \frac{1}{\text{durchschn. Dauer der Infektion (in Tagen)}} \quad (471)$$

und $\alpha = pk$, wobei

$$p = \text{Wahrscheinlichkeit der Übertragung bei eintätigem Kontakt} \quad (472)$$

und

$$k = \text{durchschn. Anzahl der Kontakte eines Infizierten an einem Tag}. \quad (473)$$

Anstatt eines Tages kann — wie bereits besprochen — auch eine beliebige andere Zeiteinheit gewählt werden. Die Größen müssen alle als Durchschnittswerte angegeben werden. Anstatt unserer Zielgröße $\frac{\beta}{\alpha}$ untersuchen wir nun den Kehrwert $\frac{\alpha}{\beta}$. Wir stellen fest

$$\frac{\alpha}{\beta} = (\text{Übertragungswahrscheinlichkeit}) \cdot (\text{Dauer}) \cdot (\text{Anzahl tägl. Kontakte}). \quad (474)$$

Die Aussagekraft dieser Größe ist wie folgt: Das Produkt der letzten beiden Faktoren, also von ‘Dauer’ und ‘täglichen Kontakten’ beschreibt, wie viele Leute eine infizierte Person während der gesamten ‘Infektionskarriere’ hat. Multipliziert man diese Anzahl der gesamten Kontakte mit der Übertragungswahrscheinlichkeit so ergibt sich

Merksatz 85. Die Größe $\frac{\alpha}{\beta}$ gibt an, wie viele Personen ein Infizierter während seiner gesamten ‘Infektionskarriere’ ansteckt.

In der Epidemiologie ist diese Kenngröße von höchstem Interesse, sie wird dort *Basisreproduktionszahl* genannt.

Notation 86. Wir nennen $\frac{\alpha}{\beta}$ die *Basisreproduktionszahl* und bezeichnen sie mit $R_0 := \frac{\alpha}{\beta}$.

Unglücklicherweise ist der Buchstabe R_0 schwierig von $R(0)$, dem Anteil der anfänglich immunen, zu unterscheiden.

Schätzung der Basisreproduktionszahl. Wie können wir die Reproduktionszahl eines Virus bestimmen? Hier gibt es mehrere Möglichkeiten: Wenn wir die gesamten virologischen Kenngrößen β, p, k zur Verfügung haben, haben wir bereits eine Formel um R_0 zu bestimmen.

Falls wir die Kenngrößen nicht kennen, so müssen wir probieren, R_0 aus den Daten herauszulesen. Dazu hier ein Paar Ideen aus der Epidemiologie.

Wir haben bereits gesehen, dass der Höhepunkt des Pandemiegeschehens sich ereignet, wenn

$$S(t^*) = \frac{\beta}{\alpha} = \frac{1}{R_0}. \quad (475)$$

So könnte man — wenn man Daten über einen kompletten Pandemieverlauf hat — auch die Reproduktionszahl schätzen: Man muss sich lediglich ansehen, wie viele Personen auf einem Höhepunkt einer Pandemie anfällig waren.

Wenn man diese Zahl der Anfälligen auch genau kennt, kann man stattdessen nicht R_0 aus der maximalen Anzahl der Infizierten berechnen. Aus (469) folgert man nämlich, dass

$$\max_{t \in [0, \infty)} I(t) = 1 + \frac{1}{R_0} \left(\log \frac{1}{R_0} - 1 \right). \quad (476)$$

Problematisch ist hierbei aber, dass sich während der Pandemie das Kontaktverhalten der Bevölkerung verändern kann, z.B. durch Maßnahmen wie Quarantäneanordnungen, Schulschließungen oder auch einfach durch individuelle Vorsicht. Dadurch kann es sein, dass es bei einer Pandemie gar nicht so schlimm kommt, wie das Modell ursprünglich vorhersagt. Dies könnte dazu führen, dass R_0 unterschätzt wird.

Eine weitere numerische Methode ist *curve fitting*. Man plottet hierbei eine Approximation der Lösung für verschiedene Parameter $\alpha, \beta > 0$ und sucht sich diejenigen Werte von α, β aus, die am besten zu den realen Daten passen.

Da sich der vorangegangene Abschnitt zur Bestimmung von R_0 mehr mit Epidemiologie als mit Mathematik beschäftigt, gebe ich hier eine Quelle an, die zeigt, dass tatsächlich alle obigen Methoden benutzt werden um R_0 zu schätzen: [LINK ZUR QUELLE](#) [COBURN, WAGNER, BLOWER (2009)]

Herdenimmunität I. Von R_0 hängt nicht nur ab, wie schlimm die Pandemie wird, sondern auch wann und wie sie endet.

Das Abklingen der Pandemie beginnt bei t^* — dort ist $S(t^*) = \frac{1}{R_0}$. Das heißt insbesondere

$$R(t^*) + I(t^*) = 1 - \frac{1}{R_0}, \quad (477)$$

also ist zum Zeitpunkt t^* exakt ein Anteil von $1 - \frac{1}{R_0}$ der Bevölkerung entweder immun oder haben eine Infektion durchgemacht oder sind derzeit infiziert. In der Epidemiologie sagt man daher, dass sich ' $1 - \frac{1}{R_0}$ *100 Prozent' der Bevölkerung erst infizieren oder immun werden

muss, damit *Herdenimmunität* besteht. Ein Beispiel hierfür: Beim SARS-CoV 2 Coronavirus ging man am Anfang der Pandemie im Jahre 2020 zunächst von $R_0 \approx 3$ aus. So wurde also vermutet, dass sich $1 - \frac{1}{3} = \frac{2}{3} \approx 66\%$ der Bevölkerung infizieren müssen, damit die Pandemie unter Kontrolle gebracht werden kann. Später mutierte das Virus — man geht davon aus, dass die sogenannte *Delta-Variante* von Sars-CoV-2 eine Basisreproduktionszahl von $R_0 \approx 5.5$ hat. Das heißt, dass der Herdenschutz erst eintreten würde, wenn 81% der Bevölkerung infiziert oder immun wären.

Es gibt auch eine stochastische Herangehensweise, die die Herdenimmunitäts-Schwelle von $1 - \frac{1}{R_0}$ auf eine andere Weise rechtfertigt. Für die (populärwissenschaftlich erklärte) Idee hierzu verweisen wir auf YOUTUBE-VIDEO “SO ENDET CORONA”.

Reaktion auf Virusmutationen. Ändern sich die Parameter des Modells, z.B. durch eine neue Virus-Variante, so muss man das die Pandemie *neu starten*. Das heißt: Man nimmt den aktuellen Werte $(S(t), I(t), R(t))$ als neuen Anfangswert und startet dort die DGL mit diesen neuen Parametern. Man setzt also die Zeit auf Null zurück und startet die Simulation mit den neuen Parametern neu. Dies muss man z.B. auch machen, wenn sich zum Beispiel das *Kontaktverhalten* der Bevölkerung ändert.

Reinfektionen und Immunitätsverlust. Bei der (gegenwärtig noch laufenden) Corona-Pandemie geht man — wie oben erwähnt — davon aus, dass 81 % der Bevölkerung infiziert oder immun sein muss, damit der Herdenschutz eintritt. Darauf ist aber in der Realität kein Verlass: Da eine *Reinfektion* mit dem Virus möglich ist, also der Zustand der vollständigen Immunität (modelliert durch die Gruppe R) quasi nicht existiert, ist unser Modell ja eigentlich nicht anwendbar. Wir müssen die Möglichkeit eines *Immunitätsverlustes* nach einer gewissen Zeit nach einer durchstandenen Infektion berücksichtigen. Ein sinnvolles Modell für Reinfektionen wäre das Folgende:

Modellierung 87. SIR-MODELL MIT IMMUNITÄTSVERLUST. Wir nehmen hier an, dass zu jedem Zeitpunkt ein gewisser Prozentsatz der Personen in Gruppe R die Immunität verliert. Damit gehen diese Personen von Gruppe R in Gruppe S über, sie sind wieder anfällig. Die Rate mit der das geschieht können wir mit den gleichen Annahmen wie in der Herleitung von (388) beziffern (HA). Wir setzen daher

$$\gamma = \frac{1}{\text{durchschn. Dauer der Immunität}}. \quad (478)$$

und erhalten ein neues DGL System

$$\begin{aligned} S'(t) &= -\alpha S(t)I(t) + \gamma R(t), \\ I'(t) &= \alpha S(t)I(t) - \beta I(t), \\ R'(t) &= \beta I(t) - \gamma R(t). \end{aligned} \quad (479)$$

In diesem Modell steckt eine problematische Annahme: Eigentlich gibt es in Gruppe R ja zwei Arten von Personen: Die *Genesenen* und die *Verstorbenen*. Während Genesene sich (wie im obigen Gesetz) reinfizieren können sind Verstorbene (drastisch formuliert) für immer immun. Sie können also nicht mehr in die Gruppe S . Die Annahme, dass wir hier mit einer konstanten Rate einen Übergang von Gruppe R nach S haben funktioniert also

nur, wenn während der Pandemie kaum jemand verstirbt. Ein gutes Beispiel wäre hier das *Herpesvirus*. Hingegen ist z.B. die *spanische Grippe* weit von dieser Annahme entfernt. Will man die Verstorbenen mit einbeziehen, so müsste man aber selbstverständlich auch mit einbeziehen, dass Menschen aus allen drei Gruppen durch andere Ereignisse sterben können und andere Menschen in die Gruppe S hineingeboren werden. Das alles sparen wir uns an dieser Stelle.

Übung 88. Erklären Sie wie in Sektion 2.1, wie (479) zustande kommt und warum

$$\gamma = \frac{1}{\text{durchschn. Dauer der Immunität}} \quad (480)$$

wie in (478) eine passende Wahl für den neuen Modellparameter γ ist.

Dieses neue (in der Literatur ausdrücklich nichtstandard-)Modell wird schwerer zu untersuchen sein. Wir werden im Laufe der Vorlesung zeigen können, dass das Modell sinnvoll ist, d.h. Die DGL (479) lässt sich stets auf $[0, \infty)$ lösen und es gilt: Falls $(S(0), I(0), R(0)) \in Z$ so ist $(S(t), I(t), R(t)) \in Z$ für alle $t \in [0, \infty)$, wobei Z wie in (410) definiert ist. Dies wird aber auch wieder fortgeschrittene Methoden benötigen.

Die Beschreibung der Pandemie gestaltet sich auch komplizierter: Sämtliche Monotonieaussagen wie in Korollar 77 werden nicht verfügbar sein. Wir erwarten auch nicht, dass die gelten — Die Pandemie wird sich *wellenartig* ausbreiten.

Falls am Ende der Vorlesung Zeit ist, werden wir auch sehen, ob die Pandemie in diesem Modell auch ausstirbt, oder ob sie ewig weitergeht. Dies kann — und wird — von den Parametern α, β, γ abhängen.

Herdenimmunität II. Wir verlieren zum Schluss noch ein Wort zum Ende der Pandemie. Hier betrachten wir wieder das ursprüngliche Modell ohne Immunitätsverlust. Was passiert, nachdem die Bevölkerung die Herdenimmunität erreicht? Wir werden nun zeigen, dass ab dem Zeitpunkt t^* ein *exponentieller Zerfall* der Pandemie beginnt.

Satz 89. HERDENIMMUNITÄT. Es sei $y \in C^1(I; \mathbb{R}^3)$ ($I \supset [0, \infty)$) eine *SIR*-Pandemie mit $y(0) \in Z$ und $S(0) > \frac{\beta}{\alpha}$. Ferner definieren wir wie in Notation 86 (p. 85) $R_0 := \frac{\alpha}{\beta}$. Es sei t^* wie in Satz 80 (p. 79) der Höhepunkt der Pandemie und $T > t^*$ beliebig. Dann gilt $S(T) - \frac{1}{R_0} < 0$ und

$$I(t) \leq I(T) e^{\alpha(S(T) - \frac{1}{R_0})(t-T)} \quad \forall t \in [T, \infty). \quad (481)$$

Insbesondere gilt

$$\lim_{t \rightarrow \infty} t^m I(t) = 0 \quad \forall m \in \mathbb{N}, \quad (482)$$

d.h. $I(t)$ strebt schneller gegen Null als jedes Polynom.

Beweis. Sei $T > t^*$ wie in der Aussage des Satzes gegeben. Dass $S(T) - \frac{1}{R_0} < 0$ liegt daran, dass $S(t^*) = \frac{1}{R_0}$ und S streng monoton fallend ist. Wir zeigen nun (481). Es sei nun $t \in [T, \infty)$. Dann gilt mit der Monotonieaussage von Korollar 77 (p. 77)

$$I'(t) \underset{(405)}{=} (\alpha S(t) - \beta) I(t) \leq (\alpha S(T) - \beta) I(t) = \alpha \left(S(T) - \frac{\beta}{\alpha} \right) I(t). \quad (483)$$

Wir können nun verwenden, dass $\frac{\beta}{\alpha} = \frac{1}{R_0}$. Dividieren wir durch $I(t)$ so folgt

$$\frac{I'(t)}{I(t)} \leq \alpha \left(S(T) - \frac{1}{R_0} \right) \quad \forall t \geq T. \quad (484)$$

Mit der Kettenregel (rückwärts gelesen) folgern wir

$$\frac{d}{dt} \log I(t) \leq \alpha \left(S(T) - \frac{1}{R_0} \right) \quad \forall t \geq T. \quad (485)$$

Wir folgern mit der Monotonie des Integrals für alle $t \geq T$

$$\log I(t) - \log I(T) = \int_T^t \frac{d}{ds} \log I(s) \, ds \leq \int_T^t \alpha \left(S(T) - \frac{1}{R_0} \right) \, ds \quad (486)$$

$$= \alpha \left(S(T) - \frac{1}{R_0} \right) (T - t). \quad (487)$$

Umgestellt bedeutet das

$$\log I(t) \leq \log I(T) + \alpha \left(S(T) - \frac{1}{R_0} \right) (T - t). \quad (488)$$

Wir wenden nun die Exponentialfunktion $\exp(x) = e^x$ auf beiden Seiten an. Da diese Funktion monoton wachsend ist schließen wir

$$I(t) \leq \exp \left(\log I(T) + \alpha \left(S(T) - \frac{1}{R_0} \right) (T - t) \right) \quad (489)$$

$$= e^{\log I(T)} e^{\alpha \left(S(T) - \frac{1}{R_0} \right) (T - t)}. \quad (490)$$

Hierbei haben wir dass Exponentialgesetz $e^{a+b} = e^a e^b$ verwendet. Es ergibt sich nun

$$I(t) \leq I(T) e^{\alpha \left(S(T) - \frac{1}{R_0} \right) (T - t)}. \quad (491)$$

Daraus folgt (481). Nun zu (482). Hierfür wiederholen wir zunächst aus der Analysis (HA), dass für alle $\delta > 0$ gilt

$$\lim_{t \rightarrow \infty} t^m e^{-\delta t} = 0 \quad (492)$$

Mit $\delta := -\alpha \left(S(T) - \frac{1}{R_0} \right) > 0$ folgern wir nun, dass $\lim_{t \rightarrow \infty} t^m e^{-\delta t} = 0$ und daher gilt

$$0 \leq t^m I(t) \leq I(T) e^{\alpha \left(S(T) - \frac{1}{R_0} \right) T} t^m e^{-\delta t} \rightarrow 0 \quad (t \rightarrow \infty). \quad (493)$$

Die Behauptung folgt aus dem ‘Sandwichlemma’ für Grenzwerte.

3 Existenz, Eindeutigkeit, Globalität

In Kapitel 2 konnten wir ein Pandemiemodell ausführlich studieren. Allerdings immer nur unter einer Annahme: Wir haben angenommen, dass wir eine Lösung in einer offenen Obermenge von $[0, \infty)$ finden. Würden wir diese nicht finden, so wäre das Modell nicht sinnvoll — denn warum sollte eine Pandemie irgendwann aufhören, beschreibbar zu sein?

Aus solchen Gründen müssen wir uns der Frage nach Existenz, Eindeutigkeit und Globalität von Lösungen widmen. Können wir aus der Differentialgleichung ablesen, ob Lösungen für immer existieren?

3.1 Der Satz von Picard-Lindelöf

Wir möchten hier in diesem Kapitel den Satz von Picard-Lindelöf zeigen. Wie mehrfach angekündigt geht es darum, Anfangswertprobleme der Form

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (494)$$

unter gewissen Voraussetzungen an I und $f : I \times G \rightarrow \mathbb{R}^n$ auf Existenz und Eindeutigkeit zu untersuchen. Wir erinnern uns daran, dass $I = (a, b)$ stets offen ist und $G \subset \mathbb{R}^n$ ein Teilgebiet ist.

Man beachte an dieser Stelle, dass wir jedoch im Allgemeinen nicht einfach ohne Weiteres $I = \mathbb{R}$ wählen sollten. Wir haben ja schon Beispiele gesehen (siehe Beispiel 5 (p. 4), wo mit einer globalen Lösung auf \mathbb{R} nicht zu rechnen ist).

Der Beweis wird sich über die nächsten paar Abschnitte hinziehen, hier soll zunächst die Formulierung diskutiert werden. Zuallererst definieren wir die essentielle Bedingung, die wir an die Funktion f stellen müssen, damit eindeutige Lösungen existieren.

Definition 89. GLOBALE LIPSCHITZ-BEDINGUNG. Eine Funktion $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ erfüllt eine *globale Lipschitz-Bedingung* auf I falls es ein $L > 0$ gibt mit

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \quad \forall t \in I \quad \forall y_1, y_2 \in \mathbb{R}^n. \quad (495)$$

Falls diese Bedingung an f erfüllt ist, so können wir von Existenz und Eindeutigkeit einer Lösung ausgehen. Das jedenfalls besagt der folgende Satz, der eine erste Version unseres Existenz- und Eindeutigkeitsresultates ist. Wir werden später noch andere Versionen dieses Satzes kennenlernen, die leicht veränderte Sachverhalte untersuchen.

Satz 90. SATZ VON PICARD-LINDELÖF, GLOBALE VERSION. Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ erfülle eine globale Lipschitz-Bedingung auf I . Dann hat für alle $t_0 \in I$ und $y_0 \in \mathbb{R}^n$ das Anfangswertproblem

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (496)$$

eine eindeutige Lösung.

Beweis. Später.

Man beachte, dass wir dieses Resultat tatsächlich nur für $G = \mathbb{R}^n$ formulieren und beweisen können. Sind kleinere Gebiete erlaubt, so kann nicht mit einer Lösung auf ganz I rechnen. Als (pathologisches) Beispiel

Übung 91. UNABDINGBARKEIT DER VORRAUSSETZUNG ‘ $G = \mathbb{R}^n$ ’ FÜR DEN GLOBALEN SATZ VON PICARD-LINDELÖF.

Sei $G = (0, 1) \subset \mathbb{R}^1$ und $f : \mathbb{R} \times (0, 1) \rightarrow \mathbb{R}$ gegeben durch

$$f(t, z) := z \quad (t, z) \in \mathbb{R} \times G. \quad (497)$$

Zeigen Sie, dass das AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in \mathbb{R}) \\ y(0) = \frac{1}{2} \end{cases} \quad (498)$$

keine Lösung auf $I = \mathbb{R}$ besitzt. Besitzt das AWP auf einer offenen Teilmenge von I eine eindeutige Lösung?

SINN DER ÜBUNG. Man sieht: Schränkt man sein Definitionsgebiet G zu stark ein, so kann man auch nicht mit einer globalen Lösung rechnen. Die Einschränkung $G = (0, 1)$ ist hier natürlich künstlich, weil man dieselbe Funktion ja auch auf einem größeren Gebiet, nämlich auf $G = \mathbb{R}$ definieren könnte. Nichtsdestotrotz sieht man, dass man sich für allgemeine Gebiete G womöglich nur mit einer Lösung auf einem Teilintervall abfinden sollte.

Trotz dieses Mankos: Viele Fälle können wir bereits mit dieser ersten Version des Satzes von Picard-Lindelöf lösen, unter Anderem auch das Fadenpendel.

Übung 92. GLOBALITÄT DES FADENPENDELS. Zeigen Sie, dass für alle $\varphi_0, \alpha_0 \in \mathbb{R}$ die Gleichung

$$\begin{cases} \varphi''(t) = -\omega^2 \sin(\varphi(t)) & (t \in \mathbb{R}) \\ \varphi(0) = \varphi_0 \\ \varphi'(0) = \alpha_0 \end{cases} \quad (499)$$

eine globale Lösung $\varphi \in C^2(\mathbb{R}; \mathbb{R})$ besitzt.

SINN DER ÜBUNG. Man sieht hier, dass der Satz von Picard-Lindelöf nicht nur für Gleichungen erster Ordnung, sondern auch für Gleichungen höherer Ordnung angewendet werden kann. Die wichtigen Ideen um Differentialgleichungen höherer Ordnung auf DGLs erster Ordnung zurückzuführen haben wir ja bereits erklärt.

Ohne Lipschitz-Bedingung verlieren wir sofort die Eindeutigkeit, wie wir in Beispiel 4 (p. 4) eindrucksvoll gesehen haben.

Vielleicht lässt sich die Lipschitz-Bedingung aber doch an manchen Stellen etwas abschwächen. Wir haben bereits erfolgreich Anfangswertprobleme studiert, bei denen die Lipschitzbedingung (495) nicht erfüllt ist. Man erinnere sich an Beispiel 5 (p. 4), wo wir das AWP

$$\begin{cases} y'(t) = y(t)^2 & (t \in \mathbb{R}) \\ y(0) = y_0 \end{cases} \quad (500)$$

studiert haben. Auch hier waren wir – trotz fehlender Lipschitzstetigkeit von $f(t, z) = z^2$ – in der Lage, Lösungen des AWP zu finden. Damals mussten wir uns dann nur damit abfinden, dass die Lösungen vielleicht nicht auf ganz \mathbb{R} definiert sind.

In anderen Fällen ist es uns aber auch ohne globale Lipschitz-Bedingung gelungen, die DGL global zu lösen. Wir erinnern uns hier an die Übung 12 (p. 13) mit

$$\begin{cases} y'(t) = y(t)(1 - y(t)) & (t \in \mathbb{R}) \\ y(0) = y_0 \in (0, 1). \end{cases} \quad (501)$$

Für solche Angelegenheiten brauchen wir auch ein Existenz- und Eindeutigkeitsresultat unter schwächeren Voraussetzungen. Hierbei kommt die folgende Bedingung ins Spiel:

Definition 93. LOKALE LIPSCHITZ-BEDINGUNG. Wir sagen eine Funktion $f \in C^0(I \times G; \mathbb{R}^n)$ erfüllt eine *lokale Lipschitz-Bedingung* falls es zu jedem $(t_0, y_0) \in I \times G$ ein $\epsilon > 0$ sowie ein $r > 0$ und ein $L > 0$ gibt mit $(t_0 - \epsilon, t_0 + \epsilon) \subset I$ sowie $B_r(y_0) \subset G$ und

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \quad \forall t \in (t_0 - \epsilon, t_0 + \epsilon) \quad \forall y_1, y_2 \in B_r(y_0). \quad (502)$$

Man beachte: ϵ, r und L dürfen in dieser Definition stets von (t_0, y_0) abhängen. Wir schreiben daher auch zukünftig $\epsilon(t_0, y_0)$, $r(t_0, y_0)$ und $L(t_0, y_0)$. Die lokale Lipschitz-Bedingung sagt nichts anderes, als dass jeder Punkt $(t_0, y_0) \in I \times G$ eine kleine Umgebung besitzt, auf der eine Lipschitz-Abschätzung gilt. Mit dieser Bedingung können wir nun folgenden Satz formulieren

Satz 94. SATZ VON PICARD-LINDELÖF, LOKALE VERSION. Es genüge $f \in C^0(I \times G; \mathbb{R}^n)$ einer lokalen Lipschitz-Bedingung. Dann gibt es zu jedem $(t_0, y_0) \in I \times G$ ein eindeutiges offenes Intervall $I_{max} = I_{max}(t_0, y_0) \subset I$ mit $t_0 \in I_{max}$ und eine eindeutige Funktion $y_{max} \in C^1(I_{max}; G)$ mit folgenden Eigenschaften:

- (Eindeutige Lösbarkeit) Das Anfangswertproblem

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I_{max}) \\ y(t_0) = y_0 \end{cases} \quad (503)$$

wird durch $y = y_{max}$ gelöst.

- (Maximalität) Falls $J \subset I$ ein offenes Intervall mit $t_0 \in J$ ist, sodass das AWP

$$\begin{cases} z'(t) = f(t, z(t)) & (t \in J) \\ z(t_0) = y_0 \end{cases} \quad (504)$$

eine Lösung $z \in C^1(J; G)$ besitzt, dann gilt $J \subset I_{max}$ und $z(t) = y_{max}(t)$ für alle $t \in J$.

Beweis. Später.

Definition 95. MAXIMALE LÖSUNG. Die eindeutige Lösung $y_{max} \in C^1(I_{max}; \mathbb{R}^n)$ aus dem vorigen Satz nennen wir die *maximale Lösung* des Anfangswertproblems

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0. \end{cases} \quad (505)$$

Zusammengefasst: Falls f lokal Lipschitz ist, so existiert zu jedem Anfangswertproblem eine *eindeutige maximale Lösung*. “Maximal” ist hierbei in folgendem Sinne gemeint: Jede weitere Lösung ist eine Einschränkung der maximalen Lösung auf ein Teilintervall.

Dieser Satz ist mehr als zufriedenstellend. Wir finden zu jedem gegebenen Anfangswertproblem stets eine *bestmögliche* Lösung $y_{max} \in C^1(I_{max}; G)$. Jedoch gibt es noch ein Problem. Nehmen wir mal an, wir haben irgendeine Lösung einer DGL gefunden. Wie können wir dann sicherstellen, ob die gefundene Lösung tatsächlich die maximale Lösung ist?

Mit anderen Worten fragen wir uns: Was muss schiefgehen, damit eine Lösung sich nicht weiter vergrößern lässt?

Der nächste Satz zeigt, dass an einem festen Randpunkt des maximalen Existenzintervall $I_{max} \subset I$ nur drei Phänomene auftreten können

- (sog. *Escape-Situation*) Die Lösung erreicht den Rand des Definitionsbereiches G . Das bedeutet $\text{dist}(y(t), \mathbb{R}^n \setminus G) \rightarrow 0$, wenn t gegen den Randpunkt von I_{max} strebt. Falls $G = \mathbb{R}^n$ so kann dieses Phänomen nicht auftreten.
- (sog. *Blow-Up-Situation*) Die Lösung explodiert, d.h. $|y(t)| \rightarrow \infty$, wenn t gegen den Randpunkt von I_{max} strebt.
- (rechts/links-Globalität) Der Randpunkt von I_{max} ist ein Randpunkt von I . Dies nennen wir dann eine globale Lösung, weil die Lösung größtmöglich ist. Genauer: Wird der linke

Randpunkt von I erreicht, so nennen wir die Lösung *linksglobal* und wird der rechte Randpunkt erreicht, so nennen wir die Lösung *rechtsglobal*. Eine Lösung heißt dann global, falls sie *linksglobal* und *rechtsglobal* ist.

Dass nur diese drei Fälle auftreten können, wird in folgendem *Kriterium für die Maximalität* formuliert.

Satz 96. MAXIMALITÄTSKRITERIUM. Es seien $I = (a, b)$, $-\infty \leq a < b \leq \infty$, und $(t_0, y_0) \in I \times G$. Ferner erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und es sei $J = (t_-, t_+) \subset I$. Es sei $y \in C^1(J; G)$ eine Lösung von

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in J) \\ y(t_0) = y_0 \end{cases} \quad (506)$$

Dann sind folgende Aussagen äquivalent:

1. $J = I_{max}$ und $y = y_{max}$.
2. Es gilt $t_+ = b$ oder

$$\liminf_{t \uparrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (507)$$

Ferner gilt $t_- = a$ oder

$$\liminf_{t \downarrow t_-} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (508)$$

Hierbei verwenden wir für den Fall $G = \mathbb{R}^n$ die Konvention $\text{dist}(z, \emptyset) = \infty$ für alle $z \in \mathbb{R}^n$.

