

Infinitesimalrechnung II

Vorlesung im Sommersemester 1993

von Prof. Dr. Rolf Leis

Inhaltsverzeichnis

| | | |
|----------|---|------------|
| 8 | Gewöhnliche Differentialgleichungen | 3 |
| 8.1 | Einführung | 3 |
| 8.2 | Klassische Existenz- und Eindeutigkeitsätze | 10 |
| 8.3 | Spezielle Lösungsmethoden | 18 |
| 8.3.1 | Lineare Systeme | 18 |
| 8.3.2 | Spezielle Gleichungen | 24 |
| 8.3.3 | Wachstumsmodelle | 27 |
| 8.4 | Qualitative Aspekte | 28 |
| 8.4.1 | Maximale Lösungen | 28 |
| 8.4.2 | Stabilität | 29 |
| 8.4.3 | Normalformen | 33 |
| 8.4.4 | Periodische Lösungen im \mathbb{R}^2 | 34 |
| 8.4.5 | Verzweigungen | 35 |
| 8.5 | Rand- und Eigenwertaufgaben | 36 |
| 8.5.1 | Sturm-Liouvillesche Randwertaufgaben | 37 |
| 8.5.2 | Die Greensche Funktion | 39 |
| 8.5.3 | Randwertaufgaben | 40 |
| 8.5.4 | Eigenwertaufgaben | 41 |
| 8.5.5 | Entwicklungssätze | 45 |
| 8.5.6 | Anwendungen | 49 |
| 8.6 | Abbildungen | 52 |
| 8.7 | Ergänzungen | 58 |
| 8.7.1 | Bahn eines Satelliten im Feld Erde-Mond | 58 |
| 8.7.2 | Rechnung zur Seite 10 oben | 60 |
| 8.7.3 | Zur Nagumobedingung | 61 |
| 8.7.4 | Der RLC-Kreis | 61 |
| 9 | Funktionen von mehreren Veränderlichen | 63 |
| 9.1 | Stetige Funktionen | 63 |
| 9.2 | Lineare Abbildungen | 66 |
| 9.3 | Gleichmäßige Konvergenz | 71 |
| 9.4 | Differenzierbare Abbildungen | 74 |
| 9.5 | Der Mittelwertsatz | 79 |
| 9.6 | Höhere Ableitungen | 81 |
| 9.7 | Lokale Umkehrbarkeit | 88 |
| 9.8 | Implizite Funktionen | 91 |
| 9.9 | Kurven und Flächen | 94 |
| 9.9.1 | Ebene Kurven | 97 |
| 9.9.2 | Kurven im Raum | 98 |
| 9.9.3 | m -dimensionale Flächen im \mathbb{R}^n | 99 |
| 9.10 | Extrema unter Nebenbedingungen | 101 |
| 9.10.1 | Quadratische Formen | 103 |
| | Bezeichnungen | 105 |
| | Index | 106 |

8 Gewöhnliche Differentialgleichungen

8.1 Einführung

Eine gewöhnliche Differentialgleichung ist — grob gesprochen — eine Gleichung, in die ein Parameter $t \in \mathbb{R}$, eine Abbildung $x : \mathbb{R} \rightarrow \mathbb{R}^n$ und ihre Ableitungen nach t eingehen; also

$$f(t, x(t), x'(t), \dots, x^{(k)}(t)) = 0.$$

Die Aufgabe besteht darin, solche Gleichungen nach x aufzulösen. Würde x von mehreren Parametern abhängen, dann träten partielle Ableitungen auf, und man würde von einer partiellen Differentialgleichung sprechen.

Gewöhnliche Differentialgleichungen findet man häufig in den Anwendungen, insbesondere in der Physik. Viele Vorgänge in der klassischen Mechanik werden durch gewöhnliche Differentialgleichungen beschrieben. Solche Vorgänge sind determiniert, endlich-dimensional und differenzierbar. Damit ist folgendes gemeint:

1. Kennt man den Zustand zur Zeit t_0 (in der Gegenwart), dann kennt man ihn für alle Zeiten (in der Zukunft und in der Vergangenheit).
2. Die Menge der Zustände zu einem festen Zeitpunkt — insbesondere die Menge der Anfangszustände — ist endlich-dimensional.
3. Der Prozeß soll durch differenzierbare Funktionen beschrieben werden.

Natürlich lassen sich nicht alle Probleme aus den Anwendungen auf diese Weise beschreiben. Insbesondere liegt die dritte Forderung in der Natur der Sache, weil wir ja Differentialgleichungen behandeln wollen. Es gibt Verallgemeinerungen des Ableitungsbegriffs, um auch nicht differenzierbare Vorgänge mit einzubeziehen.

Lassen Sie mich das Gesagte an einem Beispiel erläutern. Denken Sie an den freien Fall in Erdnähe. Nach NEWTON ist Kraft gleich Masse mal Beschleunigung, also

$$m h'' = -m g.$$

Dabei sind h die Höhe und $g = 9,81 \text{ m/sec}^2$ die Erdbeschleunigung. Wählen wir Zeit durch \sqrt{g} als neue Variable t , dann lautet unsere Differentialgleichung

$$y'' = -1.$$

Hinzu treten Anfangsbedingungen. Aus der Physik weiß man, daß der Anfangszustand durch zwei Konstanten, nämlich durch die Anfangslage und die Anfangsgeschwindigkeit

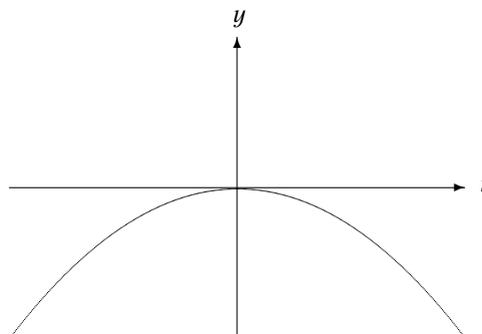
$$y(0) = a \text{ und } y'(0) = b$$

gegeben ist.

Diese Aufgabe ist so einfach, daß man sie durch Integrieren direkt lösen kann, nämlich

$$\begin{aligned} y'(t) &= -t + b \\ y(t) &= -\frac{t^2}{2} + b t + a. \end{aligned}$$

Mit $a = b = 0$ erhält man also



An diesem Beispiel sehen Sie schon, daß man zum Lösen einer Differentialgleichung integrieren muß. Und Sie wissen auch, daß sich nicht alle elementaren Funktionen geschlossen integrieren lassen. Und selbst wenn das möglich ist, erkennt man es nicht immer sofort. So gibt es eine ganze Reihe von Integrationstechniken und -tafeln. Dieses Problem potenziert sich beim Lösen von Differentialgleichungen. Es gibt viele Nachschlagewerke, und die technische Seite beim Lösen der Gleichungen hat jahrelang im Vordergrund gestanden. Dadurch hat die

Entwicklung der Theorie der gewöhnlichen Differentialgleichungen gelitten, und das ganze Gebiet war etwas in Mißkredit geraten.

Inzwischen hat sich die Situation aber geändert, und zwar einmal natürlich durch die Entwicklung in der theoretischen Mathematik selbst, zum anderen aber gerade auch durch das Hilfsmittel des Computers. Heute wird ein Anwender nicht mehr unbedingt in Nachschlagewerken nach exakten Lösungen suchen, an denen er dann oft doch nicht viel erkennen kann. Vielmehr wird er sich lieber direkt vom Computer ein Bild der Lösung geben lassen, an dem er dann qualitativ viel mehr sieht. Und auch der Theoretiker ist oft von solchen Bildern überrascht und kann eine Menge daraus lernen. So ist viel Ballast abgeworfen worden, und auch die Fragestellungen selbst haben sich verändert. Man ist gar nicht mehr so sehr an exakten Lösungen interessiert. Vielmehr stehen heute qualitative Fragen im Vordergrund, wie die Frage nach dem asymptotischen Verhalten der Lösungen, nach ihrer Stabilität, nach periodischen Lösungen oder Verzweigungen.

Leider habe ich in dieser Vorlesung über Infinitesimalrechnung nicht viel Zeit für die gewöhnlichen Differentialgleichungen. Ich kann Sie daher, gerade was moderne Fragestellungen angeht, eigentlich nur neugierig machen und muß auf Spezialvorlesungen verweisen. So werde ich zunächst klassische Existenz- und Eindeigkeitssätze zeigen und danach kurz spezielle Lösungsmethoden angeben und qualitative Aussagen streifen. In §8.5 möchte ich aber doch etwas ausführlicher Rand- und Eigenwertaufgaben behandeln.

Ich möchte Ihnen einige Literaturhinweise zu diesem Teil der Vorlesung geben. Auch jetzt nenne ich nur wenige Bücher:

Amann: Gewöhnliche Differentialgleichungen. De Gruyter, Berlin 1983.