Wir halten noch einmal fest: Ist $G = \mathbb{R}^n$, so kann die Bedingung (507) vereinfacht werden. Da $\text{dist}(y(t), \mathbb{R}^n \setminus G) = \infty$ für alle $t \in I_{max} = (t_-, t_+)$ erhält man (sofern $t_+ < b$)

$$\liminf_{t \rightarrow t_+} \frac{1}{|y(t)| + 1} = 0, \quad (509)$$

oder, äquivalent (HA),

$$\limsup_{t \rightarrow t_+} |y(t)| = \infty. \quad (510)$$

Beweis. Später.

Dieser Satz kann auch dabei helfen, die *Globalität* von Lösungen zu zeigen. Wir nennen — wie vorher angekündigt — in diesem Kontext eine maximale Lösung *global*, falls $I_{max} = I$, d.h. sie existiert auf dem gesamten Intervall, auf dem die DGL aufgestellt worden ist. Können wir nun etwa das Phänomen in (507) ausschließen, so ist $\sup(I_{max}) = t_+ = b = \sup(I)$. Wir

werden später Mittel und Wege kennenlernen, um nur anhand der Gleichung das Phänomen in (507) auszuschließen. Ein wichtiges Mittel hierbei ist, das *Wachstum* von Lösungen zu verstehen. Dies geschieht zumeist mit dem *Lemma von Gronwall*, welches wir detailliert besprechen werden.

Am Ende des Kapitels wollen wir die *stetige Abhängigkeit vom Anfangswert* diskutieren. Das heißt wir zeigen, dass sich die Lösung (zumindest kurzzeitig) nicht viel verändert, wenn wir den Anfangswert nur wenig ändern.

Nun haben wir bereits die Agenda für dieses Kapitel festgelegt. Wir müssen die obigen Sätze beweisen. Bevor wir allerdings zur Tat schreiten können sollten wir eine Sache diskutieren: Alles fußt auf der (*lokalen*) *Lipschitz-Bedingung*, siehe Definition 93 (p. 91). Es wäre also schon gut zu wissen, ob es überhaupt Funktionen gibt, die diese Bedingung erfüllen! Zum Abschluss dieses Abschnittes sehen wir daher, wie man einer großen Klasse von Funktionen sofort ansehen kann, dass sie die lokale Lipschitz-Bedingung erfüllen.

Proposition 97. DIFFERENZIERBARKEIT UND LIPSCHITZBEDINGUNG. Es sei $f \in C^0(I \times G; \mathbb{R}^n)$ so, dass für alle $t \in I$ die Abbildung $f_t : G \rightarrow \mathbb{R}^n$, $f_t(z) := f(t, z)$ auf G differenzierbar ist. Ferner sei die Abbildung $(t, z) \mapsto D_z f_t(z) \in \mathbb{R}^{n \times n}$ auf $I \times G$ stetig. Dann erfüllt f eine lokale Lipschitz-Bedingung.

In Kurznotation bezeichnet man für festes $t \in I$ die Funktion f_t auch oft mit $f(t, \cdot) : G \rightarrow \mathbb{R}^n$. Das (\cdot) stellt man sich dann als *Platzhalter* vor, in den etwas eingesetzt werden kann. Man fordert also hier, dass für alle $t \in I$ die Funktion $f(t, \cdot) : G \rightarrow \mathbb{R}^n$ auf G differenzierbar sein soll und die Abbildung $(t, z) \mapsto D_z f(t, z)$ stetig auf $I \times G$ ist.

Beweis. Es sei $(t_0, y_0) \in I \times G$. Wähle nun $\epsilon > 0$ mit $[t_0 - \epsilon, t_0 + \epsilon] \subset I$ und $r > 0$ mit $\overline{B_r(y_0)} \subset G$. Da $[t_0 - \epsilon, t_0 + \epsilon] \times \overline{B_r(y_0)} \subset I \times G$ kompakt ist und $z \mapsto D_z f(t, z) = (\partial_{z_j} f_i(t, z))_{i,j=1,\dots,n}$ in jeder Komponente stetig ist, gilt (nach dem Satz von Weierstraß)

$$\tilde{L} := \max_{i,j=1,\dots,n} \max_{(t,z) \in [t_0 - \epsilon, t_0 + \epsilon] \times \overline{B_r(y_0)}} |\partial_{z_j} f_i(t, z)| < \infty. \quad (511)$$

Wir behaupten nun, dass für alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$ und alle $z_1, z_2 \in \overline{B_r(y_0)}$ gilt, dass

$$|f(t, z_1) - f(t, z_2)| \leq n \tilde{L} |z_1 - z_2|. \quad (512)$$

Seien dazu $t \in [t_0 - \epsilon, t_0 + \epsilon]$, $z_1, z_2 \in \overline{B_r(y_0)}$ wie oben fest gewählt. Definiere für $i \in \{1, \dots, n\}$ und $\lambda \in [0, 1]$

$$h_i(\lambda) := f_i(t, z_2 + \lambda(z_1 - z_2)). \quad (513)$$

Man beachte: Da $\overline{B_r(y_0)}$ eine konvexe Menge ist, gilt $z_2 + \lambda(z_1 - z_2) \in \overline{B_r(y_0)}$ für alle $\lambda \in [0, 1]$. Ferner ist h_i wegen der Kettenregel und der Differenzierbarkeitsvoraussetzung an f auf $(0, 1)$ stetig differenzierbar. Des Weiteren gilt $h_i(1) = f_i(t, z_1)$ und $h_i(0) = f_i(t, z_2)$. Wir erhalten nun mit dem Hauptsatz der Differential- und Integralrechnung und der

Kettenregel der mehrdimensionalen Differenzierens

$$|f_i(t, z_1) - f_i(t, z_2)| = |h_i(1) - h_i(0)| = \left| \int_0^1 h'_i(\lambda) d\lambda \right| \quad (514)$$

$$= \left| \int_0^1 D_z f_i(t, z_2 + \lambda(z_1 - z_2))(z_1 - z_2) d\lambda \right| \quad (515)$$

$$= \left| \int_0^1 \langle \text{grad}_z f_i(t, z_2 + \lambda(z_1 - z_2)), z_1 - z_2 \rangle d\lambda \right| \quad (516)$$

$$\leq \int_0^1 |\langle \text{grad}_z f_i(t, z_2 + \lambda(z_1 - z_2)), z_1 - z_2 \rangle| d\lambda \quad (517)$$

$$\stackrel{C.S.U.}{\leq} \int_0^1 |\text{grad}_z f_i(z_1 + \lambda(z_2 - z_1))| |z_1 - z_2| d\lambda \quad (518)$$

$$\leq \int_0^1 |\text{grad}_z f_i(z_1 + \lambda(z_2 - z_1))| d\lambda |z_1 - z_2|. \quad (519)$$

Hierbei steht $\text{grad}_z f_i(t, z)$ für den Gradienten der Funktion $f_i(t, \cdot)$ für fixiertes t . Die Abkürzung $C.S.U.$ meint hier die ‘Cauchy-Schwarz-Ungleichung’. Nun berechnen wir

$$|\text{grad}_z f_i(z_1 + \lambda(z_2 - z_1))| = \sqrt{\sum_{j=1}^n (\partial_{z_j} f_i(z_1 + \lambda(z_2 - z_1)))^2} \leq \sqrt{n \tilde{L}^2} = \sqrt{n} \tilde{L}. \quad (520)$$

Daher gilt für alle $i \in \{1, \dots, n\}$

$$|f_i(t, z_1) - f_i(t, z_2)| \leq \sqrt{n} \tilde{L} |z_1 - z_2|. \quad (521)$$

Wir berechnen nun

$$|f(t, z_1) - f(t, z_2)| \leq \sqrt{\sum_{i=1}^n |f_i(t, z_1) - f_i(t, z_2)|^2} \quad (522)$$

$$\leq \sqrt{n(\sqrt{n} \tilde{L} |z_1 - z_2|^2)} \leq n \tilde{L} |z_1 - z_2|. \quad (523)$$

Damit ist (512) gezeigt. Wir haben also eine Lipschitz-Abschätzung mit Lipschitzkonstante $L(t_0, y_0) := n \tilde{L}$ erreicht.

Viele Differentialgleichungen sind nicht explizit zeitabhängig, d.h. $y'(t) = f(y(t))$ für ein $f \in C^0(G)$. Der obige Satz zeigt, dass in dem Fall $f \in C^1(G)$ ein hinreichendes Kriterium für die Existenz einer eindeutigen maximalen Lösung ist. Es ist insgesamt erfreulich, dass man für die Konstruktion von Lösungen also nur eine Differenzierbarkeitsvoraussetzung an f stellen muss.

Übung 98. LOKALE UND GLOBALE LIPSCHITZ-BEDINGUNG.

- (a) Geben Sie ein Beispiel für eine Funktion $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$, die eine lokale Lipschitzbedingung erfüllt, jedoch nicht einer globalen Lipschitzbedingung genügt.
- (b) Geben Sie ein Beispiel für eine beschränkte(!) Funktion $f \in C^0(\mathbb{R} \times \mathbb{R}^n; \mathbb{R}^n)$, die eine lokale Lipschitzbedingung erfüllt, jedoch nicht einer globalen Lipschitzbedingung genügt.

SINN DER ÜBUNG. Wie wir hier sehen erfüllt (dankenswerterweise) eine wirklich große Klasse von Funktionen eine lokale Lipschitz-Bedingung. Dies ist erfreulich, da wir mit den obigen Sätzen also eine wirklich große Klasse von Differentialgleichungen studieren können. Unser Beispiel

$$\begin{cases} y'(t) = 2\sqrt{|y(t)|} & (t \in \mathbb{R}) \\ y(0) = 0 \end{cases} \quad (524)$$

ist allerdings noch nicht abgehandelt — die Funktion $f(t, z) := 2\sqrt{|z|}$ erfüllt auf $\mathbb{R} \times \mathbb{R}$ nicht einmal eine lokale Lipschitz-Bedingung! Dies kann man zum Beispiel so sehen: Wenn sie eine erfüllen würde, so hätte das AWP (524) eine eindeutige maximale Lösung — Hat es aber nicht! (siehe Beispiel 4 (p. 4)).

Übung 99. EXISTENZ EINER SIR-PANDEMIE.

Es sei wie in (410)

$$Z := \{(s, i, r) \in (0, 1)^3 : s + i + r = 1\}. \quad (525)$$

Zeigen Sie: Für alle $y_0 \in Z$ existiert eine SIR-Pandemie $y \in C^1(I; \mathbb{R}^3)$ für $I \supset [0, \infty)$ mit $y(0) = y_0$. [Hierbei ist der Begriff einer SIR-Pandemie im Sinne von Definition 73 (p. 73) zu verstehen.]

HINWEIS. Es sind zwei Dinge zu zeigen: Existenz einer maximalen Lösung für die Pandemiegleichung (405) mit Anfangswert y_0 und die Rechtsglobalität der Lösung, damit $I \supset [0, \infty)$ sichergestellt ist. Für die Existenz ist nur eine lokale Lipschitz-Bedingung nachzuprüfen, siehe 94 (p. 93). Für die Rechtsglobalität muss in Satz 96 (p. 93) die Escape-Situation und die Blow-Up-Situation ausgeschlossen werden. Erinnern sie sich hierfür zurück an Satz 76 (p. 74).

SINN DER ÜBUNG. Endlich wissen wir, dass die Gleichung die wir in Kapitel 2 studiert haben, auch wirklich eine Lösung besitzt, die sich in einem physikalisch sinnvollen Bereich bewegt.

Übung 100. GLOBALITÄT DANK ENERGIEERHALTUNG.

Gegeben sei das Anfangswertproblem

$$\begin{cases} y''(t) = -\nabla U(y(t)) & (t \in \mathbb{R}) \\ y(0) = y_0 \\ y'(0) = y_1 \end{cases}$$

für ein $U \in C^2(\mathbb{R}^n; \mathbb{R})$ und $y_0, y_1 \in \mathbb{R}^n$. Zeigen Sie: Falls

$$\lim_{|x| \rightarrow \infty} U(x) = \infty \quad (526)$$

so existiert eine eindeutige globale Lösung $y \in C^1(\mathbb{R}; \mathbb{R}^n)$.

HINWEIS. Verwenden Sie die Energieerhaltung um die Blow-Up-Situation auszuschließen.
SINN DER ÜBUNG. Dies zeigt, dass *sehr viele* physikalische Phänomene eine globale Lösung haben! Die Bedingung (526) ist für die meisten physikalischen Systeme erfüllt. Der Grund hierfür ist die Energieerhaltung. Dies zeigt noch einmal, was für ein wichtiges Konzept die Energieerhaltung doch ist.

3.1.1 Das assoziierte Fixpunktproblem

Es sei für diesen Abschnitt $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$. Für den Beweis werden wir das Anfangswertproblem

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (527)$$

äquivalent mit einer Integralgleichung formulieren, nämlich durch

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad (t \in I). \quad (528)$$

Falls $n > 1$, so ist das Integral in der obigen Gleichung komponentenweise zu verstehen. Diese Integralgleichung erhalten wir leicht aus (527), indem wir in (527) von t_0 bis t integrieren. Nun definieren wir für $t \in I$

$$(Ty)(t) := y_0 + \int_{t_0}^t f(s, y(s)) \, ds. \quad (529)$$

Die Integralgleichung (528) wird dadurch zu der Gleichung

$$y(t) = (Ty)(t) \quad \forall t \in I \quad (530)$$

beziehungsweise – kurz geschrieben –

$$y = Ty. \quad (531)$$

Diese Gleichung sieht nun aus wie eine *Fixpunktgleichung*, die man etwa aus den Analysis Grundlagenvorlesungen kennt. Bei einem *Fixpunktproblem* löst man Gleichungen der Form $Tx = x$, wobei $T : X \rightarrow X$ eine Abbildung auf einem metrischen Raum (X, d) ist. Das größte Problem ist oft, den richtigen metrischen Raum X zu finden, sodass das Problem eine Lösung besitzt. An dieser Stelle haben wir noch gar nicht besprochen, welchen Raum X wir hier zugrunde legen wollen. Das soll sich erst im Laufe der Zeit ergeben.

Wir werden uns den metrischen Raum X gut überlegen müssen, denn der *Banach'sche Fixpunktsatz*, den wir benutzen, um den Fixpunkt zu konstruieren benötigt hier einige Voraussetzungen. Zur Wiederholung:

Satz 101. BANACH'SCHER FIXPUNKTSATZ. Es sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine strikte Kontraktionsabbildung, d.h. es gibt ein $q \in [0, 1)$ sodass $d(Tx, Ty) \leq qd(x, y)$ für alle $x, y \in X$. Dann hat T genau einen Fixpunkt, d.h. es gibt genau ein $x_* \in X$ mit $Tx_* = x_*$.

Beweis. Siehe Analysis-Grundlagenvorlesung (HA). Vielleicht ein kurzes Wort zur Beweisstrategie: Man gibt sich $x_0 \in X$ beliebig vor und betrachtet die *rekursiv definierte* Folge

$$x_{n+1} := Tx_n \quad (n \in \mathbb{N}_0). \quad (532)$$

Sprich: Man *iteriert* T . Man zeigt dann, dass diese Folge gegen einen Fixpunkt konvergiert.

Der Beweis ist damit in **drei Zutaten** unterteilt

1. Äquivalente Integralgleichung (528).
2. Fixpunktformulierung (531).
3. Banach'scher Fixpunktsatz.

Im Folgenden besprechen wir diese Zutaten im Detail, sodass wir sie am Ende nur noch zusammenwerfen müssen.

Zu Zutat 3, “Banach'scher Fixpunktsatz”.

Wir benutzen sogar eine Verschärfung des Banach'schen Fixpunktsatzes – es reicht nämlich als Voraussetzung aus, wenn eine beliebige *Potenz* von T eine strikte Kontraktionsabbildung ist. Mit einer Potenz von T meinen wir eine Abbildung der Form

$$T^k := \underbrace{T \circ T \circ \dots \circ T}_{k-\text{mal}} \quad (533)$$

für ein $k \in \mathbb{N}$.

Satz 102. BANACH'SCHER FIXPUNKTSATZ — VERSCHÄRFUNG. Es sei (X, d) ein vollständiger metrischer Raum und $T : X \rightarrow X$ eine Abbildung mit folgender Eigenschaft: Es gibt ein $n_0 \in \mathbb{N}$ sodass T^{n_0} eine strikte Kontraktionsabbildung ist (siehe Satz 101, p. 98). Dann hat T genau einen Fixpunkt, d.h. es gibt genau ein $x_* \in X$ mit $Tx_* = x_*$.

Beweis. Schritt 1: Existenz des Fixpunktes. Da T^{n_0} eine strikte Kontraktionsabbildung ist, gibt es nach dem Banach'schen Fixpunktsatz ein eindeutiges $\tilde{x} \in X$ mit

$$T^{n_0}\tilde{x} = \tilde{x}. \quad (534)$$

Wir behaupten nun, dass sogar $T\tilde{x} = \tilde{x}$. Dazu werden wir Folgendes zeigen: Sowohl \tilde{x} als auch $T\tilde{x}$ sind Fixpunkte von T^{n_0} . Aus der Eindeutigkeit des Fixpunktes von T^{n_0} wird dann eben $T\tilde{x} = \tilde{x}$ folgen.

Nun wissen wir bereits aus (534), dass \tilde{x} ein Fixpunkt von T^{n_0} ist. Dass auch $T\tilde{x}$ ein

Fixpunkt von T^{n_0} ist, ergibt sich aus der Rechnung

$$T^{n_0}(T\tilde{x}) = T^{n_0+1}\tilde{x} = T(T^{n_0}\tilde{x}) \stackrel{(534)}{=} T\tilde{x}. \quad (535)$$

Damit folgern wir auch, dass $T\tilde{x}$ ein Fixpunkt von T^{n_0} ist und sehen (wie oben angesprochen) $T\tilde{x} = \tilde{x}$.

Schritt 2: Eindeutigkeit des Fixpunktes Es seien $x, y \in X$ zwei Fixpunkte von T , d.h. $Tx = x$ und $Ty = y$. Dann gilt auch

$$T^{n_0}x = T^{n_0-1}(Tx) = T^{n_0-1}x = \dots = Tx = x \quad (536)$$

und analog

$$T^{n_0}y = y. \quad (537)$$

Insbesondere sind x und y beides Fixpunkte von T^{n_0} . Da T^{n_0} jedoch nur genau einen Fixpunkt haben kann, folgt $x = y$.

Nun ist das fundamentale Werkzeug — nämlich der Banach'sche Fixpunktsatz — vollends diskutiert, aber es bleibt immer noch die Wahl des Raumes X . Um zu verstehen, auf welchem Raum die Abbildung T aus (529) wirklich operiert müssen wir genau verstehen, wie die Äquivalenz von dem Anfangswertproblem und der Integralgleichung (528) wirklich zu verstehen ist. Hierzu widmen wir uns zunächst der Integralgleichung.

Zu Zutat 1, “Äquivalente Integralgleichung”.

Proposition 103. Es sei $I \subset \mathbb{R}$ und $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$. Sei dazu $y \in C^0(I; \mathbb{R}^n)$ (nur stetig!). Dann sind folgende Aussagen äquivalent

1. $y \in C^1(I; \mathbb{R}^n)$ und y löst das AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0. \end{cases} \quad (538)$$

2. $y \in C^0(I; \mathbb{R}^n)$ erfüllt die Integralgleichung

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad \forall t \in I. \quad (539)$$

Beweis. Schritt 1: “1. \Rightarrow 2.” Es sei $y \in C^1(I; \mathbb{R}^n)$ so, dass (538) erfüllt ist. Dann gilt nach dem Hauptsatz der Differential -und Integralrechnung

$$y(t) - y_0 \stackrel{(538)}{=} y(t) - y(t_0) = \int_{t_0}^t y'(s) \, ds \stackrel{(538)}{=} \int_{t_0}^t f(s, y(s)) \, ds. \quad (540)$$

Addiert man auf beiden Seiten y_0 , so folgt die Behauptung.

Schritt 2: “2. \Rightarrow 1.” Es sei $y \in C^0(I; \mathbb{R}^n)$ so, dass

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad \forall t \in I \quad (541)$$

Wir müssen zeigen, dass y eine Lösung des AWP ist. Damit das überhaupt denkbar ist, muss zuerst gezeigt werden, dass $y \in C^1(I; \mathbb{R}^n)$ (denn andernfalls könnte man y überhaupt nicht differenzieren). Hierzu erinnern wir uns an folgende Aussage aus dem Hauptsatz der Differential- und Integralrechnung: Ist $g \in C^0(I; \mathbb{R}^n)$ und $t_0 \in I$ beliebig, so ist

$$G(t) := \int_{t_0}^t g(s) \, ds \quad (542)$$

differenzierbar mit $G' = g$. Da g als stetig vorausgesetzt war, gilt sogar, dass G stetig differenzierbar ist, d.h. in $C^1(I; \mathbb{R}^n)$ liegt. Wir wollen nun die Aussage auf $g(s) := f(s, y(s))$ anwenden. Dazu muss gezeigt werden, dass diese Wahl von g stetig auf I ist. Da f stetig auf $I \times \mathbb{R}^n$ ist und nach Voraussetzung die Abbildung $s \mapsto (s, y(s))$ von I nach $I \times \mathbb{R}^n$ stetig ist, ist g tatsächlich als Verkettung stetiger Abbildungen stetig. Wir erhalten daher aus der vorigen Überlegung, dass $G : I \rightarrow \mathbb{R}^n$ definiert durch

$$G(t) := \int_{t_0}^t f(s, y(s)) \, ds \quad (t \in I) \quad (543)$$

stetig differenzierbar ist mit $G'(t) = f(t, y(t))$ für alle $t \in I$. Aus (541) folgt, dass $y(t) = y_0 + G(t)$ für alle $t \in I$. Damit ist auch y stetig differenzierbar auf I mit

$$y'(t) = G'(t) = f(t, y(t)) \quad \forall t \in I. \quad (544)$$

Außerdem gilt nach (541)

$$y(t_0) = y_0 + \int_{t_0}^{t_0} f(s, y(s)) \, ds = y_0. \quad (545)$$

Mit den vorigen beiden Gleichungen erhalten wir, dass das Anfangswertproblem (538) erfüllt ist.

Eine wichtige Erkenntnis aus dieser Proposition: Jede stetige Lösung der Integralgleichung ist automatisch stetig differenzierbar. Das heißt insbesondere, dass wir das gesamte Problem in einem Raum von stetigen Funktionen formulieren können.

Zu Zutat 2, “Fixpunktformulierung”.

Die Abbildungsvorschrift für die Abbildung T , die wir studieren möchten, haben wir bereits in (529) gesehen. Wir müssen aber noch einen vollständigen(!) metrischen Raum finden, auf dem die Abbildung definiert ist, sonst ist der Banach'sche Fixpunktsatz nicht anwendbar. Das Finden des richtigen Raumes ist in der Analysis oft ein sehr delikates Problem.

Definition 104. DER RAUM $C_b^0(I; \mathbb{R}^n)$. Es sei $I \subset \mathbb{R}^n$ offen. Wir definieren Menge

$$C_b^0(I; \mathbb{R}^n) := \{f \in C^0(I; \mathbb{R}^n) : \sup_{x \in I} |f(x)| < \infty\} \quad (546)$$

und die Abbildung $d_\infty : C_b^0(I; \mathbb{R}^n) \times C_b^0(I; \mathbb{R}^n) \rightarrow [0, \infty)$ durch

$$d_\infty(f, g) := \sup_{x \in I} |f(x) - g(x)|. \quad (547)$$

An dieser Stelle müsste man eigentlich noch anmerken, dass $d_\infty(f, g)$ wohldefiniert ist, d.h. $d_\infty(f, g) < \infty$ für alle $f, g \in C_b^0(I; \mathbb{R}^n)$. Dies sieht man aber sofort mit der Beobachtung, dass (HA)

$$\sup_{x \in I} |f(x) - g(x)| \leq \sup_{x \in I} (|f(x)| + |g(x)|) \leq \sup_{x \in I} |f(x)| + \sup_{x \in I} |g(x)|. \quad (548)$$

Proposition 105. Das Paar $(C_b^0(I; \mathbb{R}^n), d_\infty)$ definiert einen vollständigen metrischen Raum.

Beweis. Schritt 1. $C_b^0(I; \mathbb{R}^n)$ ist ein metrischer Raum. Die Symmetrie und die positive Definitheit von d_∞ ist den Lesenden als Übungsaufgabe überlassen. Zur Dreiecksungleichung es seien $f, g, h \in C^0(I; \mathbb{R}^n)$.

$$\mathbf{Z} : \quad d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h). \quad (549)$$

Dazu sei nun $x \in I$. Dann gilt

$$|f(x) - h(x)| = |(f(x) - g(x)) + (g(x) - h(x))| \leq |f(x) - g(x)| + |g(x) - h(x)|. \quad (550)$$

Nun können wir abschätzen

$$|f(x) - g(x)| \leq d_\infty(f, g), \quad |g(x) - h(x)| \leq d_\infty(g, h), \quad (551)$$

da d_∞ ja stets das Supremum der betragsmäßigen Differenzen ist. Wir folgern

$$|f(x) - h(x)| \leq d_\infty(f, g) + d_\infty(g, h) \quad \forall x \in I. \quad (552)$$

Nehmen wir das Supremum über alle $x \in I$, so erhalten wir

$$d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h). \quad (553)$$

Schritt 2: Nun zur Vollständigkeit. Es sei $(f_j)_{j=1}^\infty \subset C_b^0(I; \mathbb{R}^n)$ eine Cauchy-Folge, d.h. für alle $\epsilon > 0$ gibt es ein $n_0(\epsilon) \in \mathbb{N}$ mit der Eigenschaft, dass $d_\infty(f_j, f_m) < \epsilon$ für alle $j, m \geq n_0(\epsilon)$. Wir müssen zeigen, dass $(f_j)_{j=1}^\infty$ konvergiert, d.h. es gibt $f \in C_b^0(I; \mathbb{R}^n)$ mit $d_\infty(f_j, f) \rightarrow 0$ für $j \rightarrow \infty$.

Schritt 2.1. Zeige: Für jedes $x \in I$ ist $(f_j(x))_{j=1}^\infty \subset \mathbb{R}^n$ eine konvergente Folge in \mathbb{R}^n . Dazu

fixieren wir $x \in I$. Sei $\epsilon > 0$ beliebig vorgegeben. Wähle $n_0(\epsilon) \in \mathbb{N}$ so, dass $d_\infty(f_j, f_m) < \epsilon$ für alle $m, j \geq n_0(\epsilon)$. Nun gilt

$$|f_j(x) - f_m(x)| \stackrel{(547)}{\leq} d_\infty(f_j, f_m) < \epsilon \quad \forall j, m \geq n_0(\epsilon) \quad (554)$$

Da $\epsilon > 0$ beliebig war, folgt, dass $(f_j(x))_{j=1}^\infty \subset \mathbb{R}^n$ eine Cauchy-Folge in \mathbb{R}^n definiert. Da \mathbb{R}^n vollständig ist, konvergiert $(f_j(x))_{j=1}^\infty$ gegen einen Grenzwert in \mathbb{R}^n . Wir haben also die Zwischenbehauptung in Schritt 2.1 gezeigt. Für den weiteren Verlauf des Beweises definieren wir die Abbildung $f : I \rightarrow \mathbb{R}^n$ durch

$$f(x) := \lim_{j \rightarrow \infty} f_j(x), \quad (555)$$

die *punktweise Grenzfunktion* von $(f_j)_{j=1}^\infty$.

Schritt 2.2. Zeige $f \in C_b^0(I; \mathbb{R}^n)$, d.h. f ist stetig und beschränkt. Zur Stetigkeit: Wir wissen aus der Analysis, dass der gleichmäßige Grenzwert stetiger Funktionen wieder eine stetige Funktion ist. Wir zeigen daher, dass $(f_j)_{j=1}^\infty$ auf I gleichmäßig gegen f konvergiert. Dazu sei $\epsilon > 0$. Dann gibt es wiederum $n_0(\epsilon) \in \mathbb{N}$ mit $d_\infty(f_j, f_m) < \epsilon$ für alle $j, m \geq n_0(\epsilon)$. Für $j \geq n_0(\epsilon)$ und $x \in I$ gilt dann

$$|f_j(x) - f(x)| = \lim_{m \rightarrow \infty, m \geq n_0(\epsilon)} |f_j(x) - f_m(x)| \leq \lim_{m \rightarrow \infty, m \geq n_0(\epsilon)} d_\infty(f_j, f_m) \leq \epsilon. \quad (556)$$

Da $x \in I$ beliebig war, folgt

$$\sup_{x \in I} |f_j(x) - f(x)| \leq \epsilon \quad \forall j \geq n_0(\epsilon), \quad (557)$$

was die gleichmäßige Konvergenz impliziert. Die Stetigkeit folgt. Zur Beschränktheit: Aus der vorigen Überlegung sehen wir: Es gibt $N := n_0(1) \in \mathbb{N}$ mit $d_\infty(f_j, f_N) < 1$ für alle $j, m \geq N$. Aus (557) folgt dann auch

$\sup_{x \in I} |f_j(x) - f(x)| < 1$ für alle $j \geq N$. Damit folgert man

$$\begin{aligned} \sup_{x \in I} |f(x)| &= \sup_{x \in I} |f(x) - f_N(x) + f_N(x)| \leq \sup_{x \in I} (|f(x) - f_N(x)| + |f_N(x)|) \\ &\leq \sup_{x \in I} (|f(x) - f_N(x)| + \sup_{y \in I} |f_N(y)|) \\ &\leq \sup_{x \in I} |f(x) - f_N(x)| + \sup_{y \in I} |f_N(y)| \\ &\stackrel{(557), N=n_0(1)}{\leq} 1 + \sup_{y \in I} |f_N(y)| < \infty, \end{aligned}$$

wobei wir im letzten Schritt benutzt haben, dass $f_N \in C_b^0(I; \mathbb{R}^n)$.

Schritt 2.3. $d_\infty(f_j, f) \rightarrow 0$ für $j \rightarrow \infty$. Dies folgt bereits direkt aus (557), denn für $\epsilon > 0$ gilt

$$d_\infty(f_j, f) = \sup_{x \in I} |f_j(x) - f(x)| \stackrel{(557)}{\leq} \epsilon \quad \forall j \geq n_0(\epsilon). \quad (558)$$

Übung 106. (UN)VOLLSTÄNDIGKEIT VON $(C_b^1(I; \mathbb{R}^n), d_\infty)$. Wir definieren für $n \in \mathbb{N}$ die Menge

$$C_b^1(I; \mathbb{R}^n) := \{f \in C_b^0(I; \mathbb{R}^n) : f \text{ stetig differenzierbar auf } I\}. \quad (559)$$

Zeigen Sie, dass für $n = 1$ und $I = (-1, 1)$ das Paar $(C_b^1(I; \mathbb{R}^n), d_\infty)$ einen metrischen Raum definiert, der nicht vollständig ist.

SINN DER ÜBUNG: Am liebsten hätten wir das Fixpunktproblem (529) auf einer Menge stetig differenzierbarer Funktionen betrachtet, denn dann wäre auch sofort klar, dass die Lösung des Fixpunktproblems stetig diffbar ist.

HINWEIS: Zeigen Sie, dass die Folge gegeben durch

$$f_n(x) := \sqrt{x^2 + \frac{1}{n}} \quad (x \in (-1, 1)) \quad (560)$$

für $n \in \mathbb{N}$ eine nicht-konvergente Cauchy-Folge ist.

3.1.2 Beweis der globalen Version

Wir wollen nun mit den obigen Zutaten die Existenz einer Lösung zeigen. Es sei $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$, d.h. zunächst $G = \mathbb{R}^n$. Es sei die Differentialgleichung

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (561)$$

gegeben. Wir wollen diese auf ganz I lösen. Allerdings werden wir für die Lösungstechnik I verkleinern müssen. Wir sagen, dass ein offenes Intervall J kompakt in I enthalten ist, falls J beschränkt ist und $\overline{J} \subset I$. Wir schreiben in diesem Fall $J \subset\subset I$. Warum wir das Intervall verkleinern müssen sehen wir später. Für $J \subset\subset I$ mit $t_0 \in J$ definieren wir die Abbildung

$$T : C_b^0(J; \mathbb{R}^n) \rightarrow C_b^0(J; \mathbb{R}^n), y \mapsto Ty \quad (562)$$

wobei $Ty \in C_b^0(I; \mathbb{R}^n)$ durch die Funktionsvorschrift

$$(Ty)(t) := y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad (t \in J). \quad (563)$$

Wir müssen nachprüfen, ob diese Abbildung wohldefiniert ist. Dazu gehört insbesondere, dass Ty tatsächlich eine Funktion in $C_b(J; \mathbb{R}^n)$ definiert.

Übung 107. Es sei $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ und $J \subset\subset I$ mit $t_0 \in J$. Zeigen Sie: Für alle $y \in C_b^0(J; \mathbb{R}^n)$ ist die Abbildung $T_y : J \rightarrow \mathbb{R}^n$ gegeben durch

$$T_y(t) := y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad (t \in J) \quad (564)$$

ein Element von $C_b^0(J; \mathbb{R}^n)$. Gilt dies auch, wenn man die Voraussetzung " $J \subset\subset I$ " durch " $J \subseteq I$ " ersetzt?

SINN DER ÜBUNG: In dieser Übung lernt man, dass es notwendig ist, das Definitionssintervall zu verkleinern. Dies ist nämlich notwendig damit die in (563) definierte Abbildung eine *Selbstabbildung* ist, d.h. ihr Bild landet wieder im selben Raum.

Wenn wir unser Intervall I verkleinern müssen, stellt sich natürlich die Frage, wie wir am Ende auf das ursprüngliche Intervall zurückkommen. Die Technik dazu heißt *Ausschöpfung* und beruht auf folgender Idee

Beobachtung 108. AUSSCHÖPFUNG. Für jedes offene Intervall $I \subset \mathbb{R}^n$ gibt es eine Folge $(J_k)_{k \in \mathbb{N}}$ von offenen Intervallen mit $J_k \subset\subset I$, $J_1 \subset J_2 \subset J_3 \subset \dots$ und

$$I = \bigcup_{k=1}^{\infty} J_k. \quad (565)$$

BEGRÜNDUNG: Falls $I = (a, b)$ für $-\infty < a < b < \infty$, so kann man $J_k := (a + \frac{1}{k}, b - \frac{1}{k})$ wählen (HA). Falls $I = (a, \infty)$ für ein $a > -\infty$, so tut's $J_k := (a + \frac{1}{k}, \infty)$. Alle anderen Fälle überlassen wir den Lesenden als Übungsaufgabe.

Wir prüfen nun nach, dass wir für beschränkte Intervalle nach einem gewissen Index eine Kontraktionsabbildung vorliegen haben

Lemma 109. KONTRAKTIONSEIGENSCHAFT. Es erfülle $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung mit Lipschitzkonstante $L > 0$. Es sei $J = (a, b) \subset\subset I$ offen und beschränkt mit $-\infty < a < b < \infty$ und $t_0 \in J$. Dann gilt für alle $j \in \mathbb{N}$

$$d_{\infty}(T^j y, T^j z) \leq L^j \frac{(b-a)^j}{j!} d_{\infty}(y, z) \quad \forall y, z \in C_b^0(J; \mathbb{R}^n). \quad (566)$$

Insbesondere gibt es ein $n_0 \in \mathbb{N}$ sodass T^{n_0} eine strikte Kontraktionsabbildung ist und einen eindeutigen Fixpunkt besitzt.

Beweis. In diesem Beweis wollen wir nicht nur stumpf die Gleichung beweisen, sondern auch klären, wie man auf diese Gleichung kommt. Die ausschlaggebende Beobachtung ist hier die Folgende. Wir haben

$$(Ty)(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad \forall t \in J, \quad (567)$$

$$(Tz)(t) = z_0 + \int_{t_0}^t f(s, z(s)) \, ds \quad \forall t \in J. \quad (568)$$

Damit gilt für alle $t \in J$

$$|(Ty)(t) - (Tz)(t)| = \left| y_0 + \int_{t_0}^t f(s, y(s)) \, ds - \left(y_0 + \int_{t_0}^t f(s, z(s)) \, ds \right) \right| \quad (569)$$

$$\leq \left| \int_{t_0}^t [f(s, y(s)) - f(s, z(s))] \, ds \right| \quad (570)$$

$$\leq \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |f(s, y(s)) - f(s, z(s))| \, ds \quad (571)$$

$$\leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |y(s) - z(s)| \, ds \quad (572)$$

Die Gleichung

$$|(Ty)(t) - (Tz)(t)| \leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |y(s) - z(s)| \, ds \quad \forall t \in J \quad (573)$$

werden wir im Laufe dieses Beweises sehr häufig benutzen. Zunächst können wir damit den Fall $n = 1$ klären. Wir wissen nämlich, dass $|y(s) - z(s)| \leq d_\infty(y, z)$ für alle $s \in J$, da ja d_∞ gerade das Supremum über alls $s \in J$ ist, siehe Definition 104 (p. 101). Deswegen gilt nach (573) für alle $t \in J$

$$|(Ty)(t) - (Tz)(t)| \leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} d_\infty(y, z) \, ds \quad (574)$$

$$= L d_\infty(y, z) (\max\{t_0, t\} - \min\{t_0, t\}) = L d_\infty(y, z) |t - t_0|. \quad (575)$$

Daraufhin ergibt sich

$$|(Ty)(t) - (Tz)(t)| \leq L(b - a) d_\infty(y, z) \quad \forall t \in J. \quad (576)$$

Nehmen wir das Supremum über alle t , so erhalten wir mit Definition 104 (p. 101)

$$d_\infty(Ty, Tz) \leq L(b - a) d_\infty(y, z), \quad (577)$$

also gerade (566) im Fall $n = 1$. Für den Fall $n = 2$ gehen wir wie folgt vor. Aus (573) folgt

$$|(T^2y)(t) - (T^2z)(t)| \leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |Ty(s) - Tz(s)| \, ds \quad \forall t \in J. \quad (578)$$

Jetzt wissen wir aus (575), dass $|(Ty)(s) - (Tz)(s)| \leq L d_\infty(y, z) |t - t_0|$. Daher gilt

$$|(T^2y)(t) - (T^2z)(t)| \leq L^2 d_\infty(y, z) \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |s - t_0| \, ds \quad (579)$$

$$= L^2 d_\infty(y, z) \frac{|t - t_0|^2}{2}, \quad (580)$$

wobei wir die letzte Gleichung leicht mit einer Fallunterscheidung zwischen den Fällen “ $t \leq t_0$ ” und “ $t > t_0$ ” nachrechnen können (HA). Wir folgern daraus wiederum

$$|(T^2y)(t) - (T^2z)(t)| \leq L^2 \frac{(b-a)^2}{2} d_\infty(y, z) \quad \forall t \in J. \quad (581)$$

Nehmen wir wiederum das Supremum über alle $t \in J$, so erhalten wir wie in (566) gewünscht

$$d_\infty(T^2y, T^2z) \leq L^2 \frac{(b-a)^2}{2!} d_\infty(y, z). \quad (582)$$

Man sieht: So könnte man für beliebig große Potenzen von T weitermachen. Für allgemeine $n \in \mathbb{N}$ zeigen wir zunächst mit vollständiger Induktion die folgende

Zwischenbehauptung: Für alle $j \in \mathbb{N}$ hat man

$$|(T^jy)(t) - (T^jz)(t)| \leq L^j \frac{|t - t_0|^j}{j!} d_\infty(y, z) \quad \forall t \in J. \quad (583)$$

Für $n = 1, 2$ haben wir die Aussage bereits gezeigt. Es gelte nun die Aussage (583) für ein $j \in \mathbb{N}$. Mit (573) angewendet auf $\tilde{y} = T^jy$ und $\tilde{z} = T^jz$ erhalten wir

$$|(T^{j+1}y)(t) - (T^{j+1}z)(t)| = |(T\tilde{y})(t) - (T\tilde{z})(t)| \quad (584)$$

$$\leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |\tilde{y}(s) - \tilde{z}(s)| \, ds \quad (585)$$

$$= L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |(T^jy)(s) - (T^jz)(s)| \, ds \quad (586)$$

Benutzen wir nun die Induktionshypothese (583), so ergibt sich

$$|(T^{j+1}y)(t) - (T^{j+1}z)(t)| \leq L \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} L^j \frac{|s - t_0|^j}{j!} \, ds \quad (587)$$

$$= \frac{L^{j+1}}{j!} \int_{\min\{t_0, t\}}^{\max\{t_0, t\}} |s - t_0|^{j+1} \, ds \quad (588)$$

$$= L^{j+1} \frac{1}{j!} \frac{|t - t_0|^{j+1}}{j+1} = \frac{L^{j+1}|t - t_0|^{j+1}}{(j+1)!}. \quad (589)$$

Die Behauptung folgt per vollständiger Induktion. Nun gilt also für alle $j \in \mathbb{N}$, dass

$$|(T^jy)(t) - (T^jz)(t)| \leq L^j \frac{|t - t_0|^j}{j!} d_\infty(y, z) \quad \forall t \in J, \quad (590)$$

also

$$|(T^jy)(t) - (T^jz)(t)| \leq L^j \frac{(b-a)^j}{j!} d_\infty(y, z) \quad \forall t \in J. \quad (591)$$

Nehmen wir das Supremum über alle $t \in J$ so erhalten wir

$$d_\infty(T^j y, T^j z) \leq \frac{L^j(b-a)^j}{j!} d_\infty(y, z), \quad (592)$$

wie gewünscht. Dass es nun ein $n_0 \in \mathbb{N}$ gibt, sodass T^{n_0} eine strikte Kontraktion ist, liegt daran, dass

$$\lim_{j \rightarrow \infty} \frac{L^j(b-a)^j}{j!} = 0 \quad (593)$$

(Wie man zum Beispiel daran sehen könnte, dass die Reihe über den obigen Ausdruck gegen $e^{L(b-a)} < \infty$ konvergiert). Insbesondere muss es also $n_0 \in \mathbb{N}$ geben mit

$$\frac{L^{n_0}(b-a)^{n_0}}{(n_0)!} < 1. \quad (594)$$

Die Existenz eines eindeutigen Fixpunktes folgt dann direkt aus Satz 102 (p. 98).

Insbesondere existiert auf J jetzt nach Satz 102 (p. 98) ein eindeutiger Fixpunkt von T und damit (siehe Proposition 103, p. 99) eine eindeutige Lösung. Das verbleibende Problem ist, dass J nur eine echte Teilmenge von I ist. Wir wollen aber eine Lösung auf ganz I .

Mit Ausschöpfung können wir jetzt den globalen Satz von Picard-Lindelöf beweisen.

Wdh. Satz 90 (p. 90) SATZ VON PICARD-LINDELÖF, GLOBALE VERSION. Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ erfülle eine globale Lipschitz-Bedingung auf I . Dann hat für alle $t_0 \in I$ und $y_0 \in \mathbb{R}^n$ das Anfangswertproblem

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (595)$$

eine eindeutige Lösung.

Beweis von Satz 90 (p. 90).

Schritt 1: Existenz. Es erfülle wie in der Aussage $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung. Wähle nun eine Folge $(J_k)_{k \in \mathbb{N}}$ von Teilintervallen mit $J_k \subset\subset I$, $J_1 \subset J_2 \subset \dots$, $t_0 \in J_1$ und

$$I = \bigcup_{k=1}^{\infty} J_k \quad (596)$$

Wir definieren nun für beliebige $k \in \mathbb{N}$

$$T_k : C_b^0(J_k; \mathbb{R}^n) \rightarrow C_b^0(J_k; \mathbb{R}^n) \quad (597)$$

gegeben durch

$$(T_k y)(t) := y_0 + \int_{t_0}^t f(s, y(s)) \, ds \quad (t \in J_k). \quad (598)$$

Wir wissen durch Lemma 109 (p. 104) und Satz 102 (p. 98), dass T_k einen eindeutigen Fixpunkt $y_k \in C_b^0(J_k; \mathbb{R}^n)$ besitzt. Nach Proposition 103 gilt dann auch $y_k \in C^1(J_k; \mathbb{R}^n)$ und

$$\begin{cases} y'_k(t) = f(t, y_k(t)) & \forall t \in J_k \\ y_k(t_0) = y_0. \end{cases} \quad (599)$$

Nun wollen wir diese Lösungen zu einer Lösung auf ganz I zusammenkleben. Dazu muss aber folgendes gewährleistet sein:

Zwischenbehauptung: Für alle $l < k$ gilt $y_k(t) = y_l(t)$ für alle $t \in J_l$. Dazu zeigen wir dass $y_k|_{J_l} \in C_b^0(J_l; \mathbb{R}^n)$ ein Fixpunkt von T_l ist. In der Tat gilt für alle $t \in J_l$

$$(T_l[y_k|_{J_l}](t)) = y_0 + \int_{t_0}^t f(s, y_k(s)) \, ds \stackrel{(598)}{=} (T_k y_k)(t) = y_k(t) = y_k|_{J_l}(t). \quad (600)$$

Da T_l nur einen Fixpunkt haben kann, muss $y_k|_{J_l} = y_l$ gelten. Das ist exakt das, was wir behauptet haben. Definiere nun $y_* : I \rightarrow \mathbb{R}^n$ durch

$$y_*(t) := y_k(t) \quad \text{falls } t \in J_k. \quad (601)$$

Dies ist wohldefiniert: Zu jedem $t \in I$ gibt es ein $k \in \mathbb{N}$ mit $t \in J_k$. Die Definition ist unabhängig von k , denn falls $t \in J_k \cap J_l$ für $k \neq l$, so gilt nach der obigen Zwischenbehauptung, dass $y_k(t) = y_l(t)$. Daher ist es unerheblich, welcher Index jetzt für die Definition von y_* hergenommen wird. Nun beachte, dass y_* stetig diffbar bei jedem $t \in I$ ist, denn jedes $t \in I$ ist innerer Punkt eines Intervall J_k und wir wissen bereits, dass $y^* \in C^1(J_k; \mathbb{R}^n)$, da $y_*|_{J_k} = y_k \in C^1(J_k; \mathbb{R}^n)$. Außerdem gilt für alle $t \in J_k$

$$y'_*(t) = y'_k(t) = f(t, y_k(t)) = f(t, y_*(t)). \quad (602)$$

und, das $t_0 \in J_1$ gilt auch $y_*(t_0) = y_1(t_0) = y_0$. Mit anderen Worten ist y_* eine Lösung des AWP. Die Existenz ist gezeigt.

Schritt 2: Eindeutigkeit. Es seien $y, z \in C^1(I; \mathbb{R}^n)$ zwei Lösungen. Wir zeigen $y = z$. Dazu sei $k \in \mathbb{N}$ und $J_k \subset\subset I$ wie in Schritt 1. Da \overline{J}_k kompakt ist, sind y und z auf \overline{J}_k beschränkt (siehe Satz von Weierstraß). Damit sind y und z auch auf J_k beschränkt und es gilt $y, z \in C_b^0(J_k; \mathbb{R}^n)$. Ferner lösen $y|_{J_k}$ und $z|_{J_k}$ auf J_k das AWP, weswegen nach Proposition 103 gilt, dass

$$T_k[y|_{J_k}] = y|_{J_k}, \quad T_k[z|_{J_k}] = z|_{J_k}. \quad (603)$$

Da T_k jedoch nur einen Fixpunkt in $C_b^0(J_k; \mathbb{R}^n)$ haben kann, gilt

$$y|_{J_k} = z|_{J_k}. \quad (604)$$

Da $k \in \mathbb{N}$ beliebig war, folgt, dass $y = z$ auf $I = \bigcup_{k \in \mathbb{N}} J_k$.

Die globale Version des Satzes ist gezeigt.

Wir haben nun eine abstrakte Lösung. Können wir sie auch numerisch bestimmen? Ja:

Denn Fixpunkte von Kontraktions-Abbildungen können stets mit der *Fixpunktiteration* bestimmt werden. Die Fixpunktiteration für die Abbildung T aus (563) wird *Picard-Iteration* genannt. Allerdings gibt es hier zunächst ein Problem: Für ein beschränktes Intervall (a, b) wissen wir bisher nur, dass T^{n_0} für ein gewisses $n_0 \in \mathbb{N}$ eine strikte Kontraktionsabbildung ist (siehe Lemma 109). Daher wäre es hier eigentlich angesagt, T^{n_0} zu iterieren anstatt T . Wenn man aber bereits weiß, dass ein Fixpunkt existiert, kann man sich besser anstellen:

Gedankenexperiment 110.

NUMERISCHE APPROXIMATION VON LÖSUNGEN — DIE PICARD-ITERATION. Es sei (X, d) ein vollständiger metrischer Raum. Wir erinnern uns hier zunächst, wie wir im Beweis zum Banach'schen Fixpunktsatz den Fixpunkt einer Kontraktionsabbildung $T : X \rightarrow X$ konstruiert haben. Wir wählen $x_0 \in X$ beliebig und definieren eine rekursiv definierte Folge gegeben durch die Vorschrift

$$x_{j+1} := Tx_j \quad (j \in \mathbb{N}_0) \quad (605)$$

Mit anderen Worten *iterieren* wir T , es gilt nämlich für $n \in \mathbb{N}$

$$x_j = Tx_{j-1} = T(Tx_{j-2}) = \dots = T^j x_0. \quad (606)$$

Dieses Iterieren kann nun auch für die numerische Approximation der Lösung verwendet werden. In der Tat: Es sei $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung erfüllend, $(a, b) \subset\subset I$ ein beschränktes Intervall und $X = C_b^0((a, b); \mathbb{R}^n)$ ausgestattet mit der Metrik $d = d_\infty$. Ferner sei $T : X \rightarrow X$ wie in (563). Es sei $y \in X$ die eindeutige Lösung von

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in (a, b)) \\ y(t_0) = y_0. \end{cases} \quad (607)$$

Dann gilt nach Proposition 103 $Ty = y$. Um y zu approximieren wählen wir $z_0 \in X$ beliebig und definieren die *approximativen Lösungen* $z_j := T^j z_0$ für $j \in \mathbb{N}$. Es ist zu erwarten, dass für $j \rightarrow \infty$ $z_j \rightarrow y$ in (X, d) konvergiert. Und in der Tat gilt mit (566)

$$d_\infty(z_j, y) = d_\infty(T^j z_0, y) = d_\infty(T^j z, T^j y) \stackrel{(566)}{\leq} \frac{L^j(b-a)^j}{j!} d_\infty(z_0, y), \quad (608)$$

was für $j \rightarrow \infty$ gegen Null strebt. Hierbei haben wir benutzt, dass $T^j y = y$, weil y ja ein Fixpunkt von T ist. Man kann mit der obigen Abschätzung die Abweichung von der approximativen Lösung und der eigentlichen Lösung kontrollieren. Es gilt

$$\sup_{t \in (a, b)} |z_j(t) - y(t)| \leq \frac{L^j(b-a)^j}{j!} d_\infty(z_0, y). \quad (609)$$

Der *Fehler* (d.h. die Folge $(d_\infty(z_j, y))_{j \in \mathbb{N}}$) geht aufgrund dieser Abschätzung *sehr schnell* gegen Null — (fast) so schnell wie $\frac{1}{j!}$. Diese hohe Konvergenzgeschwindigkeit ist im Nachhinein sehr verwunderlich: Aus Lemma 109 konnten wir zunächst nur folgern, dass T^{n_0} für

ein $n_0 \in \mathbb{N}$ eine Kontraktionsabbildung ist. Aus dem Banach'schen Fixpunktsatz lernen wir daher erstmal nur, dass die *Teilfolge* $(T^{kn_0}z_0)_{k \in \mathbb{N}}$ gegen y konvergiert. Wie kommt es nun zu dieser Verbesserung, dass am Ende die *gesamte Folge* sehr schnell konvergiert? Dies ist so zu erklären: In Gleichung (608) haben wir zusätzliches Wissen reingesteckt — wir wussten nämlich bereits, dass ein Fixpunkt existiert! Dadurch konnten unsere vorherigen Abschätzungen am Fixpunkt betrachten und sie dadurch substantiell verbessern. In der Numerik spricht man von einer *a-posteriori-Abschätzung*.

Die Folge $(z_j)_{j \in \mathbb{N}}$ gegeben durch $z_j := T^j z_0$ nennen wir *Picard-Iteration mit Startwert* z_0 . Ihre Folgeglieder heißen die *Picard-Iterierten*.

Übung 111. DIE PICARD-ITERATION.

- (a) Berechnen Sie die ersten zwei Picard-Iterierten z_1, z_2 von

$$\begin{cases} y'(t) = \cos(y(t)) & (t \in (-\frac{1}{4}, \frac{1}{4})), \\ y(0) = 0. \end{cases} \quad (610)$$

Verwenden Sie den Startwert $z_0(t) := 0$ für $t \in (0, 1)$.

- (b) Zeigen Sie (ohne die DGL explizit zu lösen), dass der Fehler ihrer zweiten Picard-Iterierten kleiner gleich $\frac{1}{32}$ ist, d.h. $d_\infty(z_2, y) \leq \frac{1}{32}$.
- (c) Lösen Sie die DGL exakt und zeichnen Sie die Lösung zusammen mit den ersten zwei Picard-Lindelöf-Iterierten in ein Koordinatensystem.

SINN DER ÜBUNG: Man sieht, dass man die explizite Lösung mit der Picard-Iteration schnell und effizient approximieren kann.

HINWEIS: Für die Schranke in Aufgabenteil (b) muss $d_\infty(z_0, y)$ abgeschätzt werden. Zeigen Sie mithilfe der Differentialgleichung, dass $d_\infty(z_0, y) \leq \frac{1}{4}$. Eine solche Abschätzung nennt man in der Numerik *a-priori-Abschätzung* (als Gegenpol zur a-posteriori Abschätzung).

***-Übung 112. PICARD-ITERATION UND AUSSCHÖPFUNG.** Es sei $I \subset \mathbb{R}$ ein offenes Intervall und $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung erfüllend. Ferner sei $t_0 \in I$ und $y_0 \in \mathbb{R}^n$

- (a) Es sei $(a, b) \subset \subset I$ und $z \in C_b^0((a, b); \mathbb{R}^n)$. Es sei $T : C_b^0((a, b); \mathbb{R}^n) \rightarrow C_b^0((a, b); \mathbb{R}^n)$ wie in (563). Zeigen Sie: Es existieren

$$(Tz)(a) := \lim_{t \downarrow a}(Tz)(t), \quad \text{und} \quad (Tz)(b) := \lim_{t \uparrow b}(Tz)(t). \quad (611)$$

- (b) Wählen Sie nun $z_0 := 0 \in C^0(I; \mathbb{R}^n)$ und $(J_k)_{k \in \mathbb{N}_0}, J_0 \subset J_1 \subset J_2 \subset \dots$ eine Ausschöpfung von I , gelte etwa $J_k = (a_k, b_k) \subset \subset I$ für $k = 0, 1, \dots$ und $t_0 \in J_0$.