Arnold: Gewöhnliche Differentialgleichungen. Springer-Verlag, Berlin 1980.

Heuser: Gewöhnliche Differentialgleichungen. B.G. Teubner, Stuttgart 1989.

Koçak: Differential and difference equations through computer experiments. Springer-Verlag, Berlin 1989.

Verhulst: Nonlinear differential equations and dynamical systems. Springer-Verlag, Berlin 1990.

Wer einen Computer benutzen kann, der sollte sich unbedingt das Buch von Koçak ansehen und die dort dargestellten Methoden auch zur Lösung von Übungsaufgaben benutzen. Dem Buch ist eine Diskette mit dem Programm *Phaser* beigelegt. Es gibt vergleichbare Werke anderer Autoren.

Lassen Sie mich jetzt in der Einführung noch etwas fortfahren und das Beispiel des freien Falls genauer betrachten. Ich habe bereits von der Menge der *Anfangszustände* gesprochen. So hatten wir beim freien Fall den Anfangszustand $y(0) = a$ und $y'(0) = b$ mit $a, b \in \mathbb{R}$. Der Zustandsraum, oder der *Phasenraum* \mathcal{X} , wie wir sagen wollen, ist in diesem Falle der \mathbb{R}^2 , also $\mathcal{X} = \mathbb{R}^2$. Im allgemeinen soll \mathcal{X} endlich-dimensional sein, das war unsere zweite Forderung. Es sei daher \mathcal{X} im folgenden ein \mathbb{R}^n oder allgemeiner eine n -dimensionale Mannigfaltigkeit, zum Beispiel eine Zylinderfläche.

Die Bewegung soll nun durch eine Abbildung

$$\Phi : \mathbb{R} \times \mathcal{X} \longrightarrow \mathcal{X}$$

beschrieben werden. $\mathbb{R} \times \mathcal{X}$ nennt man auch den *erweiterten Phasenraum*. Für jedes feste t und x^0 soll also $\Phi(t, x^0) \in \mathcal{X}$ gelten. Eine solche Beschreibung der Bewegung ist angemessener als die eingangs gewählte. Die Eigenschaften der Lösungen lassen sich so besser verstehen. Sie werden das sofort sehen. Beim freien Fall setzen wir dazu

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \begin{pmatrix} y \\ y' \end{pmatrix}$$

und erhalten für x die gewöhnliche Differentialgleichung (das System)

$$x' = v(x), \quad x(0) = x^0$$

mit

$$v(x) = \begin{pmatrix} x_2 \\ -1 \end{pmatrix} \quad \text{und} \quad x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Den *Phasenfluß*

$$\Phi : \mathbb{R} \times \mathcal{X} \longrightarrow \mathcal{X}$$

erhalten wir dann durch

$$\Phi(t, x^0) := \Phi_t x^0 := x(t),$$

und es ist $\Phi_0 = id$.

Auch jetzt kann man wieder direkt integrieren. Es ist

$$x(t) = \begin{pmatrix} -\frac{t^2}{2} + x_2^0 t + x_1^0 \\ -t + x_2^0 \end{pmatrix}.$$

Damit ist die Bewegung für alle $t \in \mathbb{R}$ bekannt.

Wie gesagt, in der klassischen Mechanik soll der Fluß Φ determiniert, endlich-dimensional und differenzierbar sein. $\mathcal{X} = \mathbb{R}^2$ ist endlich-dimensional. *Determiniert* soll

$$\Phi_t : \mathcal{X} \longrightarrow \mathcal{X}, \quad \text{bijektiv}$$

$$\Phi_0 = id \quad \text{und} \quad \Phi_{t+s} = \Phi_t \Phi_s$$

bedeuten. Wir suchen also eine einparametrische Gruppe (Φ_t) von differenzierbaren bijektiven Abbildungen des Phasenraumes \mathcal{X} . Bestätigen Sie für unser Beispiel die Kompositionsvorschrift $\Phi_{t+s} = \Phi_t \Phi_s$!

Noch einige Bezeichnungen:

$x : \mathbb{R} \rightarrow \mathcal{X}$ ist die *Bewegung*.

$x(t) = \Phi_t x^0$ ist der *Zustand* zur Zeit t .

$$t \longmapsto (t, \Phi_t x^0) \in \mathbb{R} \times \mathcal{X}$$

heißt *Integralkurve* durch x^0 . Deren Projektion auf \mathcal{X} heißt *Phasenkurve*, *Trajektorie* oder *Orbit* durch x^0 ,

$$\gamma(x^0) := \{ \Phi_t x^0 \mid t \in \mathbb{R} \} \subset \mathcal{X}.$$

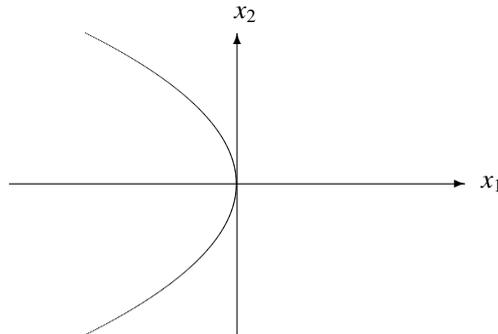
Wir wollen die Orbits für den freien Fall ausrechnen: Aus

$$\gamma(x^0) = \left\{ \begin{pmatrix} -\frac{t^2}{2} + x_2^0 t + x_1^0 \\ -t + x_2^0 \end{pmatrix} \mid t \in \mathbb{R} \right\} =: \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\}$$

folgt $t = x_2^0 - x_2$ und damit

$$(x_2)^2 = -2x_1 + (x_2^0)^2 + 2x_1^0.$$

Für $x_1^0 = x_2^0 = 0$ ergibt sich



Ein Punkt x^0 heißt *Fixpunkt* oder *stationär*, wenn

$$\forall t \in \mathbb{R} \quad \Phi_t x^0 = x^0$$

ist. Ein Orbit $\gamma(x^0)$ heißt *periodisch* mit der Periode p , wenn

$$\exists p > 0, p \text{ minimal: } \forall t \quad \Phi_{t+p} x^0 = \Phi_t x^0$$

gilt.

Wir wollen noch etwas genauer auf den Zusammenhang zwischen Phasenfluß und Differentialgleichung eingehen. Es sei also ein Phasenfluß gegeben. Dann definieren wir die *Phasengeschwindigkeit* oder das *zu Φ gehörende Vektorfeld* v durch

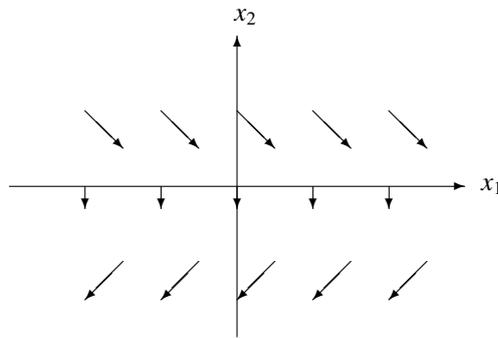
$$v(x) := \partial_t \Phi_t x \Big|_{t=0} \in \mathcal{X}$$

und nennen

$$x' = v(x)$$

die *zu Φ gehörende Differentialgleichung*. Anschaulich gesprochen bedeutet das Lösen der Differentialgleichung also, Orbits zu finden, die in das vorgegebene Vektorfeld v passen. Wir suchen also Kurven, deren Tangenten durch v gegeben sind. Für unser Beispiel ist v zu Beginn der nächsten Seite skizziert.

Man sieht sofort Satz 8.1.1



Satz 8.1.1: Es seien Φ ein Phasenfluß und $x(t) := \Phi_t x^0$. Dann löst x die zu Φ gehörende Differentialgleichung, und es ist $x(0) = x^0$.

Der Beweis folgt unmittelbar aus den Definitionen. $x(0) = x^0$ ist klar. Ferner gilt für $h \rightarrow 0$

$$\frac{x(t+h) - x(t)}{h} = \frac{1}{h} (\Phi_h - id) \Phi_t x^0 \xrightarrow{h \rightarrow 0} v(\Phi_t x^0) = v(x(t)).$$

In der Theorie der gewöhnlichen Differentialgleichungen beschäftigt man sich nun mit der Untersuchung des Zusammenhangs solcher einparametrischer Gruppen von Phasenflüssen und den dazugehörigen Vektorfeldern v . Zum Beispiel möchten wir gerne wissen, wann es zu vorgegebenem v ein Φ_t gibt, also wann $x' = v(x)$ lösbar ist.

Ich erwähne noch, daß ein Punkt $x^0 \in \mathcal{X}$ *singulärer Punkt* des Vektorfeldes v heißt, wenn $v(x^0) = 0$ ist. Man sieht sofort, daß ein Fixpunkt singulär ist. Wenn man die Differentialgleichung $x' = v(x)$ eindeutig lösen kann, ist auch die Umkehrung richtig.