Es sei T_k gewählt wie in (597). Definieren Sie nun für $k \in \mathbb{N}_0$ die rekursive Folge

$$z_{k+1}(t) := \begin{cases} T_k(z_k)(t) & t \in (a_k, b_k), \\ T_k(z_k)(b_k) & t \in I \cap [b_k, \infty), \\ T_k(z_k)(a_k) & t \in I \cap (-\infty, a_k). \end{cases} \quad (612)$$

Zeigen Sie, dass $(z_k)_{k \in \mathbb{N}}$ in $C^0(I; \mathbb{R}^n)$ liegt und punktweise gegen die Lösung des AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (613)$$

konvergiert.

SINN DER ÜBUNG. Diese Übung hinterfragt, ob wir bei der Konvergenz der Picard-Iteration auf die Annahme $(a, b) \subset \subset I$ verzichten können. Verzichten wir auf diese Annahme, so müssen wir uns aber leider damit abfinden, dass die Art der Konvergenz schlechter werden kann — schließlich zeigen wir in der Übung hier ja nur punktweise Konvergenz. Dies ist in jedem Fall schlechter als die gleichmäßige Konvergenzaussage in (609), die für Teilintervalle $(a, b) \subset \subset I$ stets erfüllt ist. Ferner wird hier die wichtige Technik der *Ausschöpfung* nochmal trainiert.

Übung 113. PICARD-ITERATION FÜR LINEARE DIFFERENTIALGLEICHUNGEN. Es sei für $A \in \mathbb{R}^{n \times n}$ und $y_0 \in \mathbb{R}^n$ das AWP

$$\begin{cases} y'(t) = Ay(t) & (t \in \mathbb{R}) \\ y(t_0) = y_0 \end{cases} \quad (614)$$

gegeben.

- (a) Zeigen Sie, dass das AWP eine eindeutige Lösung auf \mathbb{R} hat.
- (b) Beweisen Sie, dass die Picard-Iterierten mit Startwert $z_0(t) := 0$ gegeben sind durch

$$z_j(t) = \sum_{k=0}^{j-1} \frac{(t - t_0)^k}{k!} A^k y_0 \quad (j \in \mathbb{N}). \quad (615)$$

- (c) Zeigen Sie, dass für alle $(a, b) \subset \subset \mathbb{R}$ jede Komponente der matrixwertigen Folge $(M_j)_{j \in \mathbb{N}} \subset C^0(I; \mathbb{R}^{n \times n})$ gegeben durch

$$M_j(t) := \sum_{k=0}^j \frac{(t - t_0)^k}{k!} A^k \quad (616)$$

gleichmäßig auf (a, b) konvergent ist.

(d) Bezeichnen Sie für $t \in \mathbb{R}$

$$M(t) := \lim_{j \rightarrow \infty} M_j(t) := \sum_{k=0}^{\infty} \frac{(t-t_0)^k}{k!} A^k. \quad (617)$$

Zeigen Sie, dass M eine Fundamentalmatrix für die DGL
 $y'(t) = Ay(t), (t \in \mathbb{R})$ ist.

HINWEIS. Vergessen Sie in Teilaufgabe (d) nicht, die Invertierbarkeit von $M(t)$ für alle $t \in \mathbb{R}$ zu beweisen.

SINN DER ÜBUNG. Zwar haben wir für alle denkbaren Matrizen A in Sektion 1.2.2 eine Fundamentalmatrix bestimmt, aber eine kompakte Formel dafür hatten wir nie. Diese reichen wir jetzt mit der Theorie der Picard-Iteration nach: (617) ist eine kompakte Darstellung einer Fundamentalmatrix, zumindest für Systeme mit konstanten Koeffizienten. Analog zur Exponentialreihe schreibt man oft

$$e^{(t-t_0)A} := \sum_{k=0}^{\infty} \frac{(t-t_0)^k}{k!} A^k. \quad (618)$$

und nennt diese Reihe die *Matrixexponentielfunktion*.

3.1.3 Beweis der lokalen Version

Es sei $f \in C^0(I \times G; \mathbb{R}^n)$. Wir nehmen ab sofort nur an, dass f eine lokale Lipschitz-Bedingung erfüllt, siehe Definition 93 (p. 91).

Um die lokale Version des Satzes zu zeigen, zeigen wir zunächst ein *Kurzzeitexistenzresultat*, d.h. wir zeigen die Existenz eines $\delta > 0$ so, dass $(t_0 - \delta, t_0 + \delta) \subset I$

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in (t_0 - \delta, t_0 + \delta)) \\ y(t_0) = y_0 \end{cases} \quad (619)$$

eine eindeutige Lösung $y \in C^1((t_0 - \delta, t_0 + \delta); G)$ besitzt. Das heißt: Man kann zu jedem Anfangswert eine kurzzeitige Lösung um diesen Wert finden. Das klingt machbar, da ja “lokal” die Lipschitzbedingung erfüllt ist, die ja schon im letzten Abschnitt ein Garant für die Existenz von Lösungen war.

Kann man nun diese Kurzzeit-Lösungen zu einer maximalen Lösung “zusammensetzen”? Die Antwort heißt “Ja”! Allerdings ist das Wort *fortsetzen* womöglich treffender als *zusammensetzen*, je nach verwendeter Literatur. Diesen Prozess wollen wir später noch genauer studieren.

Der Existenzbeweis ist somit in zwei Teile aufgeteilt:

1. Kurzzeitexistenz.
2. Zusammensetzen der Kurzzeitlösungen zur maximalen Lösung.

Den Eindeutigkeitsbeweis werden wir am Ende separat führen, weil man beim Zusammensetzungs-Verfahren die Eindeutigkeit verliert.

Schritt 1. Kurzzeitexistenz.

Die Kurzzeitexistenz in (619) erreichen wir wie folgt: Die lokale Lipschitz-Bedingung versichert uns, dass eine Lipschitz-Abschätzung auf einer kleinen Menge der Gestalt $(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)$ für $\epsilon, r > 0$ (vgl. Definition 93) erfüllt ist. Wir zeigen nun, dass es immer möglich ist, eine *Fortsetzungsfunktion* $F : (t_0 - \epsilon, t_0 + \epsilon) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ zu wählen, die die globale(!) Lipschitzbedingung erfüllt und auf $(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)$ mit f übereinstimmt. Dankenswerterweise können wir dann das Anfangswertproblem

$$\begin{cases} y'(t) = F(t, y(t)) & (t \in (t_0 - \epsilon, t_0 + \epsilon)) \\ y(t_0) = y_0 \end{cases} \quad (620)$$

bereits auf $(t_0 - \epsilon, t_0 + \epsilon)$ lösen. Wir werden dann zeigen, dass diese Lösung zumindest für ganz kurze Zeit eine Lösung unseres ursprünglichen AWP in (619) liefert. Hierbei müssen wir das Zeitintervall vielleicht nochmal verkleinern.

Man beachte: Die genaue Gestalt der Fortsetzung hängt maßgeblich von (t_0, y_0) ab – denn F ist schließlich nur eine Fortsetzung von $f|_{(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)}$ und nicht von $f|_{I \times G}$. Letzteres wäre im Allgemeinen nicht möglich! Dies kann man zum Beispiel an den Gegenbeispielen aus Übung 98 (p. 95) sehen.

Die Existenz der in (620) benötigten Fortsetzung ist versichert durch den *Satz von Kirschbraun*.

Proposition 114. SATZ VON KIRSZBRAUN. Es sei $U \subset \mathbb{R}^n$ offen und beschränkt und $J \subset \mathbb{R}$ offen. Es sei $f : J \times U \rightarrow \mathbb{R}^n$ so, dass ein $L > 0$ existiert mit

$$|f(t, y) - f(t, z)| \leq L|y - z| \quad \forall y, z \in U \quad \forall t \in J. \quad (621)$$

Dann gibt es eine Funktion $F : J \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $F|_{J \times U} = f$ und

$$|F(t, y) - F(t, z)| \leq \sqrt{n}L|y - z| \quad \forall y, z \in \mathbb{R}^n \quad \forall t \in J. \quad (622)$$

Insbesondere erfüllt F eine globale Lipschitz-Bedingung auf J . Falls $f \in C^0(J \times U; \mathbb{R}^n)$, so ist $F \in C^0(J \times \mathbb{R}^n; \mathbb{R}^n)$.

Beweis. Der Beweis funktioniert mit einer expliziten Konstruktion von F . Wir definieren die Funktion $F = (F_1, \dots, F_n)^T$ durch

$$F_i(t, x) := \inf_{y \in U} (f_i(t, y) + L|y - x|) \quad (t \in J, x \in \mathbb{R}^n), \quad i = 1, \dots, n. \quad (623)$$

Dieses Infimum ist auf $J \times \mathbb{R}^n$ stets eine reelle Zahl, da $f_i(t, \cdot)$ auf U nach unten beschränkt ist, wie man leicht aus (621) folgert. In der Tat: Wählt man $x_0 \in U$ beliebig, so sieht man

$$f_i(t, y) = f_i(t, x_0) - (f_i(t, x_0) - f_i(t, y)) \geq f_i(t, x_0) - |f_i(t, x_0) - f_i(t, y)| \geq f_i(t, x_0) - L|x_0 - y|. \quad (624)$$

Dies schätzt man nun mit der Dreiecksungleichung ab zu $f_i(t, y) \geq f_i(t, x_0) - 2L \sup_{w \in U} |w|$.

Schritt 1: F_i ist eine Fortsetzung von f_i . Wir behaupten zunächst, dass F_i auf $J \times U$ mit

f_i übereinstimmt, also

$$F_i(t, x) = f_i(t, x) \quad \forall x \in U \quad \forall t \in J. \quad (625)$$

Dazu sei $x \in U$ beliebig aber fest. Wir zeigen zunächst " \leq " in (625). Dazu:

$$F_i(t, x) \underset{(623)}{=} \inf_{y \in U} (f_i(t, y) + L|y - x|) \underset{\text{Betrachte } y=x}{\leq} f_i(t, x) + L|x - x| = f_i(t, x). \quad (626)$$

Hierbei haben wir beim vorletzen Schritt benutzt, dass $x \in U$, damit wir $y = x$ betrachten durften. Nun zeigen wir noch " \geq " in (625). Wir behaupten dazu, dass

$$f_i(t, y) + L|y - x| \geq f_i(t, x) \quad \forall y \in U. \quad (627)$$

In der Tat gilt für $y \in U$

$$f_i(t, x) - f_i(t, y) \leq |f_i(t, x) - f_i(t, y)| \leq |f(t, x) - f(t, y)| \underset{(621)}{\leq} L|x - y|. \quad (628)$$

Addieren wir nun $f_i(t, y)$ auf beiden Seiten, so erhalten wir in der Tat (627). Aus (627) folgt nun dass

$$\inf_{y \in U} (f_i(t, y) + L|y - x|) \geq f_i(t, x), \quad (629)$$

d.h.

$$F_i(t, x) \geq f_i(t, x), \quad (630)$$

wie gewünscht.

Schritt 2: F_i genügt einer globalen Lipschitz-Bedingung. Es seien $y, z \in \mathbb{R}^n$ beliebig aber fest. Da man ein Infimum einer Menge stets mit einer Folge annähern kann, finden wir eine Folge $(w_j)_{j=1}^\infty \subset U$ mit

$$f_i(t, w_j) + L|w_j - z| \rightarrow F_i(t, z) \quad (j \rightarrow \infty). \quad (631)$$

Benutzen wir dies und die Definition von F_i , so erhalten wir

$$\begin{aligned} F_i(t, y) - F_i(t, z) &\underset{(631)}{=} \lim_{j \rightarrow \infty} F_i(t, y) - (f_i(t, w_j) + L|w_j - z|) \\ &\underset{(623)}{\leq} \liminf_{j \rightarrow \infty} [f_i(t, w_j) + L|w_j - y|] - (f_i(t, w_j) + L|w_j - z|) \\ &\leq \liminf_{j \rightarrow \infty} L(|w_j - y| - |w_j - z|) \\ &= \liminf_{j \rightarrow \infty} L(|(w_j - z) + z - y| - |w_j - z|) \\ &\leq \liminf_{j \rightarrow \infty} L(|w_j - z| + |z - y| - |w_j - z|) = L|y - z|, \end{aligned}$$

wobei wir im vorletzten Schritt die Dreiecksungleichung benutzt haben. Man kann nun die Rollen von y und z vertauschen und erhält analog, dass

$$F_i(t, z) - F_i(t, y) \leq L|y - z|. \quad (632)$$

Die beiden vorigen Gleichungen zusammengenommen ergibt

$$|F_i(t, z) - F_i(t, y)| \leq L|z - y|. \quad (633)$$

Schritt 3: Lipschitz-Bedingung für $F = (F_1, \dots, F_n)$. Die Definition der Euklidischen Norm liefert für alle $y, z \in \mathbb{R}^n$

$$|F(t, z) - F(t, y)| \leq \sqrt{\sum_{i=1}^n |F_i(t, z) - F_i(t, y)|^2} \leq \sqrt{nL^2|z - y|^2} = \sqrt{n}L|z - y|. \quad (634)$$

Schritt 4: Noch zu zeigen: Falls $f \in C^0(J \times U; \mathbb{R}^n)$ so ist auch $F \in C^0(J \times \mathbb{R}^n; \mathbb{R}^n)$. Wir zeigen dies Komponentenweise für $i = 1, \dots, n$. Die Stetigkeit von F_i in der zweiten Variablen folgt sofort aus der Bedingung (622). Für die Stetigkeit in der ersten Variablen sei $(t_j)_{j \in \mathbb{N}} \subset J$ eine Folge mit $t_j \rightarrow t$ und sei $y \in \mathbb{R}^n$ fest. Nach der Definition von F_i (siehe (623)) gibt es nun für alle $\epsilon > 0$ ein $z \in U$ mit

$$F_i(t, y) + \epsilon \geq f_i(t, z) + L|y - z|. \quad (635)$$

Es folgt, wiederum mit der Definition von F ,

$$F_i(t_j, y) - F_i(t, y) \stackrel{(623)}{\leq} (f_i(t_j, z) + L|y - z|) - F_i(t, y) \quad (636)$$

$$\stackrel{(635)}{\leq} f(t_j, z) + L|y - z| - (f_i(t, z) + L|y - z|) + \epsilon \quad (637)$$

$$= f_i(t_j, z) - f_i(t, z) + \epsilon. \quad (638)$$

Wegen der Stetigkeit von f_i folgt nun

$$\limsup_{j \rightarrow \infty} (F_i(t_j, y) - F_i(t, y)) \leq \limsup_{j \rightarrow \infty} (f_i(t_j, z) - f_i(t, z) + \epsilon) = \epsilon. \quad (639)$$

Da die linke Seite nicht mehr von ϵ abhängt, kann man den Grenzfall $\epsilon \rightarrow 0$ in der obigen Gleichung betrachten und folgert

$$\limsup_{j \rightarrow \infty} (F_i(t_j, y) - F_i(t, y)) \leq 0. \quad (640)$$

Fixiere nun wiederum $\epsilon > 0$ und wähle für alle $j \in \mathbb{N}$ ein $w_j \in U$ mit

$$F_i(t_j, y) + \epsilon \leq f(t_j, w_j) + L|y - w_j|. \quad (641)$$

Es gilt für alle $j \in \mathbb{N}$

$$F_i(t, y) - F_i(t_j, y) \stackrel{(623)}{\leq} (f_i(t_j, w_j) + L|y - w_j|) - F_i(t_j, y) \quad (642)$$

$$\stackrel{(635)}{\leq} f_i(t_j, w_j) + L|y - w_j| - (f_i(t, w_j) + L|y - w_j|) + \epsilon \quad (643)$$

$$= f_i(t_j, w_j) - f_i(t, w_j) + \epsilon. \quad (644)$$

Hier müssen wir etwas mehr Argumentieren, da nicht sofort offensichtlich ist, dass $f_i(t_j, w_j) - f_i(t, w_j)$ gegen Null konvergiert. Hierzu folgender Trick: Fixiere $k \in \mathbb{N}$. Dann gilt für $j \geq k$

$$|f_i(t_j, w_j) - f_i(t, w_j)| \quad (645)$$

$$= |f_i(t_j, w_j) - f_i(t_j, w_k) + f_i(t_j, w_k) - f_i(t, w_k) + f_i(t, w_k) - f_i(t, w_j)| \quad (646)$$

$$\leq |f_i(t_j, w_j) - f_i(t_j, w_k)| + |f_i(t_j, w_k) - f_i(t, w_k)| + |f_i(t, w_k) - f_i(t, w_j)|. \quad (647)$$

Da nun $w_j, w_k \in U$ sind, kann im ersten und dritten Summanden die Lipschitz-Abschätzung für f (siehe (621)) verwendet werden und wir erhalten

$$|f_i(t_j, w_j) - f_i(t, w_j)| \leq 2L|w_j - w_k| + |f_i(t, w_k) - f_i(t, w_k)|. \quad (648)$$

Es sei nun $(l_j)_{j \in \mathbb{N}} \subset \mathbb{N}$ eine Teilfolge mit

$$\lim_{j \rightarrow \infty} (F_i(t, y) - F_i(t_{l_j}, y)) = \limsup_{j \rightarrow \infty} (F_i(t, y) - F_i(t_j, y)). \quad (649)$$

Da nun die Folge $(w_{l_j})_{j \in \mathbb{N}} \subset U$ beschränkt ist (wegen der Beschränktheit von U), gibt es eine konvergente Teilfolge $(w_{l_j^1})_{j \in \mathbb{N}}$. Sei etwa w der Grenzwert dieser Folge. Sei nun wieder $k \in \mathbb{N}$ beliebig. Dann gilt

$$\limsup_{j \rightarrow \infty} (F_i(t, y) - F_i(t_j, y)) = \lim_{j \rightarrow \infty} (F_i(t, y) - F_i(t_{l_j}, y)) = \lim_{j \rightarrow \infty} (F_i(t, y) - F_i(t_{l_j^1}, y)) \quad (650)$$

$$\stackrel{(644)}{\leq} \limsup_{j \rightarrow \infty} (f_i(t_{l_j^1}, w_{l_j^1}) - f_i(t, w_{l_j^1})) + \epsilon \quad (651)$$

$$\stackrel{(648)}{\leq} \limsup_{j \rightarrow \infty} (2L|w_{l_j^1} - w_{l_k^1}| + |f(t_{l_j^1}, w_{l_k^1}) - f(t, w_{l_k^1})|) + \epsilon \quad (652)$$

$$= 2L|w - w_{l_k^1}| + \epsilon, \quad (653)$$

wobei wir im letzten Schritt die Stetigkeit von f am (festen) Punkt $(t, w_{l_k^1})$ ausgenutzt haben. Da nun aber k beliebig war, können wir $k \rightarrow \infty$ betrachten und erhalten

$$\limsup_{j \rightarrow \infty} (F_i(t, y) - F_i(t_j, y)) \leq \epsilon. \quad (654)$$

Betrachtet man wieder $\epsilon \rightarrow 0$ so erhält man

$$\limsup_{j \rightarrow \infty} (F_i(t, y) - F_i(t_j, y)) = 0 \quad (655)$$

Dies und (640) impliziert

$$\limsup_{j \rightarrow \infty} |F_i(t_j, y) - F_i(t, y)| = \limsup_{j \rightarrow \infty} \max\{(F_i(t, y) - F_i(t_j, y)), (F_i(t_j, y) - F_i(t, y))\} = 0 \quad (656)$$

Diese Lipschitz-Fortsetzung versetzt uns — wie bereits vorher angekündigt — in die Lage, die Kurzzeitexistenz zu zeigen.

Lemma 115. KURZZEITEXISTENZ. Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und es sei $t_0 \in I$ und $y_0 \in G$. Dann gibt es ein $\delta > 0$ so, dass

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in (t_0 - \delta, t_0 + \delta)) \\ y(t_0) = y_0 \end{cases} \quad (657)$$

eine Lösung $y \in C^1((t_0 - \delta, t_0 + \delta); \mathbb{R}^n)$ besitzt.

Beweis. Da f eine lokale Lipschitz-Bedingung erfüllt, gibt es $\epsilon = \epsilon(t_0, y_0) > 0$, $r = r(t_0, y_0) > 0$ und $L = L(t_0, y_0) > 0$ mit $\overline{B_r(y_0)} \subset G$ und

$$|f(t, y) - f(t, z)| \leq L|y - z| \quad \forall t \in (t_0 - \epsilon, t_0 + \epsilon) \quad \forall y, z \in B_r(y_0). \quad (658)$$

Wenden wir nun Proposition 114 (p. 113) mit $J = (t_0 - \epsilon, t_0 + \epsilon)$ und $U = B_r(y_0)$ an, so erhalten wir eine Funktion $F : (t_0 - \epsilon, t_0 + \epsilon) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit der Eigenschaft, dass $F|_{(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)} = f|_{(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)}$ und

$$|F(t, y) - F(t, z)| \leq \sqrt{n}L|y - z| \quad \forall t \in (t_0 - \epsilon, t_0 + \epsilon) \quad \forall y, z \in \mathbb{R}^n. \quad (659)$$

Insbesondere erfüllt F eine globale Lipschitz-Bedingung und somit gibt es eine eindeutige Lösung $\bar{y} \in C^1((t_0 - \epsilon, t_0 + \epsilon); \mathbb{R}^n)$ von

$$\begin{cases} \bar{y}'(t) = F(t, \bar{y}(t)) & (t \in (t_0 - \epsilon, t_0 + \epsilon)) \\ \bar{y}(t_0) = y_0 \end{cases} \quad (660)$$

Wir behaupten dass eine passende Einschränkung von \bar{y} auf ein Teilintervall das gewünschte AWP (657) löst. Man beobachte dazu, dass $\bar{y}(t_0) = y_0 \in B_r(y_0)$. Wegen der Stetigkeit von \bar{y} gibt es dann ein $\delta > 0$ mit $\bar{y}(t) \in B_r(y_0)$ für alle $t \in (t_0 - \delta, t_0 + \delta)$. Daraus folgern wir (mit $F|_{(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)} = f|_{(t_0 - \epsilon, t_0 + \epsilon) \times B_r(y_0)}$)

$$\bar{y}'(t) = F(t, \bar{y}(t)) = f(t, \bar{y}(t)) \quad \forall t \in (t_0 - \delta, t_0 + \delta). \quad (661)$$

Insgesamt gilt also $\bar{y} \in C^1((t_0 - \delta, t_0 + \delta); G)$ und

$$\begin{cases} \bar{y}'(t) = f(t, \bar{y}(t)) & (t \in (t_0 - \delta, t_0 + \delta)) \\ \bar{y}(t_0) = y_0. \end{cases} \quad (662)$$

Dies zeigt, dass $\bar{y}|_{(t_0-\delta, t_0+\delta)}$ eine Lösung des gewünschten AWP (657) ist.

Man beachte, dass wir mit dem obigen Beweis keine Chance haben, zu bestimmen, wie groß das Existenzintervall $(t_0 - \delta, t_0 + \delta)$ wirklich ist. Da wir dies nicht wissen, bezeichnen wir die Lösung y auch als *Kurzzeitlösung*.

Bemerkung 116. NUMERISCHE APPROXIMATION DER KURZZEITLÖSUNG. Im vorigen Lemma haben wir die Kurzzeitlösung konstruiert indem wir das Hilfsproblem (660) betrachtet haben. Dieses erfüllt, wie wir gesehen haben, eine globale Lipschitz-Bedingung. Daraus folgt insbesondere auch, dass die Lösung von (660) — und somit auch die Kurzzeitlösung unseres ursprünglichen Problems — mit der *Picard-Iteration* approximieren können, vgl. Gedankenexperiment 110. Der einzige subtile Unterschied ist, dass man den Operator T mit der Lipschitz-Fortsetzung F anstatt f konstruieren muss.

Eine alternative Möglichkeit, die Kurzzeitexistenz ohne den Satz von Kirschbraun bzw. ohne den globalen Satz von Picard-Lindelöf zu zeigen, ist eine *Lokalisierung* der Fixpunktmethodik, welche wir im vorangegangenen Abschnitt bereits kennengelernt haben. Diese Herangehensweise wird in der Literatur öfter verwendet.

Übung 117. KURZZEITEXISTENZ MIT FIXPUNKTMETHODEN.

Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung. Definieren Sie für $\eta, \delta > 0$

$$M_{\eta, \delta} := \{y \in C_b((t_0 - \delta, t_0 + \delta); \mathbb{R}^n) : d_\infty(y, y_0) \leq \eta\}. \quad (663)$$

[Hierbei wird $y_0 \in \mathbb{R}^n$ mit der konstanten Funktion $y_0(t) \equiv y_0 \in C_b((t_0 - \delta, t_0 + \delta); \mathbb{R}^n)$ identifiziert.] Definieren Sie

$$T_{\eta, \delta} : M_{\eta, \delta} \rightarrow C_b((t_0 - \delta, t_0 + \delta); \mathbb{R}^n) \quad (664)$$

durch

$$(T_{\eta, \delta} y)(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds. \quad (665)$$

Zeigen Sie

- (a) $M_{\eta, \delta}$ ist ein vollständiger metrischer Raum
- (b) Für $\eta > 0, \delta > 0$ klein genug ist $T_{\eta, \delta}(M_{\eta, \delta}) \subset M_{\eta, \delta}$
- (c) Für $\eta > 0, \delta > 0$ klein genug ist $T_{\eta, \delta}$ eine Kontraktionsabbildung.
- (d) Folgern Sie Lemma 115 ohne Benutzung des globalen Satzes von Picard-Lindelöf.

SINN DER ÜBUNG: Man sieht, dass man die Kurzzeitexistenz auch nur mit *lokalen Methoden*, d.h. ohne Kenntnis des globalen Satzes zeigen kann. Ein wichtiger Schritt hierbei

ist die Aufnahme einer *Kleinheitsbedingung* in die Definition des betrachteten metrischen Raumes. Noch einmal sieht man, dass der zugrundeliegende metrische Raum eine große Rolle spielt.

Schritt 2. Zusammenkleben zur maximalen Lösung.

Wir werden nun ein Verfahren besprechen, welches es erlaubt, die gefundene Kurzzeitlösung *größtmöglich fortzusetzen*. Was wir dann zunächst erhalten ist eine Lösung, die sich nicht weiter fortsetzen lässt. Diese nennen wir suggestiv *nichtfortsetzbare Lösung*. Eine nichtfortsetzbare Lösung ist aber etwas anderes als eine *maximale Lösung* im Sinne von Satz 94 (p. 91) — Schließlich ist im Maximalitätsbegriff auch enthalten, dass jede weitere Lösung des AWP eine Einschränkung der maximalen Lösung sein muss.

Daher ist die Existenz einer nicht-fortsetzbaren Lösung nur der erste Schritt auf dem Weg zu einer maximalen Lösung. Wir definieren hier zunächst, was wir genau unter einer nicht-fortsetzbaren Lösung verstehen.

Definition 118. NICHT-FORTSETZBARE LÖSUNG. Es sei $f \in C^0(I \times G; \mathbb{R}^n)$. Eine Lösung $y \in C^1(J; G)$, $J \subset I$ eines AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in J) \\ y(t_0) = y_0 \end{cases} \quad (666)$$

heißt *nicht-fortsetzbar*, falls es kein Intervall $\tilde{J} \supsetneq J$ gibt, sodass es $z \in C^1(\tilde{J}; G)$ gibt mit $z|_J = y$ und

$$\begin{cases} z'(t) = f(t, z(t)) & (t \in \tilde{J}) \\ z(t_0) = y_0. \end{cases} \quad (667)$$

Proposition 119. EXISTENZ EINER NICHT-FORTSETZBAREN LÖSUNG.

Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und es sei $t_0 \in I$ und $y_0 \in G$. Dann gibt es ein offenes Intervall $\hat{I} \subset I$ mit $t_0 \in \hat{I}$ und eine nichtfortsetzbare Lösung $\hat{y} \in C^1(\hat{I}; G)$.