Schließlich bemerke ich noch, daß eine Gleichung der Form

$$x' = v(x)$$

autonom heißt, weil v nur von x und nicht von t abhängt. Die Gleichung

$$x' = w(t, x)$$

mit $x \in \mathcal{X}$ nennt man dann natürlich *nichtautonom*. Auch solche Gleichungen passen in unser Schema, denn man kann jedes nichtautonome System sehr einfach als ein autonomes schreiben. Man wähle nur den Phasenraum $\mathcal{Y} := \mathcal{X} \times \mathbb{R}$ und

$$y := \begin{pmatrix} x \\ t \end{pmatrix} \in \mathcal{Y}.$$

Offenbar gilt dann

$$y' = \begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{pmatrix} w(t, x) \\ 1 \end{pmatrix} =: v(y) \in \mathcal{Y}$$

mit

$$y(0) = y^0 = \begin{pmatrix} x^0 \\ 0 \end{pmatrix}.$$

Dies wollen wir im folgenden immer berücksichtigen.

Ich möchte diese Einführung mit fünf Beispielen abschließen. Zum besseren Verständnis sei schon jetzt darauf hingewiesen, daß wir im nächsten Abschnitt Existenz- und Eindeigkeitssätze für

$$x' = v(x), \quad x(0) = x^0$$

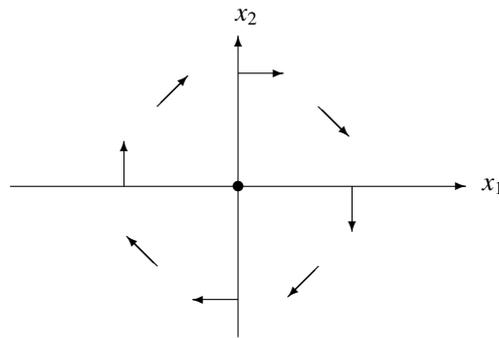
beweisen werden, und zwar unter der Voraussetzung, daß v in \mathcal{X} stetig ist und einer (globalen) *Lipschitzbedingung* genügt, nämlich

$$\exists L \in \mathbb{R}_0^+ \quad \forall x_1, x_2 \in \mathcal{X} \quad |v(x_1) - v(x_2)| \leq L |x_1 - x_2|.$$

Genügt v nur *lokal* einer solchen Bedingung, das heißt gilt nur

$$\forall x \in \mathcal{X} \quad \exists U(x) \subset \mathcal{X} \quad \exists L_x \in \mathbb{R}_0^+ \quad \forall x_1, x_2 \in U(x) \quad |v(x_1) - v(x_2)| \leq L_x |x_1 - x_2|,$$

dann gibt es auch nur lokal (in t) eine Lösung, und es stellt sich die interessante Frage, ob oder wie sie nach einiger Zeit singulär wird.

**Beispiel 8.1.2:**

1. Es seien $\mathcal{X} = \mathbb{R}^2$ und

$$v(x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x$$

Das Vektorfeld v ist linear und genügt deshalb einer globalen Lipschitzbedingung. $x' = v(x)$ ist deshalb eindeutig lösbar. $x^0 = 0$ ist Fixpunkt. Für $x^0 \neq 0$ findet man die Lösung durch einen speziellen Ansatz, nämlich

$$x_1(t) = e^{it}.$$

Aus diesem Ansatz folgt $x_1'' = -x_1$, also $\lambda^2 = -1$ oder

$$x(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} x^0.$$

Für $x^0 \neq 0$ sind alle Orbits mithin periodisch, nämlich Kreise.

2. Es seien $\mathcal{X} = \mathbb{R}^1$ und $v(x) = x^2$. Jetzt genügt v nur noch einer lokalen Lipschitzbedingung. $x^0 = 0$ ist wieder Fixpunkt. Für $x^0 \neq 0$ kann man explizit integrieren und erhält

$$x(t) = \frac{1}{\frac{1}{x^0} - t}.$$

Diese Lösung wird offenbar für $t \rightarrow 1/x_0$ singulär.

3. Es seien $\mathcal{X} = \mathbb{R}^1$ und $v(x) = 2\sqrt{|x|}$. Auch jetzt ist der Punkt $x^0 = 0$ singulär, v genügt aber in $U(x^0)$ keiner Lipschitzbedingung mehr. Die Lösungen sind nicht eindeutig bestimmt, vielmehr besitzt die Differentialgleichung $x' = v(x)$ beliebig viele Lösungen mit $x(0) = x^0$, nämlich $x(t) = 0$ und

$$x_c(t) = \begin{cases} 0 & \text{für } t \leq c \\ (t-c)^2 & \text{für } t > c \end{cases}$$

mit beliebigem $c \geq 0$. Jeder Punkt $c \in \mathbb{R}_0^+$ ist also Verzweigungspunkt.

4. Wir betrachten wieder den freien Fall, diesmal aus großer Höhe vertikal auf die als fest angenommene Erde. Es seien

$R = 6,37 \cdot 10^6$ m der Erdradius,

$M = 5,97 \cdot 10^{27}$ g die Erdmasse,

$G = 6,67 \cdot 10^{-14}$ m³/g sec² die Gravitationskonstante

und h der Abstand vom Erdmittelpunkt. Dann lautet die Newtonsche Bewegungsgleichung

$$m h'' = -m M G h^{-2}.$$

Wählen wir wieder die Variable t geeignet (oder alle Konstanten als Eins) und $\mathcal{X} = [1, \infty) \times \mathbb{R}$, dann erhalten wir im Phasenraum die Differentialgleichung

$$x' = v(x), \quad x(0) = x^0$$

mit

$$v(x) = \begin{pmatrix} x_2 \\ -\frac{1}{(x_1)^2} \end{pmatrix}.$$

Diese Gleichung läßt sich nicht mehr so einfach exakt lösen. Es gibt jedoch ein erstes Integral (d.h. eine Erhaltungsgröße $I : \mathcal{X} \rightarrow \mathbb{R}$ mit $I(x(t)) = I(x^0)$). Aus

$$2 x_2 x_2' = -\frac{2x_2}{(x_1)^2} = -\frac{2x_1'}{(x_1)^2}$$

folgt nämlich

$$I(x) := \frac{(x_2)^2}{2} - \frac{1}{x_1} = \frac{(x_2^0)^2}{2} - \frac{1}{x_1^0}.$$

Wir untersuchen zwei spezielle Fälle:

a. Ein Meteor fällt auf die Erde: Es seien $x_1^0 = \infty$, $x_2^0 = 0$ und t_1 durch $x_1(t_1) = 1$ festgelegt. Dann ist

$$x_2(t_1) = -\sqrt{2} =: -v_0.$$

In üblichen Einheiten entspricht dies einer Aufschlaggeschwindigkeit

$$11,2 \text{ km/sec.}$$

Zum Vergleich, die Schallgeschwindigkeit beträgt 0,33 km/sec und die Lichtgeschwindigkeit $3 \cdot 10^5$ km/sec.

b. Wir berechnen die Fluchtgeschwindigkeit: Es seien jetzt umgekehrt $x_1^0 = 1$ und x_2^0 so gewählt, daß

$$\lim_{t \rightarrow \infty} x_1(t) \rightarrow \infty$$

gilt. Dann folgt

$$x_2(t)^2 \rightarrow (x_2^0)^2 - 2 \stackrel{!}{\geq} 0$$

oder

$$x_2^0 \geq \sqrt{2} = v_0.$$

v_0 heißt deshalb auch Fluchtgeschwindigkeit. Man vergleiche das Phasenbild (Bild 1 in §8.6).

5. Wir betrachten die Schwingungen eines ebenen ungedämpften mathematischen Pendels. Die Punktmasse m sei an einem Faden der Länge l aufgehängt. φ sei der Ausschlagwinkel, $\varphi = 0$ entspreche der Ruhelage. Dann folgt aus den Newtonschen Gesetzen mit $\omega^2 := g/l$

$$\varphi'' + \omega^2 \sin \varphi = 0.$$

Zur Vereinfachung setzen wir $\omega = 1$ und wählen den Phasenraum $\mathcal{X} = \mathbb{R}^2$. Dann lautet unsere Differentialgleichung

$$x' = v(x), \quad x(0) = x^0$$

mit

$$v(x) = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix}.$$

Auch diese Gleichung ist schwer zu integrieren. So wählt man für kleine Schwingungen oft die Näherung

$$v_0(x) = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix},$$

und man diskutiert $x' = v_0(x)$, $x(0) = x^0$. Dies ist unser Beispiel 8.1.2.1. Man erhält für diese Näherung die Schwingungsdauer $T_0 = 2\pi$.

Doch betrachten wir jetzt $x' = v(x)$. Das Vektorfeld v genügt einer (globalen) Lipschitzbedingung. Die Aufgabe ist also für alle Anfangswerte x^0 eindeutig lösbar,

$$x^n := \begin{pmatrix} n\pi \\ 0 \end{pmatrix}, \quad n \in \mathbb{Z}$$

sind Fixpunkte. Ein erstes Integral kann man wieder leicht angeben. Aus

$$2 x_2 x_2' = -2 x_2 \sin x_1 = -2 x_1' \sin x_1$$

folgt

$$E(t) := \frac{1}{2} x_2(t)^2 + (1 - \cos x_1(t)) = \frac{1}{2} (x_2^0)^2 + (1 - \cos x_1^0) \geq 0.$$

E ist die Energie der Bewegung (Summe von kinetischer und potentieller Energie). Unser erstes Integral besagt also, daß die Energie konstant ist, und es folgt

$$x_2(t)^2 = 2\{(E - 1) + \cos x_1(t)\}.$$

Damit lassen sich die Orbits im Phasenraum skizzieren. Für $E = 0$ erhält man die Fixpunkte x_2^n . $E = 1$ ergibt

$$(x_2)^2 = 2 \cos x_1.$$

Das ist ein periodischer Orbit. $E = 2$ ist Grenzlage. Für größere E gibt es keine periodischen Orbits mehr. Man vergleiche das Phasenbild der Bewegung (Bild 2 in §8.6); man wähle auch für \mathcal{X} die Zylinderfläche $S^1 \times \mathbb{R}$ und skizziere dazu das Phasenbild. Man vergleiche auch Beispiel 8.4.13.

Wir wollen noch die Schwingungsdauer $T(x_1^0)$ für $x_2^0 = 0$, $0 < x_1^0 < \pi$, bestimmen. Für die erste Halbschwingung folgt aus dem Energieintegral

$$x_2(t) = -\sqrt{2(\cos x_1(t) - \cos x_1^0)}$$

oder

$$\int_0^t \frac{x_2(s) ds}{\sqrt{\cos x_1(s) - \cos x_1^0}} = -\sqrt{2} t.$$

Es sei τ die erste positive Nullstelle von $x_2(t)$, also

$$x_2(\tau) = 0 \text{ und } x_1(\tau) = -x_1^0.$$

Dann ist

$$\sqrt{2} t = - \int_{x_1^0}^{x_1(t)} \frac{dr}{\sqrt{\cos r - \cos x_1^0}}$$

oder

$$T(x_1^0) = 2\tau = 2\sqrt{2} \int_0^{x_1^0} \frac{dr}{\sqrt{\cos r - \cos x_1^0}}.$$

Dies ist ein elliptisches Integral, für dessen Berechnung es gute Näherungen gibt, so das Gaußsche Verfahren des arithmetisch-geometrischen Mittels. Damit findet man

$$T\left(\frac{\pi}{2}\right)/T_0 = 1,1804 \text{ und } T\left(\frac{\pi}{4}\right)/T_0 = 1,0498.$$

Wenn T konstant sein soll, muß man das Pendel entlang einer *Isochrone* führen. Das ist eine Zykloide

$$x = \begin{pmatrix} rt - a \sin t \\ r - a \cos t \end{pmatrix}.$$

Man spricht dann von einem Zykloidenpendel.

6. Ich skizziere noch kurz das Gaußsche Verfahren zur Berechnung elliptischer Integrale. Es sei

$$I = \frac{1}{\sqrt{2}} \int_0^\alpha \frac{dr}{\sqrt{\cos r - \cos \alpha}} = \frac{1}{2} \int_0^\alpha \frac{dr}{\sqrt{-\sin^2 \frac{r}{2} + \sin^2 \frac{\alpha}{2}}}.$$

Mit $k := \sin \frac{\alpha}{2}$ und $u := \frac{1}{k} \sin \frac{r}{2}$ folgt daraus

$$I = \int_0^1 \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}} = \int_0^{\pi/2} \frac{dt}{\sqrt{1-k^2 \sin^2 t}} =: K(k).$$

Es sei nun $k' := \cos \frac{\alpha}{2}$, also $k^2 + k'^2 = 1$, und

$$\bar{k} := \frac{1-k'}{1+k'} = \frac{1-k'^2}{(1+k')^2} = \frac{k^2}{(1+k')^2}.$$

Dann rechnet man nach (die Rechnung ist in §8.7.2 ausgeführt)

$$K(k) = \int_0^{\pi/2} \frac{dt}{\sqrt{\cos^2 t + k'^2 \sin^2 t}} = \frac{2}{1+k'} K(\bar{k}) = \int_0^{\pi/2} \frac{dt}{\sqrt{(\frac{1+k'}{2})^2 \cos^2 t + k' \sin^2 t}}.$$

Es seien nun

$$\begin{aligned} b_1 &:= k' & \leq & a_1 := 1 \\ b_2 &:= \sqrt{k'} & & a_2 := \frac{1+k'}{2} \\ & \vdots & & \vdots \\ b_{n+1} &:= \sqrt{a_n b_n} & & a_{n+1} := \frac{a_n + b_n}{2}. \end{aligned}$$

Dann folgt aus $a_{n+1}^2 - b_{n+1}^2 = \frac{1}{4}(a_n - b_n)^2 \geq 0$

$$b_n \leq b_{n+1} \leq a_{n+1} \leq a_n.$$

Mithin konvergieren beide Folgen (a_n) und (b_n) , und zwar gegen den gemeinsamen Grenzwert $c(k)$. Die Konvergenz ist quadratisch. Nun ist

$$K(k) = \int_0^{\pi/2} \frac{dt}{\sqrt{a_1^2 \cos^2 t + b_1^2 \sin^2 t}} = \dots = \int_0^{\pi/2} \frac{dt}{\sqrt{a_n^2 \cos^2 t + b_n^2 \sin^2 t}}.$$

Daraus folgt schließlich

$$K(k) = \frac{\pi}{2c(k)}.$$

7. Als weiteres Beispiel wird in §8.7.1 die Bahn eines Satelliten im Feld Erde-Mond berechnet.

8.2 Klassische Existenz- und Eindeutigkeitsätze

In diesem Abschnitt wollen wir klassische Existenz- und Eindeutigkeitsätze beweisen. Wir behandeln sofort nicht-autonome Systeme, also

$$x'(t) = f(t, x(t)), \quad x(0) = x^0. \quad (*)$$

Dabei seien $x \in \mathcal{X} = \mathbb{R}^n$, $J := [0, h]$, $S := J \times \mathcal{X}$ ein Streifen und

$$f : S \longrightarrow \mathcal{X}$$

stetig. Außerdem genüge f in S einer globalen Lipschitzbedingung, nämlich

$$\exists L \geq 0 \quad \forall (t, x_1), (t, x_2) \in S \quad |f(t, x_1) - f(t, x_2)| \leq L|x_1 - x_2|.$$

Dann gilt der nach E. PICARD (1856–1941) und E. LINDELÖF (1870–1946) benannte Satz, nämlich

Satz 8.2.1: *Es sei $x^0 \in \mathcal{X}$. Dann gibt es unter den angegebenen Voraussetzungen an f genau ein $x \in C_1(J, \mathcal{X})$ mit*

1. $x(0) = x^0$
2. $x' = f(\cdot, x)$.

Zum Beweis führen wir die Differentialgleichung in eine Integralgleichung über. Dieses Vorgehen liegt einmal von der Physik her nahe. Ich erinnere daran, daß man die Differentialgleichungen der klassischen Mechanik aus Integralprinzipien – wie dem Hamiltonschen – erhält.

Auch mathematisch besitzt dieses Vorgehen Vorteile. Differentialoperatoren sind unbeschränkt, die entsprechenden Integraloperatoren jedoch oft beschränkt und damit leichter zu behandeln (vgl. §9.2). Man kann auf diese Weise auch den Ableitungsbegriff verallgemeinern und nicht differenzierbare Vorgänge mit einbeziehen. Ich werde in dieser Vorlesung noch oft Differential- und Integralgleichungen nebeneinander betrachten.

Es sei also x eine Lösung der Differentialgleichung. Dann gilt offenbar

$$x(t) = x^0 + \int_0^t f(s, x(s)) ds.$$

Dafür schreiben wir kurz

$$x = T x. \quad (**)$$

Umgekehrt erfüllt eine stetige Lösung von Gl. (**) unsere Differentialgleichung.

Gl. (**) ist eine Fixpunktgleichung. Im Banachschen Fixpunktsatz haben wir bereits ein kräftiges Hilfsmittel zur Lösung solcher Gleichungen gewonnen. Wir wollen ihn anwenden und müssen dazu einen geeigneten Raum stetiger Funktionen finden, in dem T kontrahiert.