Beweis. Für $J \subset I$ bezeichnen wir mit $(AWP|J)$ das Anfangswertproblem

$$(AWP|J) : \quad \begin{cases} y'(t) = f(t, y(t)) & (t \in J), \\ y(t_0) = y_0. \end{cases} \quad (668)$$

Es sei $\delta > 0$ so, dass eine Kurzzeitlösung $y_{kurz} \in C^1((t_0 - \delta, t_0 + \delta); \mathbb{R}^n)$ existiert mit

$$\begin{cases} y'_{kurz}(t) = f(t, y_{kurz}(t)) & (t \in (t_0 - \delta, t_0 + \delta)), \\ y_{kurz}(t_0) = y_0. \end{cases} \quad (669)$$

Insbesondere ist y_{kurz} Lösung von $(AWP|(t_0 - \delta, t_0 + \delta))$. Setze nun

$$t_+ := \sup\{\bar{t} \geq t_0 : (AWP|(t_0 - \delta, \bar{t})) \text{ hat Lösung } \bar{y} \text{ mit } \bar{y}|_{(t_0 - \delta, t_0 + \delta)} = y_{\text{kurz}}\}. \quad (670)$$

und

$$t_- := \inf\{\underline{t} \leq t_0 : (AWP|(\underline{t}, t_0 + \delta)) \text{ hat Lösung } \underline{y} \text{ mit } \underline{y}|_{(t_0 - \delta, t_0 + \delta)} = y_{\text{kurz}}\}. \quad (671)$$

Man beachte, dass wegen der Kurzzeitexistenz $t_+ \geq t_0 + \delta$ und $t_- \leq t_0 - \delta$. Wir zeigen, dass $\hat{I} = (t_-, t_+)$, d.h. wir konstruieren eine nicht-fortsetzbare Lösung $\hat{y} \in C^1((t_-, t_+); \mathbb{R}^n)$. Hierzu sei $\bar{t}_j \uparrow t_+$ eine monoton wachsende Folge mit der Eigenschaft, dass $(AWP|(t_0 - \delta, \bar{t}_j))$ eine Lösung \bar{y}_j hat, die auf $(t_0 - \delta, t_0 + \delta)$ mit y_{kurz} übereinstimmt. Ferner sei $\underline{t}_j \downarrow t_-$ monoton fallend und so, dass $(AWP|(\underline{t}_j, t_0 + \delta))$ eine Lösung \underline{y}_j hat, die auf $(t_0 - \delta, t_0 + \delta)$ mit y_{kurz} übereinstimmt. Definiere für $t \in (\underline{t}_j, \bar{t}_j)$

$$y_j(t) := \begin{cases} \underline{y}_j(t) & t \in (\underline{t}_j, t_0], \\ \bar{y}_j(t) & t \in (t_0, \bar{t}_j). \end{cases} \quad (672)$$

Beachte, dass y_j stetig differenzierbar ist — weil y_j auf $(t_0 - \delta, t_0 + \delta)$ mit y_{kurz} übereinstimmt ist die Fallunterscheidung für die Differenzierbarkeit kein Problem. Ferner ist y_j eine Lösung von $(AWP|(\underline{t}_j, \bar{t}_j))$, denn da y_j und \bar{y}_j auf ihren jeweiligen Intervallen die DGL lösen können wir berechnen

$$y'_j(t) := \begin{cases} y'_j(t) & t \in (\underline{t}_j, t_0], \\ \bar{y}'_j(t) & t \in (t_0, \bar{t}_j). \end{cases} = \begin{cases} f(t, \underline{y}_j(t)) & t \in (\underline{t}_j, t_0], \\ f(t, \bar{y}_j(t)) & t \in (t_0, \bar{t}_j). \end{cases} \stackrel{(672)}{=} f(t, y_j(t)). \quad (673)$$

Die letzte Gleichheit wird hier klar, wenn man sie rückwärts liest und die Fallunterscheidung in (672) benutzt. Klarerweise ist deswegen auch $y_j(t_0) = y_{\text{kurz}}(t_0) = y_0$ für alle $j \in \mathbb{N}$. Gerne würden wir nun alle y_j zu einer Lösung auf (t_-, t_+) vereinen. Hierzu benötigen wir die folgende Zwischenbehauptung.

Zwischenbehauptung: Sei $I_j := (\underline{t}_j, \bar{t}_j)$. Dann gilt für alle $k > j$ stets $(y_k)|_{I_j} = y_j$. Zum Beweis definieren wir

$$A_{k,j} := \{t \in I_j : y_k(t) = y_j(t)\} \quad (674)$$

Zu zeigen ist, dass $A_{k,j} = I_j$. Wir werden hierzu zeigen, dass $A_{k,j}$ offen ist und auch $I_j \setminus A_{k,j}$ offen ist. Somit ist dann I_j die disjunkte Vereinigung aus den offenen Mengen $A_{k,j}$ und $I_j \setminus A_{k,j}$. Da nun I_j zusammenhängend ist, muss entweder $A_{k,j} = I_j$ oder $A_{k,j} = \emptyset$ gelten. Der Fall $A_{k,j} = \emptyset$ ist aber ausgeschlossen, da y_k und y_j auf $(t_0 - \delta, t_0 + \delta)$ beide mit y_{kurz} übereinstimmen und somit (siehe (674)) gilt $(t_0 - \delta, t_0 + \delta) \subset A_{k,j}$.

SCHRITT 1. $I_j \setminus A_{k,j}$ ist offen, denn nach (674) gilt

$$I_j \setminus A_{k,j} \stackrel{(674)}{=} \{t \in I_j : y_k(t) \neq y_j(t)\} = \{t \in I_j : |y_k(t) - y_j(t)| > 0\}. \quad (675)$$

Nun definiert $h : I_j \rightarrow \mathbb{R}$, $h(t) := |y_k(t) - y_j(t)|$, eine stetige Funktion, weswegen $h^{-1}((0, \infty))$ als Urbild einer offenen Menge stets offen ist. Nun liest man an der vorigen Gleichung ab, dass $I_j \setminus A_{k,j} = h^{-1}((0, \infty))$, wodurch also $I_j \setminus A_{k,j}$ offen ist.

SCHRITT 2. $A_{k,j}$ ist offen. Es sei $\tau \in A_{k,j}$, d.h. $y_k(\tau) = y_j(\tau)$. Wir zeigen, dass eine offene Umgebung von τ auch in $A_{k,j}$ enthalten ist. Hierzu benennen wir zunächst $y_* := y_k(\tau) = y_j(\tau)$. Dadurch $y_k|_{I_j}$ und y_j nun beides Lösungen des AWP

$$\begin{cases} v'(t) = f(t, v(t)) & (t \in I_j), \\ v(\tau) = y_*. \end{cases} \quad (676)$$

Da f die lokale Lipschitz-Bedingung erfüllt, gibt es nun ein $\epsilon = \epsilon(\tau, y_*)$, ein $r = r(\tau, y_*)$ und ein $L = L(\tau, y_*)$ so, dass $(\tau - \epsilon, \tau + \epsilon) \subset I_j$

$$|f(t, z_1) - f(t, z_2)| \leq L|z_1 - z_2| \quad \forall t \in (\tau - \epsilon, \tau + \epsilon), z_1, z_2 \in B_r(y_*). \quad (677)$$

Nach dem Satz von Kirschbraun (siehe Proposition 114) gibt es eine Funktion F die eine globale Lipschitz-Bedingung erfüllt und $F|_{(\tau - \epsilon, \tau + \epsilon) \times B_r(y_*)} = f|_{(\tau - \epsilon, \tau + \epsilon) \times B_r(y_*)}$. Ferner gibt es wegen der Stetigkeit von y_j und y_k (und wegen $y_j(\tau) = y_k(\tau) = y_*$) ein $\eta \in (0, \epsilon)$ so, dass $y_j(t), y_k(t) \in B_r(y_*)$ für alle $t \in (\tau - \eta, \tau + \eta)$. Daraus folgert man, dass $(y_k)|_{(\tau - \eta, \tau + \eta)}$ und $(y_j)|_{(\tau - \eta, \tau + \eta)}$ beides Lösungen sind von

$$\begin{cases} v'(t) = F(t, v(t)) & (t \in (\tau - \eta, \tau + \eta)), \\ v(\tau) = y_*. \end{cases} \quad (678)$$

In der Tat: Für $t \in (\tau - \eta, \tau + \eta)$ gilt $y'_j(t) = f(t, y_j(t)) = F(t, y_j(t))$, wobei wir im letzten Schritt benutzt haben, dass $y_j((\tau - \delta, \tau + \delta)) \subset B_r(y_*)$. Analoges rechnet man für y_k . Da nun aber F eine globale Lipschitzbedingung erfüllt, muss nach dem globalen Satz von Picard-Lindelöf die Lösung von (678) eindeutig sein. Daraus folgt dann, $y_k(t) = y_j(t)$ für alle $t \in (\tau - \eta, \tau + \eta) \subset I_j$. Insbesondere gilt $(\tau - \eta, \tau + \eta) \subset A_{k,j}$. Die Offenheit von $A_{k,j}$ ist damit gezeigt. Wie wir oben schon erklärt haben, folgt aus Schritt 1 und Schritt 2 auch die Zwischenbehauptung. Nun können wir \hat{y} auf $\hat{I} = (t_-, t_+)$ definieren. Da $\hat{I} = \bigcup_{j=1}^{\infty} I_j$ (wegen $\bar{t}_j \uparrow t_+$ und $\underline{t}_j \downarrow t_-$) gibt es für jedes $t \in \hat{I}$ ein $j \in \mathbb{N}$ mit $t \in I_j$. Mit dieser Erkenntnis setzen wir

$$\hat{y}(t) := y_j(t) \quad \text{falls } t \in I_j. \quad (679)$$

Aufgrund der Zwischenbehauptung ist diese Formel für \hat{y} sogar unabhängig von der Wahl von j und somit ist \hat{y} wohldefiniert. Es bleibt zu zeigen, dass $\hat{y} \in C^1(\hat{I}; \mathbb{R}^n)$ eine nicht-fortsetzbare Lösung von $(AWP|\hat{I})$ ist. Dass \hat{y} auf \hat{I} stetig differenzierbar ist, folgt daraus, dass \hat{y} auf I_j mit der stetig diff'baren Funktion y_j übereinstimmt. Da jedes $t \in \hat{I}$ auch in einem der offenen Teilintervalle I_j enthalten ist, folgt die stetige Differenzierbarkeit. Für $t \in I_j$ gilt ferner

$$\hat{y}'(t) = y'_j(t) = f(t, y_j(t)) = f(t, \hat{y}(t)). \quad (680)$$

Wiederum folgt die obige Formel auch für $t \in \hat{I}$, weil $\hat{I} = \bigcup_{j=1}^{\infty} I_j$. Ferner gilt auch $\hat{y}(t_0) = y_0$, da jedes y_j diesen Anfangswert haben muss. Noch zu zeigen ist nun, dass \hat{y} nicht fortsetzbar ist. Gäbe es nun ein $\tilde{J} \supsetneq \hat{I}$ und $z \in C^1(\tilde{J}, G)$ welches $(AWP|\tilde{J})$ löst und $z|_{\hat{I}} = \hat{y}$ erfüllt, so wäre wegen der Definition von t_+ (vgl. (670)) und t_- (vgl. (671)) $t_+ \geq \sup(\tilde{J})$ und $t_- \leq \inf(\tilde{J})$. Hierbei haben wir benutzt, dass $z|_{(t_0-\delta, t_0+\delta)} = y_{kurz}$ (da z ja auf \hat{I} mit \hat{y} übereinstimmt und \hat{y} auf $(t_0 - \delta, t_0 + \delta)$ mit y_{kurz} übereinstimmt). Benutzen wir nun die Tatsache, dass $\hat{I} = (t_-, t_+)$, so erhalten wir

$$\hat{I} = (t_-, t_+) \supset (\inf(\tilde{J}), \sup(\tilde{J})) = \tilde{J}, \quad (681)$$

ein Widerspruch. Es folgt, dass \hat{y} nicht fortsetzbar ist.

Eine nicht-fortsetzbare Lösung ist konstruiert. Nun muss nur noch deren Maximalität geklärt werden. Hierbei klären wir zunächst, dass jede weitere Lösung eine Einschränkung der nichtfortsetzbaren Lösung aus der vorigen Proposition sein muss. Aus dieser Aussage folgt auch eigentlich sofort, dass es nur eine maximale Lösung im Sinne von Satz 94 geben kann. Dies klären wir aber später im Beweis von Satz 94 nochmal genauer auf.

Lemma 120. MAXIMALITÄT EINER NICHT-FORTSETZBAREN LÖSUNG. Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung. Zusätzlich sei $t_0 \in I$ und $y_0 \in G$. Es sei $\hat{y} \in C^1(\hat{I}; G)$ eine nicht-fortsetzbare Lösung von (666). Es sei nun $J \subset I$ ein weiteres offenes Intervall mit $t_0 \in J$, sodass es eine $z \in C^1(J; G)$ von

$$\begin{cases} z'(t) = f(t, z(t)) & (t \in J) \\ z(t_0) = y_0 \end{cases} \quad (682)$$

gibt. Dann gilt $J \subset \hat{I}$ und $\hat{y}|_J = z$.

Beweis. Es seien J, z wie in der Aussage. Setze nun $O := J \cap \hat{I}$. Man beachte, dass O ein offenes Intervall ist, welches t_0 enthält. Definere nun

$$A := \{t \in O : z(t) = \hat{y}(t)\}. \quad (683)$$

Zwischenbehauptung. $A = O$. Wir werden dazu zeigen, dass A und $O \setminus A$ beide offen sind. Insbesondere ist dann O die disjunkte Vereinigung aus den offenen Mengen A und $O \setminus A$. Da O zusammenhängend ist, folgt aus einem Satz aus der Analysis, dass entweder $A = \emptyset$ oder $A = O$. Da jedoch $t_0 \in A$ ist gilt aber $A \neq \emptyset$ und somit $A = O$.

SCHRITT 1: Offenheit von $O \setminus A$. In der Tat ist

$$O \setminus A = \{t \in O : z(t) \neq \hat{y}(t)\}. \quad (684)$$

Die Mengenbeschreibung können wir umformen zu

$$O \setminus A = \{t \in O : |z(t) - \hat{y}(t)| > 0\} \quad (685)$$

Da nun $h : O \rightarrow \mathbb{R}$ definiert durch $h(t) := |z(t) - \hat{y}(t)|$ stetig auf O ist, ist nach einem Satz aus der Analysis $h^{-1}((0, \infty))$ offen. Da nach (685) $h^{-1}((0, \infty)) = O \setminus A$, ist auch $O \setminus A$ offen.

SCHRITT 2: Nun zur Offenheit von A . Es sei $\tau \in A$, d.h. $z(\tau) = \hat{y}(\tau)$. Wir werden zeigen, dass es ein $\delta > 0$ gibt mit $(\tau - \delta, \tau + \delta) \subset A$. Dazu definieren wir $y_1 := z(\tau) = \hat{y}(\tau)$. Man beachte, dass $\hat{y}|_O$ und $z|_O$ beides Lösungen von

$$\begin{cases} v'(t) = f(t, v(t)) & (t \in O) \\ v(\tau) = y_1 \end{cases} \quad (686)$$

sind. Wir wissen auch, dass es zu (τ, y_1) ein $\epsilon = \epsilon(\tau, y_1)$, ein $r = r(\tau, y_1)$ und ein $L = L(\tau, y_1)$ gibt mit

$$|f(t, v) - f(t, w)| \leq L|v - w| \quad \forall t \in (\tau - \epsilon, \tau + \epsilon), \forall v, w \in B_r(y_1). \quad (687)$$

Ferner gibt es nach dem Satz von Kirschbraun (siehe 114, p. 113) ein $F : (\tau - \epsilon, \tau + \epsilon) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, welches eine globale(!) Lipschitz-Bedingung erfüllt und mit $F|_{(\tau-\epsilon, \tau+\epsilon) \times B_r(y_1)} = f|_{(\tau-\epsilon, \tau+\epsilon) \times B_r(y_1)}$. Wähle nun $\delta \in (0, \epsilon)$ so klein, dass $\hat{y}((\tau - \delta, \tau + \delta)) \subset B_r(y_1)$ und $z((\tau - \delta, \tau + \delta)) \subset B_r(y_1)$. Dieses δ kann gewählt werden, da y_m und z beide stetig sind und bei τ den Wert y_1 annehmen (HA). Für $t \in (\tau - \delta, \tau + \delta)$ gilt nun nach (686)

$$\hat{y}'(t) = f(t, \hat{y}(t)) = F(t, \hat{y}(t)), \quad (688)$$

und analog

$$z'(t) = f(t, z(t)) = F(t, z(t)). \quad (689)$$

Das heißt — wiederum mit (686) — dass $\hat{y}|_{(\tau-\delta, \tau+\delta)}$ und $z|_{(\tau-\delta, \tau+\delta)}$ beides Lösungen von

$$\begin{cases} v'(t) = F(t, v(t)) & (t \in (\tau - \delta, \tau + \delta)) \\ v(\tau) = y_1 \end{cases} \quad (690)$$

sind. Da F nun eine globale (!) Lipschitz-Bedingung erfüllt, muss diese Lösung des AWP aber eindeutig sein, siehe Satz 90 (p. 90). Wir folgern, dass $\hat{y}|_{(\tau-\delta, \tau+\delta)} = z|_{(\tau-\delta, \tau+\delta)}$. Wir schließen also auch, dass $(\tau - \delta, \tau + \delta) \subset A$. Das zeigt, dass es zu jedem $\tau \in A$ ein $\delta > 0$ gibt mit $(\tau - \delta, \tau + \delta) \subset A$. Mit anderen Worten: A ist offen. Da wir nun gezeigt haben, dass A offen ist und $O \setminus A$ offen ist, folgt die Zwischenbehauptung wie oben diskutiert direkt daraus, dass O zusammenhängend ist. Wir wissen nun also, dass $z|_O = \hat{y}|_O$. Wir zeigen nun, dass $J \subset \hat{I}$. Nehmen wir an, dass $J \not\subset \hat{I}$, so wäre $J \cup \hat{I} \supsetneq \hat{I}$. $J \cup \hat{I}$ wäre auch ein offenes Intervall, da $t_0 \in J$ und $t_0 \in \hat{I}$. Definiere nun $w : J \cup \hat{I} \rightarrow G$ durch

$$w(t) := \begin{cases} \hat{y}(t) & t \in \hat{I} \\ z(t) & t \in J \end{cases} \quad (691)$$

Beachte: Obwohl die Intervalle \hat{I} und J einen nichtleeren Schnitt haben ist w wohldefiniert — denn auf dem Schnitt $O = \hat{I} \cap J$ stimmen \hat{y} und z ja überein. Auch zeigt man leicht (HA), dass $w \in C^1(J \cup \hat{I}; G)$ eine Lösung von

$$\begin{cases} w'(t) = f(t, w(t)) & (t \in I_m \cup J) \\ w(t_0) = y_0 \end{cases} \quad (692)$$

ist. Damit wäre nach (691) w aber eine Fortsetzung von \hat{y} über auf $J \cup \hat{I} \supsetneq \hat{I}$. Eine solche Fortsetzung kann es aber nicht geben, weil \hat{y} die Eigenschaft aus Definition 118 (p. 119) erfüllen muss. Wir erhalten einen Widerspruch und schließen $J \cup \hat{I} = \hat{I}$, also $J \subset \hat{I}$. Insbesondere gilt auch $O = J \cap \hat{I} = J$ und da \hat{y} und z auf O übereinstimmen gilt auch $\hat{y}|_J = \hat{y}|_O = z$.

Final zeigen wir endlich den lokalen Satz von Picard-Lindelöf.

Wdh. Satz 94 (p. 91) SATZ VON PICARD-LINDELÖF, LOKALE VERSION. Es genüge $f \in C^0(I \times G; \mathbb{R}^n)$ einer lokalen Lipschitz-Bedingung. Dann gibt es zu jedem $(t_0, y_0) \in I \times G$ ein eindeutiges offenes Intervall $I_{max} = I_{max}(t_0, y_0) \subset I$ mit $t_0 \in I_{max}$ und eine eindeutige Funktion $y_{max} \in C^1(I_{max}; G)$ mit folgenden Eigenschaften:

- (Eindeutige Lösbarkeit) Das Anfangswertproblem

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I_{max}) \\ y(t_0) = y_0 \end{cases} \quad (693)$$

wird durch $y = y_{max}$ gelöst.

- (Maximalität) Falls $J \subset I$ ein offenes Intervall mit $t_0 \in J$ ist, sodass das AWP

$$\begin{cases} z'(t) = f(t, z(t)) & (t \in J) \\ z(t_0) = y_0 \end{cases} \quad (694)$$

eine Lösung $z \in C^1(J; G)$ besitzt, dann gilt $J \subset I_{max}$ und $z(t) = y_{max}(t)$ für alle $t \in J$.

Beweis von Satz 94 (p. 91). Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung.

Schritt 1: Existenz einer maximalen Lösung. Aus Proposition 119 (p. 119) folgt, dass es eine nicht-fortsetzbare Lösung gibt. Nach Lemma 120 (p. 122) erfüllt diese auch die Maximalitätsbedingung aus Satz 94.

Schritt 2: Eindeutigkeit der maximalen Lösung. Angenommen es gäbe zwei maximale Lösungen $y_{max}^1 \in C^1(I_{max}^1; \mathbb{R}^n)$ und $y_{max}^2 \in C^1(I_{max}^2; \mathbb{R}^n)$, die beide die Maximalitätsbedingung aus Satz 94 erfüllen. Verwenden wir die Maximalitätsbedingung von y_{max}^1 mit $J = I_{max}^2$ und $z = y_{max}^2$, so erhalten wir, dass $I_{max}^2 \subset I_{max}^1$ und $y_{max}^2(t) = y_{max}^1(t)$ für

alle $t \in I_{max}^2$. Mit Vertauschung der Rollen folgt $I_{max}^1 \subset I_{max}^2$ und $y_{max}^1(t) = y_{max}^2(t)$ für alle $t \in I_{max}^1$. Alles zusammen ergibt $I_{max}^1 = I_{max}^2$ und $y_{max}^1 = y_{max}^2$.

Endlich haben wir den Hauptsatz dieser Vorlesung — den lokalen Satz von Picard-Lindelöf — bewiesen. Zum Abschluss bleibt hier noch die folgende Frage: Kann man die lokale Lipschitz-Bedingung auch noch abschwächen? Gibt es noch eindeutige Lösungen, falls nur gilt $f \in C^0(I \times G; \mathbb{R}^n)$?

Was die Eindeutigkeit anbelangt, lautet die Antwort auf jeden Fall *nein*, denn das Anfangswertproblem

$$\begin{cases} y'(t) = 2\sqrt{|y(t)|} \\ y(0) = 0 \end{cases} \quad (695)$$

hat nach Beispiel 4 (p. 4) unendlich viele globale Lösungen, gegeben durch

$$y_a(t) := \begin{cases} 0 & t \leq a \\ (t-a)^2 & t > a \end{cases} \quad \text{für ein } a \geq 0. \quad (696)$$

Die Existenz von (nicht-fortsetzbaren) Lösungen kann aber noch gewährleistet werden:

Satz 121. SATZ VON PEANO.

Es sei $f \in C^0(I \times G; \mathbb{R}^n)$. Dann besitzt

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad (697)$$

eine nicht-fortsetzbare Lösung $y \in C^1(J; G)$ (für ein $J \subset I$).

Beweis. Ohne Beweis.

Man kann also in der Tat fast jede DGL lösen, die man hinschreiben kann.

3.2 Globalität und stetige Abhängigkeit

Wir haben nun Existenz und Eindeutigkeit einer maximalen Lösung gezeigt. Doch wie groß ist diese maximale Lösung genau? Wann hört sie auf zu existieren? Dieser Frage wollen wir zuerst nachgehen, indem wir das *Escape-Lemma* besprechen. Dieses zeigt, dass eine maximale Lösung jedes Kompaktum $K \subset G$ verlassen muss. Daraus wird sich dann unser eingangs erwähntes Maximalitätskriterium ableiten.

3.2.1 Das Escape-Lemma

Das folgende Lemma beruht auf einer wichtigen Beobachtung: Man kann unter einer lokalen Lipschitz-Bedingung das *Nichtexplodieren der Lipschitzkonstante* auf Kompakta nachweisen. D.h. auf jedem Kompaktum kann die Lipschitzkonstante bei der Lipschitzkonstante gleichmäßig beschränkt werden. Die nächste Proposition zeigt dieses Nicht-Explodieren der Lipschitzkonstante. Die Beweistechnik, die wir hierfür verwenden, trägt auch den Namen *Prinzip der schlechten Folgen*. Sie wird häufig verwendet

Proposition 122. GLEICHMÄSSIGE LIPSCHITZ-ABSCHÄTZUNG AUF KOMPAKTA. Es erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung. Es sei nun $K \subset G$ kompakt und $A \subset I$ kompakt. Dann gibt es $L^* = L^*(A, K) > 0$ mit

$$|f(t, y) - f(t, z)| \leq L^*|y - z| \quad \forall t \in A \quad \forall y, z \in K. \quad (698)$$

Beweis. Angenommen, die Aussage ist falsch. Dann kann es zu zwei speziell vorgegebenen Kompakta $A \subset I$ und $K \subset G$ kein $L^* > 0$ geben, sodass (698) erfüllt ist. Wir behaupten, dass es dann zu jedem $n \in \mathbb{N}$ ein $y_n \in K$ und $z_n \in K$ sowie $t_n \in A$ geben muss mit

$$|f(t_n, y_n) - f(t_n, z_n)| > n|y_n - z_n|. \quad (699)$$

[In der Tat: wäre dem nicht so, so gäbe es $n_0 \in \mathbb{N}$ mit $|f(t, y) - f(t, z)| \leq n_0|y - z|$ für alle $t \in A$ und $y, z \in K$. Dann wäre aber $L^* := n_0$ eine zulässige Lipschitzkonstante, was unserer Annahme widersprechen würde]. Man beachte, dass $y_n \neq z_n$ für alle $n \in \mathbb{N}$, denn $y_n = z_n$ widerspräche (699). Wegen der Kompaktheit von A und K gibt es nun eine Teilfolge $(l_n) \subset \mathbb{N}, l_n \rightarrow \infty$, sowie $t^* \in A, y^*, z^* \in K$ mit $t_{l_n} \rightarrow t^*$, $y_{l_n} \rightarrow y^*$ und $z_{l_n} \rightarrow z^*$ (HA). Wir behaupten nun, dass $y^* = z^*$. Dazu

$$\begin{aligned} |y^* - z^*| &= \lim_{n \rightarrow \infty} |y_{l_n} - z_{l_n}| = \lim_{n \rightarrow \infty} \frac{1}{l_n} (l_n |y_{l_n} - z_{l_n}|) \stackrel{(699)}{\leq} \liminf_{n \rightarrow \infty} \frac{1}{l_n} |f(t_{l_n}, y_{l_n}) - f(t_{l_n}, z_{l_n})| \\ &\leq \liminf_{n \rightarrow \infty} \frac{1}{l_n} (|f(t_{l_n}, y_{l_n})| + |f(t_{l_n}, z_{l_n})|) = \liminf_{n \rightarrow \infty} \frac{2}{l_n} \sup_{(\tau, w) \in A \times K} |f(\tau, w)| = 0, \end{aligned} \quad (700)$$

$$\leq \liminf_{n \rightarrow \infty} \frac{1}{l_n} (|f(t_{l_n}, y_{l_n})| + |f(t_{l_n}, z_{l_n})|) = \liminf_{n \rightarrow \infty} \frac{2}{l_n} \sup_{(\tau, w) \in A \times K} |f(\tau, w)| = 0, \quad (701)$$

da $l_n \rightarrow \infty$ und $\sup_{(\tau, w) \in A \times K} |f(\tau, w)| < \infty$, da $A \times K \subset I \times G$ kompakt ist (HA) und stetige Funktionen auf kompakten Mengen stets ein endliches Maximum annehmen. Es folgt $y^* = z^*$. Wir werden nun verwenden, dass wir in einer Umgebung von (t^*, y^*) eine lokale Lipschitzabschätzung haben: Es gibt nämlich ein $\epsilon = \epsilon(t^*, y^*) > 0$ und $r = r(t^*, y^*) > 0$ sowie $L = L(t^*, y^*) > 0$ mit

$$|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2| \quad \forall t \in (t^* - \epsilon, t^* + \epsilon) \quad \forall y_1, y_2 \in B_r(y_0). \quad (702)$$

Beachte: Da $y_{l_n}, z_{l_n} \rightarrow y^*$ und $t_{l_n} \rightarrow t^*$ gibt es auch $n_1 \in \mathbb{N}$ mit $y_{l_n}, z_{l_n} \in B_r(y_0)$ sowie $t_{l_n} \in (t^* - \epsilon, t^* + \epsilon)$ für alle $n \geq n_1$. Dann gilt mit (699)

$$l_n |y_{l_n} - z_{l_n}| \stackrel{(699)}{\leq} |f(t_{l_n}, y_{l_n}) - f(t_{l_n}, z_{l_n})| \stackrel{(702)}{\leq} L |y_{l_n} - z_{l_n}| \quad \forall n \geq n_1. \quad (703)$$

Da $y_{l_n} \neq z_{l_n}$ (siehe Anfang des Beweises), gilt also $l_n \leq L$ für alle $n \geq n_1$. Dies widerspricht der Tatsache, dass $l_n \rightarrow \infty$. Die Behauptung ist gezeigt.