Es sei

$$M := C(J, X)$$

mit der Metrik

$$d(f, g) := \sup_{t \in J} |e^{-(L+1)t} (f(t) - g(t))|.$$

(M, d) ist ein vollständiger metrischer Raum; offenbar gilt

$$T : M \longrightarrow M,$$

sowie

$$\forall m \in M \quad (Tm)(0) = x^0.$$

Außerdem folgt aus $f \in C(S, X)$ die Differenzierbarkeit von Tm mit

$$(Tm)'(t) = f(t, m(t)).$$

Wir müssen daher nur noch die Kontraktion von T nachweisen, also

$$\exists k \in [0, 1) \quad d(Tm_1, Tm_2) \leq k d(m_1, m_2).$$

Das folgt mit $k := L/(L+1)$ aus folgender Abschätzung

$$\begin{aligned} d(Tm_1, Tm_2) &= \sup_{t \in J} \left\{ e^{-(L+1)t} \left| \int_0^t \{f(s, m_1(s)) - f(s, m_2(s))\} ds \right| \right\} \\ &\leq \sup_{t \in J} \left\{ e^{-(L+1)t} L \int_0^t |m_1(s) - m_2(s)| ds \right\} \\ &\leq L d(m_1, m_2) \sup_{t \in J} \left\{ e^{-(L+1)t} \int_0^t e^{(L+1)s} ds \right\} \\ &= k d(m_1, m_2) \sup_{t \in J} e^{-(L+1)t} \{e^{(L+1)t} - 1\} \\ &\leq k d(m_1, m_2). \end{aligned}$$

Damit ist der Satz von Picard-Lindelöf bewiesen. Der Beweis ist konstruktiv. Man kann mit einem beliebigen m_0 starten, bildet $m_{n+1} := Tm_n$ und erhält die Lösung von Gl. (**) durch

$$x = \lim_{n \rightarrow \infty} m_n.$$

Es gilt auch die Fehlerabschätzung

$$d(x, m_n) \leq \frac{k^n}{1-k} d(m_1, m_0).$$

All das folgt aus dem Banachschen Fixpunktsatz.

Beispiel 8.2.2: Es seien $X = \mathbb{R}^1$ und $x' = id + x$, $x(0) = 0$. Wir wählen $m_0 = 0$ und erhalten

$$\begin{aligned} m_1(t) &= \int_0^t s ds = \frac{t^2}{2} \\ m_2(t) &= \int_0^t \left(s + \frac{s^2}{2}\right) ds = \frac{t^2}{2} + \frac{t^3}{3!} \\ m_n(t) &= \frac{t^2}{2} + \cdots + \frac{t^{n+1}}{(n+1)!} \end{aligned}$$

und damit $x(t) = e^t - 1 - t$.

Das Resultat von Satz 8.2.1 ist besonders schön, weil die Existenz einer bzgl. t globalen Lösung geliefert wird (das heißt, x existiert für alle $t \in J$). Für die Anwendungen stört es jedoch, daß f einer Lipschitzbedingung überall in S genügen muß. Damit wäre

$$f(t, x) = t + x^2$$

bereits ausgeschlossen. Wie schon in der Einführung angedeutet, wollen wir daher sowohl diese Voraussetzung als auch die Aussage des Satzes abschwächen und lokalisieren. Das ist leicht möglich.

Wir sagen, daß f in $J \times \{x\}$ einer *Lipschitzbedingung* genügt, wenn

$$\exists U(x) \subset \mathcal{X} \quad \exists L_x \in \mathbb{R}_0^+ \quad \forall (t, x_1), (t, x_2) \in J \times U(x) \quad |f(t, x_1) - f(t, x_2)| \leq L_x |x_1 - x_2|$$

gilt. f genügt in S einer Lipschitzbedingung, wenn f für alle $x \in \mathcal{X}$ einer Lipschitzbedingung in $J \times \{x\}$ genügt.

f genüge in S einer Lipschitzbedingung. Dann wählen wir eine Testfunktion $\varphi \in \mathring{C}_\infty(\mathcal{X})$ mit $\varphi = 1$ in einer Umgebung von $x^0 \in \mathcal{X}$. Die Testfunktionen $\mathring{C}_\infty(\mathcal{X})$ sind beliebig oft differenzierbar und haben kompakten Träger. Es sei

$$F := \varphi f.$$

Dann hat F bzgl. \mathcal{X} einen kompakten Träger und genügt deshalb einer globalen Lipschitzbedingung in S . Mithin ist

$$x' = F(\cdot, x), \quad x(0) = x^0$$

nach Satz 8.2.1 eindeutig lösbar. In einer Umgebung von (x^0) ist $F(\cdot, x) = f(\cdot, x)$. Mithin gilt in einer Umgebung von $(0, x^0)$

$$x' = f(\cdot, x), \quad x(0) = x^0.$$

Damit haben wir also eine lokale Lösung unserer Differentialgleichung gefunden.

Diese Lösung ist auch eindeutig. Das folgt – sogar unter allgemeineren Voraussetzungen – aus folgender Überlegung: Es seien x_1, x_2 lokale Lösungen mit $x_i(0) = x^0$. Es sei $[0, a]$ im gemeinsamen Existenzintervall enthalten. Dann gilt für alle $t \in [0, a]$

$$\begin{aligned} |x_1(t) - x_2(t)| &= \left| \int_0^t \{f(s, x_1(s)) - f(s, x_2(s))\} ds \right| \\ &\leq L_{x^0} \int_0^t |x_1(s) - x_2(s)| ds. \end{aligned}$$

Für $\varphi := |x_1 - x_2|$ gilt also $\varphi \geq 0$, $\varphi(0) = 0$ und

$$\varphi(t) \leq L_{x^0} \int_0^t \varphi(s) ds.$$

Dann verschwindet φ . Das folgt aus dem Gronwallschen Lemma, das wir anschließend beweisen. Damit ist gezeigt

Satz 8.2.3: *Es seien $x^0 \in \mathcal{X}$ und $f \in C(S, \mathcal{X})$. f genüge in S einer Lipschitzbedingung. Dann gibt es in einer Umgebung $U(0, x^0)$ genau eine stetig differenzierbare Lösung unserer Differentialgleichung.*

Bevor wir weitermachen, beweise ich das GRONWALLSCHE Lemma (1918), das wir noch mehrfach benutzen werden.

Gronwallsches Lemma: *Es seien $\varphi \in C([0, a], \mathbb{R})$, $h \in \mathcal{L}^1([0, a], \mathbb{R}_0^+)$, $\alpha \in \mathbb{R}$ und*

$$\forall t \in [0, a] \quad \varphi(t) \leq \alpha + \int_0^t h(s) \varphi(s) ds.$$

Dann gilt

$$\varphi(t) \leq \alpha \exp\left(\int_0^t h(s) ds\right).$$

Beweis: Mit $\varepsilon > 0$ sei (zunächst für stetiges h)

$$\psi(t) := (\alpha + \varepsilon) \exp\left(\int_0^t h(s) ds\right).$$

Dann ist $\psi' = h\psi$ oder

$$\psi(t) = \alpha + \varepsilon + \int_0^t h(s) \psi(s) ds.$$

Wir wollen $\varphi < \psi$ zeigen. Für $t = 0$ ist das wegen

$$\varphi(0) \leq \alpha < \alpha + \varepsilon = \psi(0)$$

richtig. Es sei $t_0 \in [0, a]$ die erste Stelle mit $\varphi(t_0) = \psi(t_0)$. Dann ist $\varphi(t) < \psi(t)$ in $[0, t_0)$, und es würde

$$\varphi(t_0) \leq \alpha + \int_0^{t_0} h(s)\varphi(s)ds < \alpha + \varepsilon + \int_0^{t_0} h(s)\psi(s)ds = \psi(t_0)$$

folgen. Also gilt $\varphi < \psi$ in $[0, a]$. Mit $\varepsilon \downarrow 0$ folgt daraus die Behauptung des Lemmas.

Der Eindeutigkeitsnachweis zu Satz 8.2.3 zeigt nun, daß dazu eine Lipschitzbedingung nicht erforderlich ist. So genügt zum Beispiel die Nagumobedingung (nach M. NAGUMO, 1905–)

$$\forall (t, x_1), (t, x_2) \in S \quad |t| \cdot |f(t, x_1) - f(t, x_2)| \leq |x_1 - x_2|.$$

Wir werden in einer Übungsaufgabe (bzw. in §8.7.3) darauf eingehen. Hier gebe ich nur

Beispiel 8.2.4: Es sei $X = \mathbb{R}^1$. Dann gibt es höchstens ein $x \in C_1(J, X)$ mit $x' = f(\cdot, x)$, $x(0) = x^0$ und

$$f(t, x) = \sqrt{t} \sin \frac{x}{t}.$$

f genügt also keiner Lipschitzbedingung. Es gilt aber

$$|f(s, x_1) - f(s, x_2)| = \sqrt{s} \left| \frac{x_1 - x_2}{s} \right| \cdot \left| \cos \frac{\theta}{s} \right| \leq \frac{|x_1 - x_2|}{\sqrt{s}}.$$

Damit erhalten wir

$$|x_1(t) - x_2(t)| = \left| \int_0^t \{f(s, x_1(s)) - f(s, x_2(s))\} ds \right| \leq \int_0^t \frac{|x_1(s) - x_2(s)|}{\sqrt{s}} ds,$$

und aus dem Gronwallschen Lemma folgt $|x_1(t) - x_2(t)| = 0$.

Wir wollen Satz 8.2.3 zur Illustration auf den \mathbb{R}^1 spezialisieren und die *Lebensdauer* der Lösung abschätzen. Es seien $k > 0$ und Q der Quader

$$Q := [0, h] \times [x^0 - k, x^0 + k].$$

f genügt in $J \times \{x^0\}$ einer Lipschitzbedingung. Das heißt, es gilt

$$\exists U(x^0) \quad \exists L_{x^0} \quad \forall (t, x_1), (t, x_2) \in J \times U(x^0) \quad |f(t, x_1) - f(t, x_2)| \leq L_{x^0} |x_1 - x_2|.$$

Wir wählen k so klein, daß

$$Q \subset J \times U(x^0)$$

ist. Es seien ferner

$$A := \max_Q |f| \quad \text{und} \quad a := \min \left(h, \frac{k}{A} \right).$$

Dann gilt für die nach Satz 8.2.3 existierende lokale Lösung x

$$x \in C_1([0, a], X) \quad \text{und} \quad |x - x^0| \leq k.$$

Zum Beweis wählen wir

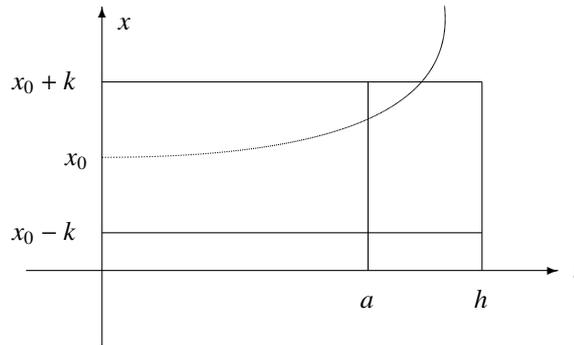
$$F(\cdot, x) = \begin{cases} f(\cdot, x^0 + k) & \text{für } x > x^0 + k \\ f(\cdot, x) & \text{für } x^0 - k \leq x \leq x^0 + k \\ f(\cdot, x^0 - k) & \text{für } x < x^0 - k. \end{cases}$$

Dann genügt F in S einer globalen Lipschitzbedingung, und $y' = F(\cdot, y)$, $y(0) = x^0$ kann eindeutig gelöst werden. Für $t \in [0, a]$ ist aber

$$|y(t) - x^0| = \left| \int_0^t F(s, y(s)) ds \right| \leq A a \leq k.$$

Mithin bleibt y im Quader $[0, a] \times [x^0 - k, x^0 + k]$ und stimmt dort mit x überein. x existiert also in $[0, a]$.

Man wird nun versuchen, x über $t = a$ hinaus fortzusetzen oder mögliche Singularitäten jenseits von a zu charakterisieren. Auf solche Fragen werden wir in §8.4 ein wenig eingehen.



Beispiel 8.2.5: Es seien $\mathcal{X} = \mathbb{R}^1$, $x' = f(\cdot, x)$ und $x(0) = 0$ mit

$$f(t, x) = t^2 + x^2.$$

Dies ist eine *Riccatische Differentialgleichung* (nach J.F. RICCATI, 1676–1754). Wählt man $h = k = 2$, dann ist $A = 8$ und $a = 1/4$. Dieser Wert ist zu klein, die Lebensdauer der Lösung liegt bei 2. Interessant ist der Zusammenhang der nichtlinearen Gleichung mit linearen Gleichungen zweiter Ordnung, den *Besselschen Gleichungen*. Dazu macht man den Ansatz

$$u(t) := \exp\left(-\int_0^t x(s) ds\right)$$

und erhält $u' = -xu$ sowie

$$u''(t) + t^2 u(t) = 0, \quad u(0) = 1, \quad u'(0) = 0.$$

u ist eine *Besselfunktion*, nämlich

$$u(t) = \sqrt{t} Z_\nu(t^2/2) \text{ mit } \nu = \pm 1/4.$$

Die Bessel- oder *Zylinderfunktionen* Z_ν treten in der Physik oft auf. Sie stellen gedämpfte Schwingungen dar und genügen der Differentialgleichung

$$t^2 Z_\nu''(t) + t Z_\nu'(t) + (t^2 - \nu^2) Z_\nu(t) = 0.$$

Unabhängige Lösungen sind für $\nu \notin \mathbb{Z}$ die Besselfunktionen

$$\begin{aligned} Z_\nu(t) &= t^\nu \{c_{\nu 1} + O(t^2)\}, \\ Z_{-\nu}(t) &= t^{-\nu} \{c_{\nu 2} + O(t^{2\nu} + t^2)\}. \end{aligned}$$

Für $u(t)$ kommen also in Betracht (mit $c_i \neq 0$)

$$\begin{aligned} u_1(t) &= \sqrt{t} Z_\nu(t^2/2) = t \{c_1 + \dots\}, \\ u_2(t) &= \sqrt{t} Z_{-\nu}(t^2/2) = c_2 + \dots, \end{aligned}$$

und aus $u(0) = 1, u'(0) = 0$ folgt mit $\nu = 1/4$

$$u(t) = c \sqrt{t} Z_{-\nu}(t^2/2).$$

Daraus erhält man

$$x(t) = -\frac{u'(t)}{u(t)}.$$

Die erste Nullstelle von u ist also die Lebensdauer unserer Lösung x .

Wir kehren nun zum Existenzproblem zurück. Ich erinnere an Beispiel 8.1.2.3, das ich leicht abgewandelt wiederhole.

Beispiel 8.2.6: Es sei $\mathcal{X} = \mathbb{R}^1$, $x' = f(\cdot, x)$ und $x(0) = 0$ mit

$$f(t, x) = \begin{cases} 0 & \text{für } x < 0 \\ 2\sqrt{t} & \text{für } 0 \leq x \leq 1 \\ 2 & \text{für } 1 < x. \end{cases}$$

f genügt in $J \times \{0\}$ keiner Lipschitzbedingung, und zwar liegt die Schwierigkeit bei diesem Beispiel nicht am Verhalten für große x , sondern bei $x = 0$. Für $x^0 \neq 0$ genügt f einer Lipschitzbedingung, und man kann lokal

eindeutig lösen. Wie wir aber am Beispiel 8.1.2.3 gesehen haben, gibt es zu $x(0) = 0$ beliebig viele Lösungen. Ich wiederhole das Beispiel, um zu betonen, daß die Lipschitzbedingung auch zum Nachweis der Existenz einer Lösung nicht notwendig ist.

Wir wollen deshalb den Existenzsatz ebenfalls verbessern und den Satz von Peano beweisen, der die Existenz einer Lösung schon für nur stetige f liefert. Wir hatten den Satz von Picard-Lindelöf mit Hilfe des Fixpunktsatzes hergeleitet. Auch zum Beweis des Satzes von Peano benötigen wir ein tieferliegendes analytisches Hilfsmittel, nämlich einen Auswahlssatz, den Satz von Arzelà-Ascoli. Diesen Satz werden wir im weiteren Verlauf der Vorlesung noch mehrfach benutzen. Ich unterbreche deshalb hier die Behandlung der Differentialgleichungen kurz und schiebe einen wichtigen Satz von allgemeinem Interesse ein, der auf C. ARZELÀ (1847–1912) und G. ASCOLI (1843–1896) zurückgeht. Dazu geben wir folgende

Definition 8.2.7: Es seien $G \subset \mathbb{R}^n$ und $f_n \in C(G, \mathbb{R}^m)$. Dann heißt (f_n) gleichartig stetig in $G : \iff$

$$\forall x \in G \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall y \in G, |x - y| < \delta \quad \forall n \in \mathbb{N} \quad |f_n(x) - f_n(y)| < \varepsilon.$$

Die Folge (f_n) heißt gleichmäßig gleichartig stetig in $G : \iff$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in G, |x - y| < \delta \quad \forall n \in \mathbb{N} \quad |f_n(x) - f_n(y)| < \varepsilon.$$