Der Kniff ist nun: Bleibt die Lösung in einem Kompaktum, so lässt sich die lokale Lipschitz-Bedingung durch eine globale Lipschitz-Bedingung “ersetzen”. Da Lösungen bei einer globalen Lipschitzbedingung stets global sind, muss also in diesem Fall die Globalität

folgen. Man schließt daraus, dass eine maximale Lösung entweder global ist oder jedes Kom-
paktum verlassen muss. Hierzu ist zu sagen, dass diese Aussage *beidseitig* gelten muss, für
Details verweisen wir auf das folgende Lemma: Das sogenannte *Escape-Lemma*.

Lemma 123. ESCAPE-LEMMA. Es sei $I = (a, b)$, $f \in C^0(I \times G; \mathbb{R}^n)$ erfülle eine lokale
Lipschitz-Bedingung und $y \in C^1(I_{max}; G)$ die eindeutige maximale Lösung des AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I_{max}) \\ y(t_0) = y_0. \end{cases} \quad (704)$$

Ferner sei $I_{max} = (t_-, t_+)$ mit $t_- < t_0 < t_+$. Falls es ein $K \subset G$ kompakt gibt, sodass
 $y(t) \in K$ für alle $t \in [t_0, t_+]$, so gilt $t_+ = b$. Analog: Falls $y(t) \in K$ für alle $t \in (t_-, t_0]$, so
gilt $t_- = a$.

Beweis. Es sei $K \subset G$ kompakt mit $y(t) \in K$ für alle $t \in [t_0, t_+]$. Dann gibt es (HA)
 $\epsilon > 0$ mit

$$\text{dist}(x, \mathbb{R}^n \setminus G) \geq \epsilon \quad \forall x \in K. \quad (705)$$

Setze nun

$$B_{\frac{\epsilon}{2}}(K) := \{x \in \mathbb{R}^n : \text{dist}(x, K) < \frac{\epsilon}{2}\} \quad (706)$$

und

$$\tilde{K} := \{x \in \mathbb{R}^n : \text{dist}(x, K) \leq \frac{\epsilon}{2}\}. \quad (707)$$

Man zeigt leicht (HA), dass $B_{\frac{\epsilon}{2}}(K)$ offen und beschränkt ist und \tilde{K} kompakt ist. Ange-
nommen $t_+ < b$. Wähle $t_{++} \in (t_+, a)$ und setze $A := [t_0, t_{++}]$. Nach Proposition 122 (p.
126), erfüllt f dann eine Lipschitz-Abschätzung (wie (698)) auf $A \times \tilde{K}$, insbesondere also
auch auf der Teilmenge $(t_0, t_{++}) \times B_{\frac{\epsilon}{2}}(K)$. Das heißt, es gibt ein $L^* > 0$ mit

$$|f(t, y_1) - f(t, y_2)| \leq L^* |y_1 - y_2| \quad \forall t \in (t_0, t_{++}) \quad y_1, y_2 \in B_{\frac{\epsilon}{2}}(K). \quad (708)$$

Nach dem Satz von Kirschbraun (Proposition 114, p. 113) gibt es dann $F \in C^0((t_0, t_{++}) \times
\mathbb{R}^n; \mathbb{R}^n)$ derart, dass F eine globale Lipschitz-Bedingung erfüllt und $F|_{(t_0, t_{++}) \times B_{\frac{\epsilon}{2}}(K)} =
f|_{(t_0, t_{++}) \times B_{\frac{\epsilon}{2}}(K)}$. Wähle nun $t_1 \in [t_0, t_+)$. Das AWP

$$\begin{cases} z'(t) = F(t, z(t)) & (t \in (t_0, t_{++})) \\ z(t_1) = y(t_1) \end{cases} \quad (709)$$

hat nach der globalen Version von Picard-Lindelöf eine eindeutige Lösung
 $z \in C^1((t_0, t_{++}); \mathbb{R}^n)$. Beachte, dass für $t \in (t_0, t_+)$ gilt

$$y'(t) = f(t, y(t)) \underset{y(t) \in K}{=} F(t, y(t)). \quad (710)$$

Da auch $y(t_1) = y(t_1)$ gilt also

$$\begin{cases} y'(t) = F(t, y(t)) & (t \in (t_0, t_+)) \\ y(t_1) = y(t_1) \end{cases} \quad (711)$$

Deswegen (und wegen (709)) lösen sowohl $y|_{(t_0, t_+)}$ als auch $z|_{(t_0, t_+)}$ dasselbe AWP (711). Damit gilt wegen der Eindeutigkeit $y|_{(t_0, t_+)} = z|_{(t_0, t_+)}$. Insbesondere gilt $z(t) \in K$ für alle $t \in (t_0, t_+)$ und wegen der Abgeschlossenheit von K gilt daher auch $z(t) \in K$ für alle $t \in (t_0, t_+]$. Wähle nun $\tilde{t} \in (t_+, t_{++})$ sodass $z(t) \in B_{\frac{\epsilon}{2}}(K)$ für alle $t \in (t_0, \tilde{t})$. Ein solches \tilde{t} gibt es wegen der Stetigkeit von z und der Tatsache, dass $\text{dist}(K, B_{\frac{\epsilon}{2}}(K)^c) \geq \frac{\epsilon}{2} > 0$. Deswegen muss $z(t) \in B_{\frac{\epsilon}{2}}(K)$ noch “über t_+ hinaus” gelten. Wir definieren nun $w : (t_-, \tilde{t}) \rightarrow \mathbb{R}^n$ durch

$$w(t) := \begin{cases} y(t) & t \in (t_-, t_+), \\ z(t) & t \in (t_0, \tilde{t}). \end{cases} \quad (712)$$

Dass sich die beiden Fälle überschneiden spielt keine Rolle, denn wir wissen ja, dass y und z auf (t_0, t_+) übereinstimmen. Es gilt daher $w \in C^1((t_-, \tilde{t}); \mathbb{R}^n)$. Ferner gilt $w(t_0) = y(t_0) = y_0$ und für $t \in (t_-, t_+)$

$$w'(t) = y'(t) = f(t, y(t)) = f(t, w(t)), \quad (713)$$

sowie für $t \in (t_0, \tilde{t})$

$$w'(t) = z'(t) = F(t, z(t)) \underset{z(t) \in B_{\frac{\epsilon}{2}}(K)}{=} f(t, z(t)) = f(t, w(t)). \quad (714)$$

Daher hat man

$$\begin{cases} w'(t) = f(t, w(t)) & t \in (t_-, \tilde{t}) \\ w(t_0) = y_0, \end{cases} \quad (715)$$

d.h. w ist eine Lösung auf (t_-, \tilde{t}) . Dies ist ein Widerspruch zur Maximalität von y , da $(t_-, \tilde{t}) \supsetneq (t_-, t_+) = I_{max}$. Man kann analog vorgehen wenn man die spiegelbildliche Aussage für t_- zeigen will. Dies führen wir nicht im Detail vor.

Das eingangs formulierte Maximalitätskriterium (siehe Satz 96 (p. 93)) fällt sofort als Konsequenz heraus

Wdh. Satz 96 (p. 93). MAXIMALITÄTSKRITERIUM. Es seien $I = (a, b)$, $-\infty \leq a < b \leq \infty$, und $(t_0, y_0) \in I \times G$. Ferner erfülle $f \in C^0(I \times G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und es sei $J = (t_-, t_+) \subset I$. Es sei $y \in C^1(J; G)$ eine Lösung von

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in J) \\ y(t_0) = y_0 \end{cases} \quad (716)$$

Dann sind folgende Aussagen äquivalent:

1. $J = I_{max}$ und $y = y_{max}$.

2. Es gilt $t_+ = b$ oder

$$\liminf_{t \uparrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (717)$$

Ferner gilt $t_- = a$ oder

$$\liminf_{t \downarrow t_-} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (718)$$

Hierbei verwenden wir für den Fall $G = \mathbb{R}^n$ die Konvention $\text{dist}(z, \emptyset) = \infty$ für alle $z \in \mathbb{R}^n$.

Beweis von Satz 96 (p. 93). Es seien $I = (a, b)$, $f \in C^0(I \times G; \mathbb{R}^n)$ lokal Lipschitz und J, y wie in der Aussage. Wir zeigen zunächst $1. \Rightarrow 2.$ Dazu sei $y = y_{max} \in C^1(I_{max}; G)$ die maximale Lösung mit $I_{max} = (t_-, t_+)$. \mathbb{Z}_2 : Es gilt $t_+ = b$ oder

$$\liminf_{t \rightarrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (719)$$

Angenommen (719) gilt nicht. Dann gibt es $\epsilon \in (0, 1)$ und $\tilde{t} \in I_{max}$, $\tilde{t} < t_+$ mit

$$\min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} \geq \epsilon \quad \forall t \in (\tilde{t}, t_+), \quad (720)$$

d.h. für alle $t \in (\tilde{t}, t_+)$ hat man

$$\text{dist}(y(t), \mathbb{R}^n \setminus G) \geq \epsilon, \quad \text{und} \quad \frac{1}{|y(t)| + 1} \geq \epsilon \quad \left[\Leftrightarrow |y(t)| \leq \frac{1}{\epsilon} - 1 \right]. \quad (721)$$

Setze nun

$$K_\epsilon := \overline{B_{\frac{1}{\epsilon}-1}(0)} \cup \{z \in \mathbb{R}^n : \text{dist}(z, \mathbb{R}^n \setminus G) \geq \epsilon\}. \quad (722)$$

Man überzeugt sich mit (721) leicht davon, dass K_ϵ kompakt ist (HA) und $y(t) \in K_\epsilon$ für alle $t \in (\tilde{t}, t_+)$. Ferner gilt, dass $y([t_0, \tilde{t}])$ kompakt ist, denn y ist stetig auf $[t_0, \tilde{t}]$ und stetige Funktionen bilden kompakte Mengen stets auf kompakte Mengen ab. Setze also nun $K := y([t_0, \tilde{t}]) \cup K_\epsilon$. Die Menge K ist als Vereinigung kompakter Mengen wiederum kompakt (HA). Ferner gilt, wie eine leichte Fallunterscheidung zwischen $t \in [t_0, \tilde{t}]$ und $t \in (\tilde{t}, t_+)$ zeigt, dass

$$y(t) \in K \quad \forall t \in [t_0, t_+]. \quad (723)$$

Aus dem Escape Lemma (siehe) folgt nun, dass $t_+ = b$. Wir haben damit also gezeigt, dass $t_+ = b$ oder

$$\liminf_{t \rightarrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (724)$$

Mit exakt derselben Argumentation zeigen wir auch $t_- = a$ oder

$$\liminf_{t \rightarrow t_-} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (725)$$

Die Richtung $1. \Rightarrow 2.$ ist gezeigt. Nun zu $2. \Rightarrow 1.$ Falls $t_+ = b$ und $t_- = a$, so haben wir bereits eine größtmögliche Lösung, und somit ist klarerweise $J = (a, b) = I_{max}$ ein maximales Lösungsintervall. Falls

$$\liminf_{t \rightarrow t_-} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0 \quad (726)$$

so ist die Lösung y eine *nicht-fortsetzbare* Lösung, denn entweder ist sie unbeschränkt oder jede Fortsetzung müsste am Rand des Gebietes G beginnen. Beides ist nicht zulässig. Laut Lemma 120 (p. 122) ist nun aber jede nicht-fortsetzbare Lösung gleich die maximale Lösung.

Wir haben nun ein Kriterium kennengelernt, welches äquivalent zur Maximalität ist. Ferner kann es benutzt werden, um Globalität zu zeigen: Gilt für eine maximale Lösung

$$\liminf_{t \rightarrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^n \setminus G), \frac{1}{|y(t)| + 1} \right\} \neq 0, \quad (727)$$

so gilt $t_+ = b$, d.h. die Lösung existiert so lang wie nur möglich. Wir werden nun lernen, wie wir (727) in konkreten Anwendungen nachprüfen können, ohne die Lösung explizit zu kennen.

Beispiel 124. RECHTSGLOBALITÄT MIT HILFSFUNKTION. Wir betrachten das DGL-System

$$\begin{cases} y'_1(t) = -y_1(t)^3 + y_1(t)y_2(t)^2 \cosh(y_2(t)) \\ y'_2(t) = -y_2(t) - y_1(t)^2 y_2(t) \cosh(y_2(t)) \end{cases} \quad (728)$$

BEHAUPTUNG: Für alle $y_0 = (y_{1,0}, y_{2,0})^T \in \mathbb{R}^2$ gibt es eine maximale Lösung $y : I_{max} \rightarrow \mathbb{R}^2$ von (728) mit $y(0) = y_0$ und $I_{max} \supset [0, \infty)$.

BEGRÜNDUNG.

SCHRITT 1. Existenz einer maximalen Lösung. Hierfür müssen wir nachrechnen, dass $f : \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definiert durch

$$f(t, (z_1, z_2)) := \begin{pmatrix} -z_1^3 + z_1 z_2^2 \cosh(z_2) \\ -z_2 - z_1^2 z_2 \cosh(z_2) \end{pmatrix} \quad (729)$$

Eine lokale Lipschitzbedingung erfüllt. Hierfür genügt es nach Proposition 97 (p. 94) zu

zeigen, dass f in beiden Variablen z_1, z_2 stetig differenzierbar ist. Dies ist den Lesenden als Übungsaufgabe überlassen. Mit der lokalen Version von Picard-Lindelöf schließen wir die Existenz einer eindeutigen maximalen Lösung von

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(0) = (y_{0,1}, y_{0,2}). \end{cases} \quad (730)$$

Diese sei im Folgenden gegeben durch $y \in C^1(I_{max}; \mathbb{R}^2)$.

SCHRITT 2. Rechtsglobalität. $\mathbb{Z}\mathbb{Z}$: $I_{max} \supset [0, \infty)$. Es sei $I_{max} = (t_-, t_+)$ mit $t_- < 0 < t_+$, da $t_0 = 0$ ja der Anfangswert ist. Angenommen $t_+ < \infty$. Dann gilt nach (731)

$$\liminf_{t \rightarrow t_+} \min \left\{ \text{dist}(y(t), \mathbb{R}^2 \setminus G), \frac{1}{|y(t)| + 1} \right\} = 0. \quad (731)$$

Da $G = \mathbb{R}^2$ ist der erste Term in dem Minimumsausdruck konstant gleich unendlich und damit für die Minimumsbildung in (731) vernachlässigbar. Wir schließen

$$\liminf_{t \rightarrow t_+} \frac{1}{|y(t)| + 1} = 0, \quad (732)$$

d.h.

$$\limsup_{t \rightarrow t_+} |y(t)| = \infty. \quad (733)$$

Betrachte nun $h(t) := |y(t)|^2 = y_1(t)^2 + y_2(t)^2$ und berechne

$$\begin{aligned} h'(t) &= 2(y_1(t)y'_1(t) + y_2(t)y'_2(t)) \\ &\stackrel{(728)}{=} 2(y_1(t)(-y_1(t)^3 + y_1(t)y_2(t)^2 \cosh(y_2(t))) + y_2(t)(-y_2(t) - y_1(t)^2 y_2(t) \cosh(y_2(t)))) \\ &= 2(-y_1(t)^4 + y_1(t)^2 y_2(t)^2 \cosh(y_2(t)) - y_2(t)^2 - y_1(t)^2 y_2(t)^2 \cosh(y_2(t))) \\ &= -2y_1(t)^4 - 2y_2(t)^2 \leq 0. \end{aligned}$$

Damit ist h monoton fallend und es gilt für alle $t \in (0, t_+)$

$$|y(t)|^2 = h(t) \leq h(0) = |y_0|^2 = y_{0,1}^2 + y_{0,2}^2. \quad (734)$$

Insbesondere gilt

$$|y(t)| \leq \sqrt{y_{0,1}^2 + y_{0,2}^2} \quad \forall t \in (0, t_+). \quad (735)$$

Daher gilt auch

$$\limsup_{t \rightarrow t_+} |y(t)| \leq \sqrt{y_{0,1}^2 + y_{0,2}^2}, \quad (736)$$

ein Widerspruch zu (733). Es folgt, dass die Annahme falsch sein muss und daher gilt $t_+ = \infty$. Die Behauptung ist gezeigt. Bei dem Beweis war die Monotonie der Hilfsfunktion h von entscheidender Bedeutung. Diese muss man leider für jedes Problem aufs Neue “erraten”. In manchen Fällen kann aber auch ein *Integral der Bewegung* die Rolle einer solchen Hilfsfunktion übernehmen, siehe Übung 100 (p. 96).

***Übung 125.** RECHTSGLOBALITÄT DURCH KONTROLLE DER NORM. Betrachten Sie für einen beliebigen Anfangswert $y_0 = (y_{0,1}, y_{0,2})^T \in \mathbb{R}^2$ das Anfangswertproblem

$$\begin{cases} y'_1(t) = -y_2(t) \\ y'_2(t) = y_1(t) + (1 - y_1^2(t) - y_2^2(t))y_2(t). \\ (y_1(0), y_2(0)) = (y_{0,1}, y_{0,2}). \end{cases} \quad (737)$$

Zeigen Sie: Es existiert eine maximale Lösung $y_{max} = (y_1, y_2) : I_{max} \rightarrow \mathbb{R}^2$, die auch rechtsglobal ist, d.h. $I_{max} \supset [0, \infty)$.

HINWEIS: Betrachten Sie $f'(t)$ für $f(t) := |y(t)|^2$.

SINN DER ÜBUNG. Hier sieht man, dass man eine Kontrolle der Norm auch erreichen kann, wenn sie nicht zwingend monoton fallend ist.

Übung 126. LYAPUNOW-FUNKTIONEN. Es erfülle $f \in C^0(\mathbb{R} \times \mathbb{R}^n; \mathbb{R}^n)$ eine lokale Lipschitzbedingung. Eine Funktion $V \in C^1(\mathbb{R}^n; \mathbb{R})$ heißt *Lyapunow-Funktion* für f , falls

- $\text{grad}V(z) \cdot f(t, z) \leq 0$ für alle $(t, z) \in \mathbb{R} \times \mathbb{R}^n$
- $V^{-1}((-\infty, c])$ ist kompakt für alle $c \in \mathbb{R}$.

1. Es sei nun für ein beliebiges $y_0 \in \mathbb{R}^n$ das AWP

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad (738)$$

gegeben, wobei f eine Lyapunow-Funktion besitzt. Zeigen Sie, dass $I_{max} \supset [t_0, \infty)$.

2. Es sei $E \in C^1(\mathbb{R}^n; \mathbb{R})$ und

$$f(t, z) = -\text{grad } E(z) \quad \forall (t, z) \in \mathbb{R}^n. \quad (739)$$

Zeigen Sie, dass E eine Lyapunow-Funktion für f ist, falls $E^{-1}((-\infty, c])$ kompakt ist.

SINN DER ÜBUNG. Hier wird die Technik aus dem vorigen Beispiel verallgemeinert: Man zeigt Globalität, indem man die Bedingung (731) ausschließt. Um dies tun zu können, muss man das *Explodieren* der Lösung verhindern. Im vorigen Beispiel haben wir das Explodieren dadurch verhindert, dass die euklidische Norm entlang einer Lösung abnimmt.

In dieser Übung werden wir dies verallgemeinern: Die allgemeine Lyapunow-Funktion V nimmt entlang der Lösung ab. Ein großer Anwendungsbereich sind hier *Gradientenflüsse*, d.h. Lösungen von

$$y'(t) = -\text{grad } E(y(t)). \quad (740)$$

3.2.2 Das Lemma von Gronwall

Wir haben im letzten Abschnitt gesehen, dass wir die Globalität dadurch zeigen können, dass Kompakta in endlicher Zeit nicht verlassen werden. Um das Verlassen von Kompakta auszuschließen, müssen wir das *Wachstum von Lösungen* verstehen. Ein fundamentales Tool hierbei ist das *Lemma von Gronwall*, benannt nach dem schwedisch-amerikanischen Mathematiker *T.H. Grönwall (1877-1932)*. Mit ihm kann man nicht nur das Wachstum von Lösungen verstehen, sondern auch das Wachstum von *Funktionen von Lösungen*. Dies erlaubt uns zum Beispiel auch, zu quantifizieren, wie stark sich Lösungen zu verschiedenen Anfangswerten unterscheiden.

Lemma 127. DAS LEMMA VON GRONWALL. Es seien $u_0, t_0, T \in \mathbb{R}$, $t_0 < T$, und $u, v \in C([t_0, T]; \mathbb{R})$ zwei Funktionen mit $v(t) \geq 0$ für alle $t \in [t_0, T]$. Es gelte für alle $t \in [t_0, T]$

$$u(t) \leq u_0 + \int_{t_0}^t u(s)v(s) \, ds. \quad (741)$$

Dann folgt für alle $t \in [t_0, T]$

$$u(t) \leq u_0 \exp \left(\int_{t_0}^t v(s) \, ds \right). \quad (742)$$

Beweis. Es sei $\varepsilon > 0$. Betrachten wir die Funktion $h_\varepsilon : [t_0, T] \rightarrow \mathbb{R}$ gegeben durch

$$h_\varepsilon(t) = (u_0 + \varepsilon) \exp \left(\int_{t_0}^t v(s) \, ds \right). \quad (743)$$

Wir stellen fest, dass $h_\varepsilon(t_0) = u_0 + \varepsilon$ und

$$h'_\varepsilon(t) = (u_0 + \varepsilon) \exp \left(\int_{t_0}^t v(s) \, ds \right) v(t) = v(t)h_\varepsilon(t) \quad (744)$$

Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$h_\varepsilon(t) = h_\varepsilon(t_0) + \int_{t_0}^t h'_\varepsilon(s) \, ds = (u_0 + \varepsilon) + \int_{t_0}^t h_\varepsilon(s)v(s) \, ds. \quad (745)$$

Zwischenbehauptung. Wir behaupten, dass $u(t) \leq h_\varepsilon(t)$ für alle $t \in [t_0, T]$.

Beweis der Zwischenbehauptung. Wir führen den Beweis durch Widerspruch. Angenommen, es gäbe ein $\tilde{t} \in [t_0, T]$ mit $u(\tilde{t}) > h_\varepsilon(\tilde{t})$. Wir arbeiten mit dem *Prinzip des*

kleinsten Verbrechers, d.h. wir setzen

$$\tau := \inf\{t \in [t_0, T] : u(t) > h_\varepsilon(t)\}. \quad (746)$$

Dies ist wohldefiniert, da wegen unserer Annahme das Infimum über eine nichtleere Menge gebildet wird. Beachte, dass $\tau > t_0$, denn mit $t = t_0$ in (741) sieht man $u(t_0) \leq u_0 < u_0 + \varepsilon = h_\varepsilon(t_0)$, weswegen aus Stetigkeitsgründen $u(t) < h_\varepsilon(t)$ für $t \in [t_0, t_0 + \delta]$, $\delta > 0$ klein genug, gilt. Da τ der kleinste Zeitpunkt des Übersteigens ist, muss gelten, dass

$$u(t) \leq h_\varepsilon(t) \quad \text{für alle } t \in [t_0, \tau]. \quad (747)$$

Dies impliziert wegen der Stetigkeit auch, dass $u(\tau) \leq h_\varepsilon(\tau)$. Da u aber h_ε direkt nach τ übersteigen wird (siehe (746)), gilt sogar

$$u(\tau) = h_\varepsilon(\tau). \quad (748)$$

Andererseits gilt wegen der Annahme (741)

$$u(\tau) \leq u_0 + \int_{t_0}^{\tau} u(s)v(s) \, ds \stackrel{(747), v \geq 0}{=} u_0 + \int_{t_0}^{\tau} h_\varepsilon(s)v(s) \, ds \quad (749)$$

$$< u_0 + \varepsilon + \int_{t_0}^{\tau} h_\varepsilon(s)v(s) \, ds \stackrel{(745)}{=} h_\varepsilon(\tau). \quad (750)$$

Dies zeigt $u(\tau) < h_\varepsilon(\tau)$, was ein offensichtlicher Widerspruch zu (748) ist. Die Zwischenbehauptung ist gezeigt. Wir folgern daher für alle $t \in [t_0, T]$ und $\varepsilon > 0$

$$u(t) \leq h_\varepsilon(t) = (u_0 + \varepsilon) \exp\left(\int_{t_0}^t v(s) \, ds\right) \quad (751)$$

Lassen wir $\varepsilon \rightarrow 0$ streben, so erhalten wir (742).

Übung 128. LINEARE DIFFERENTIALGLEICHUNGEN.

Es seien $a, b \in C^0(I; [0, \infty))$ und $z \in C^1(I; \mathbb{R})$ so, dass

$$\begin{cases} z'(t) \leq a(t)z(t) + b(t) & \forall t \in I. \\ z(t_0) \leq y_0 \end{cases} \quad (752)$$

Ferner sei $y \in C^1(I, \mathbb{R})$ die eindeutige Lösung von

$$\begin{cases} y'(t) = a(t)y(t) + b(t) & (t \in I). \\ y(t_0) = y_0 \end{cases} \quad (753)$$

Zeigen Sie, dass $z(t) \leq y(t)$ für alle $t \in I$ mit $t \geq t_0$.

HINWEIS. Zeigen Sie eine Abschätzung der Form wie in (741) für $u := z - y$

SINN DER ÜBUNG. Dies ist ein wichtiges Tool, um Lösungen abzuschätzen. Wir werden es im Laufe dieser Vorlesung zwar nicht benutzen, aber Differentialungleichungen haben in der Forschung durchaus eine große Bedeutung. Manchmal wird in der Literatur das Gronwall-Lemma so formuliert!

Wir wollen nun lernen, das Gronwall-Lemma zu benutzen, um Globalität von Lösungen zu zeigen. Wir werden in der folgenden Proposition eine Wachstumsabschätzung für f diskutieren, die benutzt werden kann, um eine Blow-Up-Situation auszuschließen.

Proposition 129. GLOBALITÄT DURCH LINEARES WACHSTUM. Es erfülle $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und sei $t_0 \in I$, $y_0 \in \mathbb{R}^n$. Ferner gebe es $c \in C^0(I; [0, \infty))$ mit

$$|f(s, z)| \leq c(s)(1 + |z|) \quad \forall s \in I \quad \forall z \in \mathbb{R}^n. \quad (754)$$

Es sei das AWP

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (755)$$

gegeben. Dann gibt es eine eindeutige globale Lösung $y \in C^1(I; \mathbb{R}^n)$.