Dann gilt

Satz 8.2.8: Es seien $G \subseteq \mathbb{R}^n$, $f_n \in C(\bar{G}, \mathbb{R}^m)$ sowie (f_n) punktweise beschränkt und gleichartig stetig in \bar{G} . Dann gibt es eine Teilfolge $(f_{n'})$ und ein $f \in C(\bar{G}, \mathbb{R}^m)$ mit

$$\|f_{n'} - f\| \rightarrow 0.$$

Der Beweis des Satzes von Arzelà-Ascoli erfolgt in zwei Schritten:

1. Wir zeigen, daß (f_n) gleichmäßig gleichartig stetig ist: Indirekt, es gelte

$$\exists \varepsilon \quad \forall \delta \quad \exists x, y, |x - y| < \delta \quad \exists n \quad |f_n(x) - f_n(y)| \geq \varepsilon.$$

Dann wählen wir $\delta_k := 1/k$ und dazu x_k, y_k, n_k . Weil G beschränkt ist, gibt es eine konvergente Teilfolge der (x_k) . Es sei bereits (x_k) diese Folge und $x_k \rightarrow z \in \bar{G}$. Dann gilt auch $y_k \rightarrow z$ und

$$|f_{n_k}(x_k) - f_{n_k}(y_k)| \geq \varepsilon$$

im Widerspruch zu

$$|f_{n_k}(x_k) - f_{n_k}(y_k)| \leq |f_{n_k}(x_k) - f_{n_k}(z)| + |f_{n_k}(z) - f_{n_k}(y_k)| < \varepsilon,$$

denn (f_n) ist in \bar{G} gleichartig stetig.

2. Es sei nun Y eine überall dichte abzählbare Punktmenge in \bar{G} , zum Beispiel $Y := \mathbb{Q}^n \cap \bar{G}$. Es seien $Y = \{y_1, y_2, \dots\}$, $f_{n,0} := f_n$, $(f_{n,i})$ eine in y_i konvergente Teilfolge von $(f_{n,i-1})$ und $(f_{n,n})$ die Diagonalfolge. Es sei bereits (f_n) diese Diagonalfolge. Dann konvergiert (f_n) in \bar{G} gleichmäßig. Um das zu zeigen, wählen wir zu $\varepsilon > 0$ ein $\delta(\varepsilon)$ mit

$$\forall x, y \in \bar{G}, |x - y| < \delta(\varepsilon) \quad \forall n \quad |f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$$

und eine endliche Teilmenge Z von Y mit

$$\forall x \in \bar{G} \quad \exists z \in Z \quad |x - z| < \delta(\varepsilon).$$

Ferner wählen wir ein $N(\varepsilon)$ mit

$$\forall z \in Z \quad \forall m, n \geq N(\varepsilon) \quad |f_n(z) - f_m(z)| < \frac{\varepsilon}{3}.$$

Dann folgt für alle $x \in \bar{G}$ und alle $n, m \geq N(\varepsilon)$

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f_m(z)| + |f_m(z) - f_n(z)| + |f_n(z) - f_n(x)| < \varepsilon.$$

Das war zu zeigen.

Wir kehren nun zur Theorie der Differentialgleichungen zurück und sind in der Lage, folgenden Satz von PEANO zu beweisen:

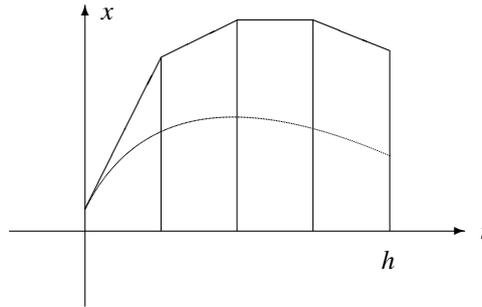
Satz 8.2.9: Es seien $x^0 \in X$ und $f \in C_b(S, X)$. Dann gibt es ein $x \in C_1(J, X)$ mit $x(0) = x^0$ und

$$x' = f(\cdot, x).$$

Der Beweis ist sehr anschaulich und benutzt im Grunde ein numerisches Verfahren. Solche Existenzbeweise durch Approximation findet man oft in der Theorie der Differentialgleichungen, deshalb auch die große Bedeutung der Computer in diesem Gebiet. Es seien

$$t_{m,\mu} := \frac{\mu h}{2^m} \quad \text{mit } \mu = 0, 1, \dots, 2^m.$$

Wir approximieren x nun durch einen Polygonzug, der in diesen Punkten jeweils den richtigen Anstieg hat (vgl. die Abbildung),



nämlich $x_m(0) := x^0$ und für $t_{m,\mu} < t \leq t_{m,\mu+1}$, $0 \leq \mu \leq 2^m - 1$

$$x_m(t) := x_m(t_{m,\mu}) + (t - t_{m,\mu})f(t_{m,\mu}, x_m(t_{m,\mu})) \quad (1)$$

x_m ist stetig und in jedem Teilintervall linear. Zur Abkürzung schreiben wir

$$\begin{aligned} g_m(0) &:= f(0, x^0) \\ g_m(t) &:= f(t_{m,\mu}, x_m(t_{m,\mu})) \quad \text{für } t_{m,\mu} < t \leq t_{m,\mu+1}. \end{aligned}$$

Dann gilt

$$x_m(t) = x_m(t_{m,\mu}) + \int_{t_{m,\mu}}^t g_m(s) ds.$$

Nun ist

$$\begin{aligned} x_m(t_{m,\mu}) &= x_m(t_{m,0}) + \sum_{\nu=1}^{\mu} \{x_m(t_{m,\nu}) - x_m(t_{m,\nu-1})\} \\ &= x^0 + \sum_{\nu=1}^{\mu} \int_{t_{m,\nu-1}}^{t_{m,\nu}} g_m(s) ds, \end{aligned}$$

also

$$x_m(t) = x^0 + \int_0^t g_m(s) ds. \quad (2)$$

Wir wollen zeigen, daß (x_m) eine gleichmäßig konvergente Teilfolge enthält. Mit

$$c := \sup_{(t,x) \in S} |f(t, x)| < \infty$$

ist auch

$$\forall t \in J \quad |g_m(t)| \leq c.$$

Mithin sind (g_m) und (x_m) in J beschränkt. Für $t_1, t_2 \in J$ gilt

$$|x_m(t_1) - x_m(t_2)| = \left| \int_{t_1}^{t_2} g_m(s) ds \right| \leq c |t_1 - t_2|.$$

Die Folge (x_m) ist deshalb bezüglich m gleichartig stetig in J , und aus dem Satz von Arzelà–Ascoli folgt die Existenz einer konvergenten Teilfolge. Es sei (x_m) bereits diese konvergente Teilfolge. Dann gibt es ein $x \in C(J, X)$ mit

$$\|x_m - x\| \rightarrow 0.$$

Für festes $t \in J$ war

$$g_m(t) = f(t_{m,\mu}, x_m(t_{m,\mu}))$$

mit $\mu = \mu(t)$, so daß $t_{m,\mu} < t \leq t_{m,\mu+1}$. Mithin ist

$$|t_{m,\mu(t)} - t| \leq \frac{h}{2^m} \rightarrow 0,$$

und aus Gl. (1) erhalten wir

$$|x_m(t_{m,\mu}) - x_m(t)| \leq c |t - t_{m,\mu}| \leq \frac{ch}{2^m} \rightarrow 0.$$

f ist in jedem $K \Subset S$ gleichmäßig stetig. Mithin konvergiert (g_m) in J gleichmäßig,

$$\|g_m - f(\cdot, x)\| \rightarrow 0,$$

und aus Gl. (2) folgt

$$x(t) = x^0 + \int_0^t f(s, x(s)) ds.$$

Das war zu zeigen.

Wenn man weiß, daß die Aufgabe eindeutig lösbar ist (wenn f also zum Beispiel eine Nagumo-Bedingung erfüllt), dann ist das angegebene Verfahren konstruktiv. Das sei noch einmal betont. Denn dann enthält jede Folge (x_m) eine Teilfolge, die gegen denselben Grenzwert konvergiert. Mithin konvergiert (x_m) selbst.

Das angegebene Verfahren läßt sich für numerische Zwecke verbessern. Ich möchte solche Fragen in dieser Vorlesung nicht weiter vertiefen. Das geschilderte Verfahren heißt *explizites Euler-Verfahren* und hat die Konvergenzordnung eins. *Phaser* verwendet auch das *verbesserte Euler-Verfahren* mit der Konvergenzordnung zwei und das *Runge-Kutta-Verfahren* mit der Konvergenzordnung vier.