Beweis. Es sei $I = (a, b)$. Existenz einer eindeutigen maximalen Lösung $y \in C^1(I_{max}; \mathbb{R}^n)$ folgt aus der lokalen Lipschitz-Bedingung, siehe Satz 94 (p. 91). Es sei $I_{max} = (t_-, t_+) \subset (a, b)$.

Schritt 1. Wir zeigen $t_+ = b$. Angenommen $t_+ < b$. Das heißt auch insbesondere $t_+ < \infty$. Dann gilt nach dem Maximalitätskriterium (Satz 96, p. 93) (angewendet auf den Spezialfall $G = \mathbb{R}^n$)

$$\liminf_{t \rightarrow t_+} \frac{1}{|y(t)| + 1} = 0, \quad (756)$$

d.h.

$$\limsup_{t \rightarrow t_+} |y(t)| = \infty. \quad (757)$$

Wir werden nun mithilfe des Gronwall-Lemmas zeigen, dass (757) nicht eintreten kann. Wähle $T \in [t_0, t_+)$ beliebig. Wir machen für $t \in [t_0, T]$ folgende Beobachtung

$$|y(t)| = |(y(t) - y_0) + y_0| \leq |y(t) - y_0| + |y_0| \quad (758)$$

$$= |y_0| + \left| \int_{t_0}^t y'(s) \, ds \right| \leq |y_0| + \left| \int_{t_0}^t f(s, y(s)) \, ds \right| \quad (759)$$

$$\leq |y_0| + \int_{t_0}^t |f(s, y(s))| \, ds = |y_0| + \int_{t_0}^t c(s)(1 + |y(s)|) \, ds \quad (760)$$

$$= |y_0| + \int_{t_0}^t c(s) \, ds + \int_{t_0}^t c(s)|y(s)| \, ds \quad (761)$$

$$\leq |y_0| + \int_{t_0}^{t_+} c(s) \, ds + \int_{t_0}^t |y(s)|c(s) \, ds. \quad (762)$$

Wir sehen nun, dass (741) erfüllt ist für $u(t) := |y(t)|$, $v(t) := c(t)$ und $u_0 := |y_0| + \int_{t_0}^{t_+} c(s) \, ds$. Man beachte, dass $u_0 \in \mathbb{R}$, da c auf dem kompakten Intervall $[t_0, t_+]$ stetig ist und daher beschränkt ist — weswegen $\int_{t_0}^{t_+} c(s) \, ds < \infty$ sofort aus $t_+ < \infty$ folgt. Für diese Wahlen von u, v, u_0 gilt nun also

$$u(t) \leq u_0 + \int_{t_0}^t u(s)v(s) \, ds. \quad \forall t \in [t_0, T] \quad (763)$$

Mit dem Lemma von Gronwall 127 folgt

$$u(t) \leq u_0 \exp\left(\int_{t_0}^t v(s) \, ds\right) \quad \forall t \in [t_0, T], \quad (764)$$

d.h.

$$|y(t)| \leq \left(|y_0| + \int_{t_0}^{t_+} c(s) \, ds\right) \exp\left(\int_{t_0}^t c(s) \, ds\right) \quad \forall t \in [t_0, T]. \quad (765)$$

Da $T < t_+$ beliebig war, folgern wir

$$|y(t)| \leq \left(|y_0| + \int_{t_0}^{t_+} c(s) \, ds\right) \exp\left(\int_{t_0}^t c(s) \, ds\right) \quad \forall t \in [t_0, t_+). \quad (766)$$

Daher berechnen wir

$$\limsup_{t \rightarrow t_+} |y(t)| \leq \left(|y_0| + \int_{t_0}^{t_+} c(s) \, ds\right) \exp\left(\int_{t_0}^{t_+} c(s) \, ds\right) < \infty, \quad (767)$$

ein Widerspruch zu (757). Dies zeigt $t_+ = b$.

Schritt 2. Wir zeigen $t_- = a$. Betrachte dazu das AWP

$$\begin{cases} z'(t) = -f(-t, z(t)) \\ z(-t_0) = y_0. \end{cases} \quad (768)$$

Noch präziser

$$\begin{cases} z'(t) = g(t, z(t)) \\ z(-t_0) = y_0, \end{cases} \quad (769)$$

wobei $g \in C^0((-b, -a) \times \mathbb{R}^n; \mathbb{R}^n)$ definiert ist durch $g(t, w) := -f(-t, w)$. Unser neues Intervall auf dem das AWP gestellt ist, ist also $(-b, -a)$. Man zeigt leicht, dass g auf $(-b, -a) \times \mathbb{R}^n$ eine lokale Lipschitz-Bedingung erfüllt und es daher eine eindeutige maximale Lösung auf einem Teilintervall von $(-b, -a)$ gibt. Beachte nun, dass $z(t) := y(-t)$, $t \in (-t_+, -t_-)$, eine Lösung von (768) ist. In der Tat:

$$z'(t) = -y'(-t) = -f(-t, y(-t)) = -f(-t, z(t)), \quad (770)$$

$$z(-t_0) = y(t_0) = y_0. \quad (771)$$

Zwischenbehauptung. Die Lösung $z : (-t_+, t_-) \rightarrow \mathbb{R}^n$, $z(t) = y(-t)$ ist auch die maximale Lösung für (768). In der Tat: Es sei $\tilde{z} \in C^1((c, d); \mathbb{R}^n)$ eine weitere Lösung von (768). Definiere $\tilde{y} : (-d, -c) \rightarrow \mathbb{R}^n$

$$\tilde{y}(t) := \tilde{z}(-t) \quad \text{für } t \in (-d, -c). \quad (772)$$

Dann gilt

$$\tilde{y}'(t) = -\tilde{z}'(-t) = -[-f(t, \tilde{z}(-t))] = f(t, \tilde{y}(t)), \quad (773)$$

d.h. \tilde{y} löst (755). Damit ist — aufgrund der Maximalität von $y - \tilde{y}$ eine Einschränkung von y , d.h. $(-d, -c) \subset (t_-, t_+)$ und $\tilde{y}(t) = y(t)$ für alle $t \in (-d, -c)$. Wir folgern, dass $(c, d) \subset (-t_+, -t_-)$ und für alle $t \in (c, d)$ gilt

$$\tilde{z}(t) = \tilde{y}(-t) = y(-t) = z(t). \quad (774)$$

Es folgt die Maximalität von z . Daher ist $(-t_+, -t_-)$ auch das maximale Existenzintervall für (768). Nehmen wir nun an, dass $t_- > a$. Dann gilt $-t_- < -a$. Allerdings erfüllt die Funktion g aus (769)

$$|g(s, z)| = |-f(-s, z)| = |f(-s, z)| \leq c(-s)(1 + |z|) \underset{\tilde{c}(s) := c(-s)}{=} \tilde{c}(s)(1 + |z|) \quad (775)$$

Für ein $\tilde{c} := c(-(\cdot)) \in C^0((-b, -a); [0, \infty))$. Aus Beweisschritt 1 folgt, dass dann die maximale Lösung von (769) bis zum größtmöglichen Zeitpunkt $-a$ existieren muss. Dies zeigt

$$-t_- = \text{Maximalpunkt der maximalen Lösung von (769)} = -a, \quad (776)$$

mit anderen Worten $t_- = a$. Die Behauptung ist gezeigt.

Beweisreflexion 130. Die Technik aus Schritt 2 des vorigen Beweises kann allgemein als *Umkehren der Zeit* beschrieben werden. Wollen wir Aussagen über den Minimalpunkt t_- machen, so können wir stattdessen die *zeitumgekehrte Gleichung* (768) betrachten und Aussagen über den Maximalpunkt der Lösung der zeitumgekehrten Gleichung machen.

Der Maximalpunkt der zeitumgekehrten Gleichung ist dann — wie wir im Beweis gesehen haben — exakt das Negative vom Minimalpunkt der ursprünglichen Gleichung. Somit müssen wir eigentlich nur Maximalpunkte studieren. Daher kümmert es uns auch nicht, dass wir im Lemma von Gronwall $T > t_0$ vorraussetzen müssen.

Die Globalitätsbedingung (754) nennen wir ab sofort *lineare Wachstumsbedingung*. Sie trifft automatisch zu, wenn die Funktion f eine globale Lipschitz-Bedingung erfüllt. Daher kann die Globalitätsaussage in der vorigen Proposition auch als eine Erweiterung der Globalitätsaussage im globalen Satz von Picard-Lindelöf verstanden werden.

Übung 131. LIPSCHITZBEDINGUNG UND LINEARES WACHSTUM. Es erfülle $f \in C^0(I \times$

$\mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung. Zeigen Sie: Dann gibt es $c \in C^0(I; [0, \infty))$ mit

$$|f(s, z)| \leq c(s)(1 + |z|) \quad \forall s \in I \quad \forall z \in \mathbb{R}^n. \quad (777)$$

SINN DER ÜBUNG: Man sieht, dass das lineare Wachstumskriterium der vorigen Proposition die Globalitätsaussage im globalen Picard-Lindelöf erweitert.

Jetzt sollten wir vielleicht noch feststellen, dass die Menge von Funktionen, die die lineare Wachstumsbedingung erfüllen auch *strikt größer* ist als die Menge der Funktionen, die eine globale Lipschitz-Bedingung erfüllen. Ansonsten hätte Proposition 129 ja keinen Mehrwert für uns. In der nächsten Übung sehen wir also ein Beispiel für eine nicht global Lipschitzstetige Funktion, die aber eine lineare Wachstumsbedingung erfüllt.

Übung 132. ANWENDUNG DES GLOBALITÄTSKRITERIUMS. Gegeben sei die DGL

$$y'(t) = \sin(y(t)^2) + t \quad (778)$$

- (a) Zeigen Sie: $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $f(t, z) = \sin(z^2) + t$ erfüllt eine lokale Lipschitz-Bedingung, aber keine globale Lipschitz-Bedingung.
- (b) Zeigen Sie: Für alle $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}$ gilt, dass das AWP

$$\begin{cases} y'(t) = \sin(y(t)^2) + t & t \in \mathbb{R} \\ y(t_0) = y_0 \end{cases} \quad (779)$$

eine eindeutige globale Lösung $y \in C^1(\mathbb{R}; \mathbb{R})$ besitzt.

SINN DER ÜBUNG. Bislang war nur eine globale Lischitz-Bedingung ein “Blankoscheck” für die Globalität einer Lösung. Im Falle von nicht global Lipschitzstetigen Funktionen gibt es einen solchen Blankoscheck nicht: Wir mussten stets mit der Struktur der Gleichung arbeiten, um Escape-Situation und Blow-Up-Situation auszuschließen. Die lineare Wachstumskriterium (754) ist nun ein weiterer “Blankoscheck”. In der Übung soll man sehen, dass mit dieser Bedingung tatsächlich eine größere Klasse von Funktionen untersucht werden kann.

Anwendung des Globalitätskriteriums auf lineare Gleichungen.

Wir sind nun endlich in der Lage, Satz 37 (p. 32) zu zeigen, den wir lange aufgeschoben haben. Dieser Satz besagt, dass jede lineare Differentialgleichung der Form

$$y'(t) = A(t)y(t) \quad (t \in I) \quad (780)$$

wobei $A \in C^0(I; \mathbb{C}^{n \times n})$ eine *Fundamentalmatrix* besitzt. Zur Erinnerung: Eine Fundamentalmatrix ist eine $n \times n$ -matrixwertige Abbildung $t \mapsto M(t)$, $t \in I$ mit den folgenden Eigenschaften

1. $M(t)$ ist invertierbar für alle $t \in I$.

2. Die allgemeine Lösung von (780) ist gegeben durch

$$\mathbb{L} = \{y(t) = M(t)c, (t \in I) \mid c \in \mathbb{C}^n\}. \quad (781)$$

Wir werden den Satz hier nur für *reellwertige* Matrixabbildungen $A \in C^0(I; \mathbb{R}^{n \times n})$ zeigen. Für komplexwertige Matrixabbildungen könnten wir aber genauso vorgehen — wir müssten nur den Satz von Picard-Lindelöf nur im Komplexen formulieren. Dies ist ohne Abänderungen möglich. Wir verzichten an dieser Stelle aber aus pädagogischen Gründen darauf.

Wdh. Satz 37, p. 32 (für reellwertige Matrizen).

EXISTENZ EINER FUNDAMENTALMATRIX. Es sei $A \in C^0(I; \mathbb{R}^{n \times n})$. Zu jedem linearen DGL-System der Form $y'(t) = A(t)y(t), (t \in I)$ gibt es eine Fundamentalmatrix $t \mapsto M(t)$ in $C^1(I; \mathbb{R}^{n \times n})$. Insbesondere bildet die Menge aller Lösungen einen ($\mathbb{R}-$)Vektorraum der Dimension n . Ferner hat für alle $y_0 \in \mathbb{R}^n$ und $t_0 \in I$ das Anfangswertproblem

$$\begin{cases} y'(t) = A(t)y(t) & (t \in I) \\ y(t_0) = y_0 \end{cases} \quad (782)$$

eine eindeutige Lösung, gegeben durch

$$y(t) = M(t)M(t_0)^{-1}y_0 \quad (t \in I). \quad (783)$$

Beweis von Satz 37, p. 32 (für reellwertige Matrizen). Wie vorhin angesprochen machen wir die (unwesentliche) Vereinfachung, dass $A \in C^0(I; \mathbb{R}^{n \times n})$ eine reellwertige Matrixabbildung ist.

Schritt 1: Beschaffung linear unabhängiger Lösungen auf I . Wähle nun $t_0 \in I$ beliebig. Für $i \in \{1, \dots, n\}$ betrachte nun die Anfangswertprobleme

$$\begin{cases} y'_i(t) = A(t)y_i(t) \\ y_i(t_0) = e_i, \end{cases} \quad (784)$$

wobei e_i wie immer der i -te Einheitsvektor ist. Wir behaupten, dass es für alle $i = 1, \dots, n$ Lösungen $y_i \in C^1(I; \mathbb{R}^n)$ von (784) gibt. Insbesondere muss also gezeigt werden, dass die maximalen Lösungen y_i von (784) global, d.h. auf ganz I definiert, sind. Hierzu verwenden wir das Kriterium aus Proposition 129. Es sei $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ gegeben durch

$$f(t, z) = A(t)z. \quad (785)$$

Mit Proposition 97 (p. 94) zeigt man leicht, dass f einer lokalen Lipschitz-Bedingung genügt. (Beachte hierfür, dass die Funktion f in Proposition 97 nur nach der z -Variablen differenzierbar sein muss — Differenzierbarkeit nach der t -Variablen wird nicht benötigt). Wir müssen zeigen, dass f das Kriterium (754) erfüllt, d.h. es existiert $c \in C^0(I; [0, \infty))$ mit

$$|f(t, z)| = |A(t)z| \leq c(t)(1 + |z|) \quad \forall t \in I \quad \forall z \in \mathbb{R}^n. \quad (786)$$

Dazu schreiben wir $A(t) = (a_{jk}(t))_{j,k=1}^n$, $t \in I$, wobei $a_{kj} \in C^0(I; \mathbb{R})$ die einzelnen Einträge sind. Dann gilt nach Definition der Matrix-Vektor-Multiplikation

$$f_j(t, z) = \sum_{k=1}^n a_{jk}(t) z_k \quad (787)$$

Mit der Dreiecksungleichung und der Cauchy-Schwarz-Ungleichung finden wir

$$|f_j(t, z)| = \left| \sum_{k=1}^n a_{jk}(t) z_k \right| \leq \sum_{k=1}^n |a_{jk}(t)| |z_k| \quad (788)$$

$$\leq \sqrt{\sum_{k=1}^n |a_{jk}(t)|^2} \sqrt{\sum_{k=1}^n |z_k|^2} = |z| \sqrt{\sum_{k=1}^n |a_{jk}(t)|^2}. \quad (789)$$

Daher gilt

$$|f(t, z)| \leq \sqrt{\sum_{j=1}^n |f_j(t, z)|^2} \stackrel{(789)}{\leq} \sqrt{\sum_{j=1}^n \left(|z|^2 \sum_{k=1}^n |a_{jk}(t)|^2 \right)} \quad (790)$$

$$= |z| \sqrt{\sum_{k,j=1}^n |a_{jk}(t)|^2} = c(t) |z| \leq c(t)(1 + |z|), \quad (791)$$

wobei wir hier

$$c(t) := \sqrt{\sum_{k,j=1}^n |a_{jk}(t)|^2} \quad (792)$$

gesetzt haben. Wir haben (754) damit nachgewiesen und somit folgt die Globalität der Lösungen y_i für alle $i = 1, \dots, n$.

Schritt 2. Konstruktion der Fundamentalmatrix. Wir definieren nun für $t \in I$ die Matrix

$$M(t) := (y_1(t) | \dots | y_n(t)) \in \mathbb{R}^{n \times n} \quad (793)$$

und behaupten, dass M eine Fundamentalmatrix ist.

Schritt 2.1: Beweis der Invertierbarkeit. Dazu zeigen wir zunächst, dass $M(t)$ für alle $t \in I$ invertierbar ist. Hierzu fixieren wir $t_1 \in I$ beliebig. Nach einem Satz aus der linearen Algebra ist die Invertierbarkeit von $M(t_1)$ äquivalent zu der Tatsache, dass alle Spalten von $M(t_1)$, sprich $\{y_1(t_1), \dots, y_n(t_1)\}$ linear unabhängig sind. Dies zeigen wir im Folgenden. Angenommen es gibt $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit

$$\sum_{i=1}^n \lambda_i y_i(t_1) = 0. \quad (794)$$

Zu zeigen ist, dass $\lambda_1 = \dots = \lambda_n = 0$. Wir behaupten zunächst, dass gilt

$$\sum_{i=1}^n \lambda_i y_i(t) = 0 \quad \forall t \in I. \quad (795)$$

Hierfür definiere man

$$\xi(t) := \sum_{i=1}^n \lambda_i y_i(t) \quad (796)$$

Wir werden im Folgenden nachrechnen, dass

$$\begin{cases} \xi'(t) = A(t)\xi(t) & (t \in I) \\ \xi(t_1) = 0 \end{cases} \quad (797)$$

Hieraus würde bereits folgen, dass $\xi \equiv 0$. In der Tat: die Nullfunktion ist eine Lösung desselben AWP (797). Da die Lösung von (797) aber eindeutig sein muss, folgt sofort $\xi \equiv 0$. Nun zeigen wir noch (797). Dazu:

$$\xi'(t) = \sum_{i=1}^n \lambda_i y'_i(t) \stackrel{(784)}{=} \sum_{i=1}^n \lambda_i A(t)y_i(t) = A(t) \left(\sum_{i=1}^n \lambda_i y_i(t) \right) = A(t)\xi(t). \quad (798)$$

Dass $\xi(t_1) = 0$ ist folgt aus (794). Wir haben (795) gezeigt. Setzen wir in (795) aber $t = t_0$ ein und verwenden, dass nach (784) $y_i(t_0) = e_i$, so erhalten wir

$$0 = \sum_{i=1}^n \lambda_i y_i(t_0) = \sum_{i=1}^n \lambda_i e_i. \quad (799)$$

Daraus folgt also $(\lambda_1, \dots, \lambda_n) = 0$, d.h. $\lambda_1 = \dots = \lambda_n = 0$. Die Behauptung ist gezeigt.

Schritt 2.2: Beweis von (781). Zunächst zeigen wir, dass für alle $c = (c_1, \dots, c_n) \in \mathbb{R}^n$ die Funktion $y(t) := M(t)c$, $(t \in I)$, eine Lösung der DGL definiert. Dies folgt direkt aus der Beobachtung, dass

$$\begin{aligned} y(t) &= M(t)c = M(t)(c_1 e_1 + \dots + c_n e_n) = c_1 M(t)e_1 + \dots + c_n M(t)e_n \\ &= c_1 y_1(t) + \dots + c_n y_n(t). \end{aligned} \quad (800)$$

Da $y_1, \dots, y_n \in \mathbb{L}$ und \mathbb{L} ein Vektorraum ist, folgt mit der obigen Gleichung $y \in \mathbb{L}$.

Als nächstes zeigen wir, dass jede beliebige Lösung die Gestalt in (781) hat. Es sei dazu $y \in C^1(I; \mathbb{R}^n)$ eine beliebige Lösung. Definiere nun $c := y(t_0) \in \mathbb{R}^n$. Dann löst y das AWP

$$\begin{cases} y'(t) = A(t)y(t) & (t \in I), \\ y(t_0) = c. \end{cases} \quad (801)$$

Setze nun $z(t) := M(t)c$ $\underset{(800)}{=} c_1y_1(t) + \dots + c_ny_n(t)$, $(t \in I)$. Mit (800) sieht man

$$z(t_0) = c_1y_1(t_0) + \dots + c_ny_n(t_0) = c_1e_1 + \dots + c_ne_n = c. \quad (802)$$

Da z aber nun auch (wieder wegen (800)) $z'(t) = A(t)z(t)$ erfüllt, löst z dasselbe AWP wie y (also (801))! Mit der Eindeutigkeit der maximalen Lösung folgt $z = y$. Die Existenz einer Fundamentalmatrix $t \mapsto M(t)$ ist somit bewiesen. Der Rest der Aussage wurde bereits in Kapitel 1 auf Seite 32 f. (im Teilbeweis zu Satz 37) gezeigt.

Hiermit haben wir unsere Überbleibsel aus Kapitel 1 final geklärt.

3.3 Wohlgestelltheit

Wir möchten dieses Kapitel mit einer Bemerkung zum *Wohlgestelltheitsbegriff* im Sinne von Hadamard abschließen.

Der französische Mathematiker *Jacques Hadamard* (1865-1963) nennt ein mathematisches Problem *wohlgestellt* falls folgende Eigenschaften erfüllt sind

- (1) Es existiert stets eine Lösung.
- (2) Lösungen sind eindeutig.
- (3) Lösungen hängen stetig von den Eingangsdaten ab.

Das Problem, welches wir hier untersuchen, ist das Finden einer maximalen Lösung für ein *Anfangswertproblem*. Der Satz von Picard-Lindelöf klärt hierbei die Punkte (1) und (2). Punkt (3) ist in unserem Fall als die stetige Abhängigkeit der Lösung vom Anfangswert zu verstehen. Diese ist auch zutreffend, allerdings muss man hier zunächst klarstellen wie man *stetige Abhängigkeit* versteht. „Moralisch“ ist damit folgendes gemeint: Verändert man das Anfangsdatum nur wenig, so verändert sich auch die Lösung nur wenig.

Unter einer globalen Lipschitz-Bedingung ist dies eine leichte Konsequenz aus dem Lemma von Gronwall (Lemma 127, p. 133).

Satz 133. STETIGE ABHÄNGIGKEIT. Es erfülle $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine globale Lipschitz-Bedingung mit Lipschitz-Konstante $L > 0$. Ferner seien $y, z \in C^1(I; \mathbb{R}^n)$ die eindeutigen Lösungen zu

$$\begin{cases} y'(t) = f(t, y(t)) & (t \in I), \\ y(t_0) = y_0. \end{cases} \quad (803)$$

und

$$\begin{cases} z'(t) = f(t, z(t)) & (t \in I), \\ z(t_0) = z_0. \end{cases} \quad (804)$$

Dann gilt

$$|y(t) - z(t)| \leq e^{L|t-t_0|} |y_0 - z_0| \quad \forall t \in I. \quad (805)$$

Bevor wir dieses Resultat beweisen, sollten wir klären, inwiefern es als stetige Abhängigkeit aufgefasst werden kann. Es sei dazu $J \subset\subset I$ beliebig. Die Lösungen y und z wie im vorigen Satz liegen dann in $C_b(J; \mathbb{R}^d)$, da sie ja auf \bar{J} beschränkt sind. Aus (805) folgt dann (mit $d_\infty(f, g) := \sup_{t \in J} |f(t) - g(t)|$)

$$d_\infty(y, z) \leq \left(\sup_{t \in J} e^{L(t-t_0)} \right) |y_0 - z_0|, \quad (806)$$

d.h.

$$d_\infty(y, z) \leq K d_{\mathbb{R}^n}(y_0, z_0), \quad (807)$$

wobei $d_{\mathbb{R}^n}(y_0, z_0) := |y_0 - z_0|$, was (HA) eine Metrik auf \mathbb{R}^n definiert. Dies zeigt: Für eine DGL

$$y'(t) = f(t, y(t)), \quad (t \in I), \quad (808)$$

wobei f eine globale Lipschitz-Bedingung erfüllt, und für $J \subset\subset I$ ist die Abbildung

$$A : (\mathbb{R}^n, d_{\mathbb{R}^n}) \rightarrow (C_b(J; \mathbb{R}^n), d_\infty), \quad A(y_0) := \text{Lösung von (808) zum Anfangswert } y_0$$

eine *stetige Abbildung* zwischen den beiden oben genannten metrischen Räumen.

Beweis von Satz 133 (p. 142).

Schritt 1. Abschätzung (805) für $t \geq t_0$. Wir arbeiten mit dem Lemma von Gronwall. Es seien y und z wie in der Aussage. Wir definieren für $t \in I$

$$u(t) := |y(t) - z(t)|. \quad (809)$$

Es sei $T > t_0$ beliebig. Nun gilt für $t \in [t_0, T]$

$$u(t) = |y(t) - z(t)| = \left| \left(y_0 + \int_{t_0}^t y'(s) \, ds \right) - \left(z_0 + \int_{t_0}^t z'(s) \, ds \right) \right| \quad (810)$$

$$= \left| y_0 - z_0 + \int_{t_0}^t (y'(s) - z'(s)) \, ds \right| \quad (811)$$

$$\leq |y_0 - z_0| + \left| \int_{t_0}^t (y'(s) - z'(s)) \, ds \right| \quad (812)$$

$$\leq |y_0 - z_0| + \int_{t_0}^t |y'(s) - z'(s)| \, ds \quad (813)$$

$$\stackrel{(803),(804)}{\leq} |y_0 - z_0| + \int_{t_0}^t |f(s, y(s)) - f(s, z(s))| \, ds. \quad (814)$$

Nun verwenden wir, dass f eine globale Lipschitzbedingung mit Lipschitzkonstante $L > 0$ erfüllt (siehe Definition 3.1) und erhalten

$$u(t) \leq |y_0 - z_0| + \int_{t_0}^t L |y(s) - z(s)| \, ds, \quad (815)$$

und mit (809) umgeschrieben erhalten wir

$$u(t) \leq |y_0 - z_0| + \int_{t_0}^t u(s) \cdot L \, ds. \quad (816)$$

Nun verwenden wir das Lemma von Gronwall (Lemma 127, p. 133) mit $u_0 := |y_0 - z_0|$ und $v(s) := L$ (für alle $s \in [t_0, T]$). Daher haben wir für alle $t \in [t_0, T]$

$$u(t) \leq u_0 \exp\left(\int_{t_0}^t v(s) \, ds\right) = |y_0 - z_0| e^{L(t-t_0)}, \quad (817)$$

mit anderen Worten (da $T > t_0$ beliebig war)

$$|y(t) - z(t)| \leq |y_0 - z_0| e^{L(t-t_0)} \quad \forall t \geq t_0, \quad (818)$$

also auch

$$|y(t) - z(t)| \leq |y_0 - z_0| e^{L|t-t_0|} \quad \forall t \geq t_0, \quad (819)$$

Schritt 2: Abschätzung für $t < t_0$. Dazu lassen wir wiederum die Zeit rückwärts laufen: Definiere $\tilde{y}(t) := y(-t)$ und $\tilde{z}(t) := z(-t)$. Es lösen \tilde{y} und \tilde{z} die DGLn

$$\tilde{y}'(t) = -y'(-t) = -f(-t, y(-t)) = -f(t, \tilde{y}(t)) \quad (820)$$

und analog

$$\tilde{z}'(t) = -f(-t, \tilde{z}(t)). \quad (821)$$

Man hat also, falls $I = (a, b)$

$$\begin{cases} \tilde{y}'(t) = -f(t, \tilde{y}(t)) & (t \in (-b, -a)) \\ \tilde{y}(-t_0) = y_0 \end{cases} \quad (822)$$

und

$$\begin{cases} \tilde{z}'(t) = -f(t, \tilde{z}(t)) & (t \in (-b, -a)) \\ \tilde{z}(-t_0) = z_0. \end{cases} \quad (823)$$

Es ist nun leicht zu sehen, dass die Abbildung

$$\tilde{f} : (-b, -a) \times \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad \tilde{f}(t, z) := -f(-t, z) \quad (824)$$

auch eine globale Lipschitz-Bedingung mit derselben Lipschitzkonstante $L > 0$ erfüllt (HA). Daher gilt nach Schritt 1 (wobei wir hier der Zeit den Namen τ geben)

$$|\tilde{y}(\tau) - \tilde{z}(\tau)| \leq e^{L(\tau-(-t_0))} |y_0 - z_0| \quad \forall \tau \geq -t_0. \quad (825)$$

Substituieren wir $\tau := -t$ und verwenden $\tilde{y}(\tau) = y(-\tau) = y(t)$, sowie $\tilde{z}(\tau) = z(t)$, so erhalten wir

$$|y(t) - z(t)| \leq e^{L(t_0-t)} |y_0 - z_0| \quad \forall t \in \mathbb{R} : -t \geq -t_0, \quad (826)$$

das heißt

$$|y(t) - z(t)| \leq e^{L|t-t_0|} |y_0 - z_0| \quad \forall t \leq t_0. \quad (827)$$

Dies und (819) zeigen die Behauptung.

Wie so oft ist hier wieder die Frage, ob die globale Lipschitz-Bedingung zu einer lokalen Lipschitz-Bedingung abgeschwächt werden kann. Eine Verallgemeinerung der Aussage von Satz 133 ist jedoch nicht offensichtlich: Denn die (maximalen) Lösungen zu zwei verschiedenen Anfangswerten sind ja nichteinmal auf demselben Intervall definiert. Aus diesem Grund sollte man zufrieden sein, wenn die Lösungen *für kurze Zeit* stetig voneinander abhängen

Auf kleinen Intervallen kann die stetige Abhängigkeit auch nur unter einer lokalen Lipschitz-Bedingung erhalten werden.

Übung 134. STETIGE ABHÄNGIGKEIT UNTER LOKALER LIPSCHITZ-BEDINGUNG. Es erfülle $f \in C^0(I \times \mathbb{R}^n; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung mit Lipschitz-Konstante $L > 0$. Ferner seien y, z die eindeutigen maximalen Lösungen zu

$$\begin{cases} y'(t) = f(t, y(t)), \\ y(t_0) = y_0. \end{cases} \quad (828)$$

und

$$\begin{cases} z'(t) = f(t, z(t)), \\ z(t_0) = z_0. \end{cases} \quad (829)$$

Zeigen Sie: Dann gibt es ein $\delta > 0$, ein $L > 0$ und ein $r > 0$ mit

$$|y(t) - z(t)| \leq e^{L|t-t_0|} |y_0 - z_0| \quad \forall t \in (t_0 - \delta, t_0 + \delta) \quad \forall z_0 \in B_r(y_0) \quad (830)$$

SINN DER ÜBUNG. Dies ist nochmal eine gute Übung für *global-to-local Techniken*. Kann man die Abschätzung im vorigen Satz für kleine Mengen *lokalisieren*?

4 Ausblick: Autonome Differentialgleichungen

Eine wichtige Klasse von Differentialgleichungen ist gegeben durch DGL's der Form

$$y'(t) = f(y(t)) \quad (t \in I) \quad (831)$$

wobei $f \in C^0(G; \mathbb{R}^n)$ für ein $G \subset \mathbb{R}^n$. Die Funktion f ist hier also nicht explizit zeitabhängig. Wir werden sehen, dass wir uns Lösungen dieser DGLs vorstellen können wie Partikel in einer *ortsabhängigen Strömung*. Diese Partikel erzeugen dann *Bahnen*, d.h. Stromlinien. Dieses Profil von Stromlinien ist für jede Differentialgleichung ein Anderes.

Wir werden sehen: Die Stromlinien sagen viel über das Verhalten von Lösungen aus. In der Forschung stellt man sich für ein gegebenes $f \in C^0(G; \mathbb{R}^n)$ zum Beispiel folgende

Fragen: Sind alle Stromlinien beschränkt? Nähern sich alle Stromlinien einem Punkt an, wie bei einem Strudel? Bleiben bestimmte Stromlinien in einer vorgegebenen Menge? All diese Eigenschaften können aus der Funktion f herausgelesen werden. In diesem Kontext wird f auch häufig *Strömungsfeld* oder *Vektorfeld* genannt.

Definition 135. AUTONOME DIFFERENTIALGLEICHUNG. Es sei $f \in C^0(G; \mathbb{R}^n)$. Eine *autonome Differentialgleichung* ist eine DGL der Gestalt

$$y'(t) = f(y(t)). \quad (832)$$

Eine Lösung ist eine Funktion $y \in C^1(I; \mathbb{R}^n)$ (für irgendein Intervall $I \subset \mathbb{R}$), welche (832) für alle $t \in I$ erfüllt.

Ein großer Unterschied bei autonomen Differentialgleichungen: Es ist wegen der Zeitunabhängigkeit von f nicht notwendig, ein bestimmtes Zeitintervall vorzuschreiben, auf dem man die DGL lösen will, sondern man lässt einfach beliebige Intervalle $I \subset \mathbb{R}$ zu. Die grundlegende Beobachtung auf der das Studium autonomer Differentialgleichungen beruht ist die Folgende

Beobachtung 136. SHIFT-INVARIANZ. Es sei $y \in C^1((a, b); G)$ eine Lösung von (832). Definiere nun für ein fest gewähltes $\tau \in \mathbb{R}$ die *geshiftete* oder *zeitverschobene* Funktion $z(t) := y(t + \tau)$, ($t \in (a - \tau, b - \tau)$). Dann gilt für alle $t \in (a - \tau, b - \tau)$

$$z'(t) = y'(t + \tau) \underset{(832)}{=} f(y(t + \tau)) = f(z(t)). \quad (833)$$

Insbesondere ist die geshiftete Funktion z wieder eine Lösung der Differentialgleichung (832). Dies ist eine Besonderheit autonomer DGLs und würde nicht funktionieren, wenn f explizit zeitabhängig wäre. Eine wichtige Konsequenz ist, dass man sich bei Anfangswertproblemen immer auf den Anfangszeitpunkt $t_0 = 0$ setzen kann, ohne die Eigenschaften der Lösung zu verändern. In der Tat: Ist für $y_0 \in G$ und $t_0 \in \mathbb{R}$ y eine Lösung des Anfangswertproblems

$$\begin{cases} y'(t) = f(y(t)) & (t \in (a, b)) \\ y(t_0) = y_0 \end{cases} \quad (834)$$

so definiert $z : (a - t_0, b - t_0) \rightarrow \mathbb{R}^n$ gegeben durch $z(t) = y(t + t_0)$ eine Lösung von

$$\begin{cases} z'(t) = f(z(t)) & (t \in (a - t_0, b - t_0)) \\ z(0) = y_0. \end{cases} \quad (835)$$

Dies zeigt, dass es ausreicht, den Anfangszeitpunkt $t_0 = 0$ zu betrachten, um das Verhalten aller Lösungen zu verstehen.

Als nächstes wollen wir die oben angesprochenen *Bahnen* definieren. Diese sind Wertemengen maximaler Lösungen. Damit maximale Lösungen existieren, müssen wir aber ab jetzt annehmen, dass f eine lokale Lipschitz-Bedingung erfüllt. Damit meinen wir (präziser), dass

die Abbildung $g : \mathbb{R} \times G \rightarrow \mathbb{R}^n$ gegeben durch $g(t, z) := f(z)$ eine lokale Lipschitz-Bedingung erfüllt, denn die DGL hat dann die gewohnte Form $y'(t) = g(t, y(t))$. Da die Zeitkomponente hier überhaupt keine Rolle spielt, bedeutet das konkret, dass es für alle $z_0 \in G$ ein $r > 0$ und ein $L > 0$ gibt mit

$$|f(z_1) - f(z_2)| \leq L|z_1 - z_2| \quad \forall z_1, z_2 \in B_r(z_0). \quad (836)$$

Definition 137. BAHNEN UND FLÜSSE. Es erfülle $f \in C^0(G; \mathbb{R}^n)$ eine lokale Lipschitz-Bedingung und sei $y_0 \in G$. Es sei ferner $y_{max} \in C^1(I_{max}; G)$ die eindeutige maximale Lösung zum Anfangswertproblem

$$\begin{cases} y'(t) = f(y(t)) \\ y(0) = y_0. \end{cases} \quad (837)$$

Dann definieren wir den *Fluss* von y_0 durch $\phi_{y_0} : I_{max} \rightarrow \mathbb{R}^n$ durch

$$\phi_{y_0}(t) := y_{max}(t) \quad (t \in I_{max}). \quad (838)$$

Wir nennen den Fluss ϕ_{y_0} *global*, falls $I_{max} = \mathbb{R}$. Ferner definieren wir die *Bahn* (oder das *Orbit*) von y_0 durch

$$\mathcal{O}(y_0) := \{\phi_{y_0}(t) : t \in I_{max}\} = \phi_{y_0}(I_{max}). \quad (839)$$

Zudem definieren wir die *positive Bahn* durch

$$\mathcal{O}^+(y_0) := \{\phi_{y_0}(t) : t \in I_{max} \cap [0, \infty)\} = \phi_{y_0}(I_{max} \cap [0, \infty)) \quad (840)$$

Die *Flussabbildung* $\Phi : A \rightarrow \mathbb{R}^n$ ist gegeben durch $\Phi(t, y_0) := \phi_{y_0}(t)$, wobei hier

$$A := \{(t, y_0) : t \in I_{max}(0, y_0)\}, \quad (841)$$

wobei $I_{max}(0, y_0)$ das (von y_0 abhängige maximale Existenzintervall von ϕ_{y_0} bezeichnet.

Wir stellen uns das also so vor: Ein Partikel im Strömungsfeld f , welches wir bei $y_0 \in G$ starten, bewegt sich in der Bahn $\mathcal{O}(y_0)$. Eigenschaften solcher Bahnen und deren strukturelle Bedeutung sollen das Hauptaugenmerk dieses kurzen Ausblicks sein. Im Folgenden sprechen wir kurz über *Äquilibria*, d.h. über Bahnen, die $\mathcal{O}(y_0) = \{y_0\}$ erfüllen.

Beispiel 138. ÄQUILIBRIA. Es sei $\bar{y} \in G$ so, dass $f(\bar{y}) = 0$. Dann ist (HA) die konstante Funktion $y_{max} : \mathbb{R} \rightarrow \mathbb{R}^n$ definiert durch $y_{max}(t) := \bar{y}$ für alle $t \in \mathbb{R}$ die eindeutige maximale Lösung von

$$\begin{cases} y'(t) = f(y(t)) \\ y(0) = \bar{y}. \end{cases} \quad (842)$$

Die Bahn von \bar{y} ist dann also gegeben durch $\mathcal{O}(\bar{y}) = \{\bar{y}\}$ (HA). Auch gilt $\mathcal{O}^+(\bar{y}) = \{\bar{y}\}$. Ein solches $\bar{y} \in G$ mit $f(\bar{y}) = 0$ nennen wir daher *Äquilibrium* oder *Gleichgewichtspunkt*

der Differentialgleichung.

Ausblick 139. STABILITÄT. Es sei $\bar{y} \in G$ ein Äquilibrium, d.h. $f(\bar{y}) = 0$. In der Forschung interessiert man sich dafür, was Flüsse in einer Umgebung von \bar{y} machen, d.h. zum Beispiel: wie sieht $\mathcal{O}^+(w)$ für w in einer kleinen Umgebung von \bar{y} aus? Zum Beispiel heißt ein Äquilibrium $\bar{y} \in G$ *stabil* falls

$$\forall \epsilon > 0 \exists \delta > 0 : (|w - \bar{y}| < \delta \Rightarrow \mathcal{O}^+(w) \subset B_\epsilon(\bar{y})). \quad (843)$$

Ein Äquilibrium $\bar{y} \in G$ heißt *attraktiv*, falls

$$\exists \epsilon > 0 : \bar{y} \in \overline{\mathcal{O}^+(w)} \quad \forall w \in B_\epsilon(\bar{y}), \quad (844)$$

wobei hier $\overline{\mathcal{O}^+(w)}$ den Abschluss von der positiven Bahn $\mathcal{O}^+(w)$ in \mathbb{R}^n meint. Ferner heißt ein Äquilibrium *asymptotisch stabil*, falls es stabil und attraktiv ist. Man kann Kriterien an f formulieren, diese (asymptotische) Stabilität bei y gewährleisten.

***-Übung 140. STABILITÄT LINEARER SYSTEME.** Gegeben sei für eine beliebige Matrix $A \in \mathbb{R}^{n \times n}$ die autonome DGL

$$y'(t) = Ay(t). \quad (845)$$

- (a) Zeigen Sie, dass $\bar{y} = 0$ ein Äquilibrium ist.
- (b) Zeigen Sie: Ist A diagonalisierbar $\text{Re}(\lambda) \leq 0$ für alle Eigenwerte λ von A , so ist $\bar{y} = 0$ stabil.
- (c) Kann in Teilaufgabe (b) auf die Diagonalisierbarkeit von A verzichtet werden?
- (d) Zeigen Sie: Ist $\text{Re}(\lambda) < 0$ für alle Eigenwerte λ von A , so ist $\bar{y} = 0$ asymptotisch stabil.

SINN DER ÜBUNG. Die Stabilitätsuntersuchung, die wir hier für lineare Systeme gemacht haben, lässt sich tatsächlich auf nicht-lineare Systeme verallgemeinern! Man kann beweisen, dass die DGL

$$y'(t) = f(y(t)) \quad (846)$$

in einer kleinen Umgebung eines Äquilibriums \bar{y} ein ähnliches Verhalten zeigt wie die *um \bar{y} linearisierte* DGL

$$z'(t) = Df(\bar{y})z(t) \quad (847)$$

zeigt. Solche Linearisierungsverfahren sind nicht nur für gewöhnliche, sondern auch für partielle Differentialgleichungen relevant.

Wir wollen als nächstes noch die Anschauung von Bahnen als Stromlinien etwas genauer diskutieren. Dass wir uns Bahnen als Stromlinien vorstellen können, liegt daran, dass sich jeder Fluss komplett in einer Bahn bewegt.

Beispiel 141. FLUSSINVARIANZ VON BAHNEN. Es sei $t_0 \in \mathbb{R}$ und $y_0 \in G$ so, dass $(t_0, y_0) \in A$.

BEHAUPTUNG: Dann gilt $\mathcal{O}(y_0) = \mathcal{O}(\Phi(t_0, y_0))$.

BEWEIS: Es ist $\phi_{y_0} : I_{max} \rightarrow \mathbb{R}^n$, $I_{max} = (t_-, t_+)$, die eindeutige maximale Lösung von

$$\begin{cases} y'(t) = f(y(t)) \\ y(0) = y_0. \end{cases} \quad (848)$$

Daher ist $\mathcal{O}(y_0) = \phi_{y_0}(I_{max})$. Ferner gilt auch $\Phi(t_0, y_0) = \phi_{y_0}(t_0)$. Die Bahn $\mathcal{O}(\Phi(t_0, y_0))$ ist nun gegeben durch $\phi_{\Phi(t_0, y_0)}(J_{max})$, wobei $\phi_{\Phi(t_0, y_0)} : J_{max} \rightarrow \mathbb{R}^n$ die eindeutige maximale Lösung von

$$\begin{cases} z'(t) = f(z(t)) \\ z(0) = \Phi(t_0, y_0) \end{cases} \quad \text{das heißt} \quad \begin{cases} z'(t) = f(z(t)) \\ z(0) = \phi_{y_0}(t_0) \end{cases} \quad (849)$$

ist. Man überzeugt sich nun leicht, dass die maximale Lösung von (849) durch $z_{max}(t) = \phi_{y_0}(t + t_0)$, für $t \in J_{max} = (t_- - t_0, t_+ - t_0)$ gegeben ist. In der Tat: Dass $\phi_{y_0}(\cdot + t_0)$ die DGL löst, folgt direkt aus (848) und der Shiftinvarianz (siehe Beobachtung 136, p. 146). Die Anfangswertbedingung überprüft man leicht durch Einsetzen (HA). Die Maximalität sieht man so: Würde sich die geshiftete Lösung z_{max} weiter fortsetzen lassen, so ließe sich auch die Lösung ϕ_{y_0} durch Zurückshiften dieser Fortsetzung weiter fortsetzen lassen. Das kann aber nicht sein, da ϕ_{y_0} nicht fortsetzbar ist. Die Details überlassen wir hier den Lesenden als Übungsaufgabe. In jedem Fall folgern wir aus der Eindeutigkeit der maximalen Lösung, dass

$$\phi_{\Phi(t_0, y_0)} = z_{max} = \phi_{y_0}(\cdot + t_0) \quad \text{auf } J_{max} = (t_- - t_0, t_+ - t_0). \quad (850)$$

Daher gilt

$$\mathcal{O}(\Phi(t_0, y_0)) = \phi_{\Phi(t_0, y_0)}((t_- - t_0, t_+ - t_0)) \underset{(850)}{=} \phi_{y_0}((t_-, t_+)) = \mathcal{O}(y_0). \quad (851)$$

Beobachtung 142. FLUSSEIGENSCHAFT. Im vorigen Beispiel, genauer in (850), haben wir gesehen, dass $\phi_{\Phi(t_0, y_0)}(t) = \phi_{y_0}(t + t_0)$, für alle t wo beide Ausdrücke definiert sind. Mit der Flussabbildung geschrieben erhält man

$$\Phi(t, \Phi(t_0, y_0)) = \phi_{\Phi(t_0, y_0)}(t) = \phi_{y_0}(t + t_0) = \Phi(t + t_0, y_0), \quad (852)$$

also insgesamt

$$\Phi(t, \Phi(t_0, y_0)) = \Phi(t + t_0, y_0). \quad (853)$$

Diese Formel nennen wir *Flusseigenschaft*. Ihre Interpretation ist wie folgt: Lassen wir ein Partikel von einem Startort y_0 aus erst t_0 Sekunden fließen und danach nochmal t Sekunden fließen (linker Ausdruck), so ergibt sich dieselbe Position als wenn wir dasselbe Partikel nochmal bei y_0 starten würden und direkt $t + t_0$ Sekunden fließen lassen würden (rechter Ausdruck).

Wir werden nun sehen, dass G in eine Familie disjunkter(!) *Bahnen* zerfällt. Plottet man

alle Bahnen in G (bzw. möglichst viele Bahnen), so sprechen wir von einem *Phasenporträt*. Aus diesem Phasenporträt kann man meist viel über das Verhalten von Lösungen lernen. Zum Beispiel kann man damit manchmal auch Globalität von Lösungen nachweisen, wie wir am Ende noch sehen werden.

Proposition 143. BAHNEN KREUZEN SICH NICHT. Es seien $y_0, y_1 \in G$. Dann gilt entweder $\mathcal{O}(y_0) = \mathcal{O}(y_1)$ oder $\mathcal{O}(y_0) \cap \mathcal{O}(y_1) = \emptyset$.

Beweis. Nehmen wir mal an, es gilt $\mathcal{O}(y_0) \cap \mathcal{O}(y_1) \neq \emptyset$, sei etwa $y_* \in \mathcal{O}(y_0) \cap \mathcal{O}(y_1)$. Dann gibt es $t_0 \in \mathbb{R}$ mit $y_* = \phi_{y_0}(t_0) = \Phi(t_0, y_0)$ und $t_1 \in \mathbb{R}$ mit $y_* = \phi_{y_1}(t_1) = \Phi(t_1, y_1)$. Dann gilt nach Beispiel 141 (p. 149)

$$\mathcal{O}(y_0) = \mathcal{O}(\Phi(t_0, y_0)) = \mathcal{O}(y_*) = \mathcal{O}(\Phi(t_1, y_1)) = \mathcal{O}(y_1). \quad (854)$$

Die Behauptung ist gezeigt.

Dass Bahnen sich nicht kreuzen (d.h. entweder gleich oder disjunkt sind), kann wie angekündigt auch zum Nachweis von Globalität verwendet werden. Es ist nämlich jetzt möglich, Lösungen mit geschickt gewählten Bahnen in beschränkten Mengen *einzusperren*. Dieses Einsperren kann dann genutzt werden, um Escape-Situationen und Blow-Up-Situationen auszuschließen.

Beispiel 144. Wir betrachten das autonome DGL-System

$$\begin{cases} y'_1(t) = -y_2(t) \\ y'_2(t) = y_1(t) + y_1(t)^2 + y_2(t)^2 - 1 \end{cases}. \quad (855)$$

Man rechnet leicht nach, dass $t \mapsto (\cos(t), \sin(t))$ eine globale Lösung auf ganz \mathbb{R} definiert. Somit ist $\phi_{(1,0)^T}(t) = (\cos(t), \sin(t))$, da die Lösung $t \mapsto (\cos(t), \sin(t))$ den Anfangswert $(1, 0)^T$ annimmt. Somit gilt

$$\mathcal{O}((1, 0)^T) = \{(\cos(t), \sin(t)) : t \in \mathbb{R}\} = \partial B_1(0). \quad (856)$$

BEHAUPTUNG: Für alle $y_0 = (y_{0,1}, y_{0,2})^T \in B_1(0)$ hat das Anfangswertproblem

$$\begin{cases} y'_1(t) = -y_2(t) \\ y'_2(t) = y_1(t) + (y_1(t)^2 + y_2(t)^2 - 1) \\ (y_1(0), y_2(0)) = (y_{0,1}, y_{0,2}) \end{cases} \quad (857)$$

eine globale Lösung $y \in C^1(\mathbb{R}; \mathbb{R}^2)$.

BEWEIS DER BEHAUPTUNG: Dazu überprüfen wir zunächst, dass $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch

$$f(z_1, z_2) = \begin{pmatrix} -z_2 \\ z_1 + z_1^2 + z_2^2 - 1 \end{pmatrix} \quad (858)$$

eine lokale Lipschitzbedingung erfüllt. Dies folgt direkt aus Proposition 97 (p. 94). Somit finden wir eine maximale Lösung $y_{max} \in C^1(I_{max}; \mathbb{R}^2)$. Es bleibt zu zeigen, dass $I_{max} = \mathbb{R}$. Hierzu verwenden wir das Maximalitätskriterium, also Satz 96 (p. 93). Da f auf ganz \mathbb{R}^2 definiert ist, kann eine Escape-Situation nicht vorkommen. Es muss nur die Blow-Up-Situation ausgeschlossen werden. Hierfür zeigen wir $y_{max}(t) \in B_1(0)$ für alle $t \in I_{max}$. Dies hat folgenden Grund: Da $\partial B_1(0) = \mathcal{O}((1, 0)^T)$ eine Bahn ist, darf $\partial B_1(0)$ nicht gekreuzt werden. Somit ist die Lösung im Inneren von $\partial B_1(0)$ eingesperrt! Mit anderen Worten: Sie muss die ganze Zeit in $B_1(0)$ bleiben. Die Blow-Up-Situation ist somit ausgeschlossen und die Globalität ist gezeigt.

Wem die obige Erklärung zum Sachverhalt " $y_{max}(t) \in B_1(0)$ für alle $t \in I_{max}$ " nicht ausreichend rigoros war: hier folgt noch einmal eine mathematisch wasserdichte Herleitung. Angenommen, es gäbe ein $t_0 \in I_{max}$ mit $y_{max}(t_0) \notin B_1(0)$. Dann wäre (wegen $y_0 = (y_{0,1}, y_{0,2})^T \in B_1(0)$)

$$|y_{max}(0)| = |(y_{0,1}, y_{0,2})^T| < 1 \quad \text{und} \quad |y_{max}(t_0)| \geq 1. \quad (859)$$

Nach dem Zwischenwertsatz muss dann ein $t^* \in (0, t_0]$ existieren mit $|y_{max}(t^*)| = 1$, d.h. $y_{max}(t^*) \in \partial B_1(0)$. Insbesondere gilt

$$y_{max}(t^*) = \phi_{y_0}(t^*) \in \mathcal{O}(y_0) \quad (860)$$

und gleichzeitig

$$y_{max}(t^*) \in \partial B_1(0) = \mathcal{O}((1, 0)^T). \quad (861)$$

Daher folgert man $\mathcal{O}(y_0) \cap \mathcal{O}((1, 0)^T) \neq \emptyset$. Mit der Proposition 143 (p. 150) folgern wir $\mathcal{O}(y_0) = \mathcal{O}((1, 0)^T) = \partial B_1(0)$. Jedoch gilt dann

$$y_0 \in \mathcal{O}(y_0) = \partial B_1(0), \quad (862)$$

ein Widerspruch zur Tatsache, dass $y_0 \in B_1(0)$. Man beachte, dass es uns nur gelungen ist, die Lösungen einzusperren, weil wir eine Bahn gefunden haben, die die Lösungen einsperrt.

Ausblick 145. INVARIANTE MENGEN. Im vorigen Beispiel haben wir davon profitiert dass wir Lösungen in einer gewissen Menge einsperren können. Wir bezeichnen solche Mengen als *invariante Mengen*. Konkreter heißt eine Menge $C \subset G$ *invariant* falls

$$\mathcal{O}(w) \subset C \quad \forall w \in C. \quad (863)$$

Wiederum ist es möglich, Kriterien an eine Menge C zu formulieren, die hinreichend für die Invarianz von C sind. Plottet man ein hinreichend genaues Phasenportät, so werden invariante Mengen zumeist sichtbar: Es sind diejenigen Mengen, aus denen Bahnen nicht hinausführen. Ein weiterer wichtiger Begriff ist der Begriff der *positiven Invarianz*. Man nennt $C \subset G$ *positiv invariant* falls $\mathcal{O}^+(w) \subset C$ für alle $w \in C$. Tatsächlich sind nicht alle positiv invarianten Mengen gleich invariant – In diesem Zusammenhang ist ein *Umkehren*

der Zeitachse also nicht ohne Veränderung des Verhaltens möglich.

***-Übung 146. INVARIANTE MENGEN EINER SIR-PANDEMIE.** Es sei das SIR-Pandemiemodell gegeben. Zeigen Sie, dass die Menge $C = (0, 1)^3$ positiv invariant ist, jedoch nicht invariant.

SINN DER ÜBUNG: Hier sieht man, dass die Theorie invarianter Mengen auch für die Anwendung sehr nützlich sein kann. Für die SIR-Pandemie kann man die positive Invarianz recht gut “zu Fuß” zeigen (siehe Satz 76, p. 74). Betrachtet man aber das SIR-Modell mit *Reinfektionen* (siehe 87, p. 86), so muss ein tiefliegenderes Invarianzkriterium benutzt werden.

In dieser Ausblick-Sektion haben wir die Fragen der *Stabilität von Äquilibria* und der *Invarianz von Mengen* aufgeworfen. Das Studium der *dynamischen Systeme* beschäftigt sich mit diesen Fragen und noch vielen mehr, z.B. auch was das *asymptotische Verhalten* von Flüssen anbelangt. Wir hoffen dieser Ausblick hat Ihnen gefallen und Lust auf mehr gemacht. Außerdem bedanken wir uns noch einmal ganz herzlich für das Verfolgen dieser Vorlesung!