Neben der Frage nach der Existenz und Eindeutigkeit der Lösungen ist in den Anwendungen besonders auch die Frage nach ihrer stetigen Abhängigkeit von den Anfangswerten und Koeffizienten von besonderer Bedeutung. Wir wollen auch diese Probleme anschneiden. Als erstes zeigen wir

Satz 8.2.10: f genüge einer globalen Lipschitzbedingung in S . Es seien $x^0, y^0 \in X$ und

$$\begin{aligned} x' &= f(\cdot, x), & x(0) &= x^0 \\ y' &= f(\cdot, y), & y(0) &= y^0. \end{aligned}$$

Dann gilt mit $k := L/(L+1) < 1$

$$d(x, y) \leq \frac{|x^0 - y^0|}{1 - k}.$$

Der Satz besagt also die stetige Abhängigkeit von den Anfangswerten. Der Beweis folgt dem von Satz 8.2.1. Es ist

$$x(t) - y(t) = x^0 - y^0 + \int_0^t \{f(s, x(s)) - f(s, y(s))\} ds,$$

also

$$\begin{aligned} d(x, y) &\leq |x^0 - y^0| + L \sup_{t \in J} e^{-(L+1)t} \int_0^t |x(s) - y(s)| ds \\ &\leq |x^0 - y^0| + k d(x, y). \end{aligned}$$

Es gilt auch

Satz 8.2.11: Es seien $f, g \in C(S, X)$ mit $\|f - g\| < \infty$, und f genüge in S einer globalen Lipschitzbedingung mit der Lipschitzkonstanten L . Es seien ferner x, y in J Lösungen von

$$\begin{aligned} x' &= f(\cdot, x), & x(0) &= x^0 \\ y' &= g(\cdot, x), & y(0) &= x^0. \end{aligned}$$

Dann gilt mit $k = L/(L+1) < 1$

$$d(x, y) \leq \frac{h}{1 - k} \|f - g\|.$$

Der Beweis folgt wieder aus

$$x(t) - y(t) = \int_0^t \{f(s, x(s)) - f(s, y(s))\} ds + \int_0^t \{f(s, y(s)) - g(s, y(s))\} ds$$

oder

$$d(x, y) \leq k d(x, y) + h \|f - g\|.$$

8.3 Spezielle Lösungsmethoden

In diesem Abschnitt wollen wir uns einige spezielle Gleichungen näher ansehen und die entsprechenden Lösungsmethoden kennenlernen. Es sei wieder $\mathcal{X} = \mathbb{R}^n$ und

$$x' = v(x), \quad x(0) = x^0.$$

Das Vektorfeld v möge einer globalen Lipschitzbedingung in \mathcal{X} genügen. Dann wissen wir, daß es für jedes $T > 0$ eine eindeutige Lösung in $[0, T]$ und $[-T, 0]$ gibt. Weil $x'(t)$ sowohl für $t \downarrow 0$ als auch für $t \uparrow 0$ existiert und beide Limites übereinstimmen, gilt $x \in C_1(\mathbb{R}, \mathcal{X})$. Es sei

$$\Phi_t x^0 := x(t)$$

der zugehörige Fluß. Dann ist $\Phi_0 = id$ und

$$\Phi_{t+s} = \Phi_t \Phi_s. \quad (*)$$

Letzteres folgt aus der eindeutigen Lösbarkeit der Differentialgleichung. Mit

$$y(t) := \Phi_{t+s} x^0 \quad \text{und} \quad z(t) := \Phi_t(\Phi_s x^0)$$

folgt nämlich $y(0) = z(0)$ und

$$\begin{aligned} y'(t) &= \frac{d}{d(t+s)} \Phi_{t+s} x^0 = v(\Phi_{t+s} x^0) = v(y(t)) \\ z'(t) &= v(\Phi_t(\Phi_s x^0)) = v(z(t)). \end{aligned}$$

Mithin ist in diesem einfachsten Fall (Φ_t) wirklich eine einparametrische Gruppe stetig differenzierbarer bijektiver Abbildungen. Die Gleichung für das ungedämpfte Pendel (Beispiel 8.1.2.5) gehört hierhin. Noch einfacher sind natürlich lineare Gleichungen, die wir zuerst diskutieren wollen.

8.3.1 Lineare Systeme

Es sei wie im ersten Abschnitt $J := [0, h]$ und $S := J \times \mathcal{X}$. Wir beginnen mit dem nichtautonomen linearen System

$$x' = F x + f, \quad x(0) = x^0. \quad (1)$$

Dabei sei $F \in C(J, \mathcal{L}(\mathcal{X}))$ eine $n \times n$ Matrix und $f \in C(J, \mathcal{X})$ ein Vektor.

$$v(\cdot, x) := F x + f$$

genügt in S einer globalen Lipschitzbedingung.

Wir behandeln zuerst das homogene System

$$x' = F x, \quad x(0) = x^0. \quad (2)$$

Offenbar besitzt Gl. (2) genau n linear unabhängige Lösungen. Es seien für $i = 1, \dots, n$ die y^i Lösungen mit $y^i(0) = e^i$ und

$$\Psi_t := (y^1(t), \dots, y^n(t)).$$

Dann gilt $\Psi_0 = id$, und

$$x(t) := \Psi_t x^0 = \sum_{i=1}^n y^i(t) x_i^0$$

ist die eindeutige Lösung von Gl. (2). Weil diese Gleichung als nichtautonom angenommen wurde, erfolgt die Komposition von Ψ jedoch im allgemeinen nicht gemäß (*).

Es gilt

Satz 8.3.1: Es seien x^1, \dots, x^n Lösungen von $x' = F x$ und

$$\Delta(t) := \det(x^1(t), \dots, x^n(t)).$$

Dann sind x^1, \dots, x^n in J linear abhängig \iff

$$\exists t_0 \in J \quad \Delta(t_0) = 0.$$

In diesem Falle gilt sogar

$$\forall t \in J \quad \Delta(t) = 0.$$

Beweis: Daß die Determinante verschwindet, wenn die Lösungen linear abhängig sind, ist klar. Es sei also für ein $t_0 \in J$ $\Delta(t_0) = 0$. Dann gibt es Konstanten $(c_1, \dots, c_n) \neq (0, \dots, 0)$ und ein

$$x(t) := \sum_{i=1}^n c_i x^i(t)$$

mit $x(t_0) = 0$. Wegen der Eindeutigkeit folgt daraus aber $x(t) = 0$ für alle $t \in J$.

Natürlich gilt ein solches Kriterium nur für Lösungen linearer Gleichungssysteme. Es seien

$$x^1(t) := \begin{pmatrix} t \\ 0 \end{pmatrix} \quad \text{und} \quad x^2(t) := \begin{pmatrix} t^2 \\ 0 \end{pmatrix}.$$

Dann ist $\Delta(t) = 0$; x^1 und x^2 sind jedoch in J linear unabhängig.

Die Determinante in Satz 8.3.1 ist nach H. WRONSKI (1776–1853) benannt. Sie läßt sich leicht berechnen. Es gilt nämlich

$$\Delta(t) = \Delta(t_0) \exp\left(\int_{t_0}^t \text{Spur } F(s) ds\right).$$

Der Beweis erfolgt durch Nachrechnen. Es ist

$$\det(x^1, \dots, x^n) = \det(x^1, \dots, x^n)^t =: \det(z^1, \dots, z^n)$$

und

$$(z^j)' = \begin{pmatrix} x_j^1 \\ \vdots \\ x_j^n \end{pmatrix}' = \sum_{i=1}^n F_{ji} \begin{pmatrix} x_i^1 \\ \vdots \\ x_i^n \end{pmatrix} = \sum_{i=1}^n F_{ji} z^i.$$

Daraus folgt

$$\begin{aligned} \Delta' &= \sum_{j=1}^n \det(z^1, \dots, z^{j'}, \dots, z^n) \\ &= \sum_{j=1}^n \det(z^1, \dots, F_{jj} z^j, \dots, z^n) \\ &= \Delta \cdot \text{Spur } F. \end{aligned}$$

Wir behandeln nun die inhomogene Gleichung (1). Es sei p eine partikuläre Lösung dieser Gleichung; das heißt, es sei $p' = Fp + f$ mit beliebigem $p(0) =: p^0$. Dann erhält man die eindeutige Lösung von Gl. (1) durch

$$x = p + \Psi_t(x^0 - p^0).$$

Eine partikuläre Lösung kann man mit der *Methode der Variation der Konstanten* berechnen. Dazu macht man den Ansatz

$$p(t) = \Psi_t c(t).$$

Wäre c konstant, dann würde man die homogene Gleichung lösen. So aber erhält man

$$p'(t) = F(t)\Psi_t c(t) + \Psi_t c'(t) = F(t)p(t) + \Psi_t c'(t).$$

Wir erhalten also eine partikuläre Lösung, wenn wir c mit

$$\Psi_t c'(t) = f(t) \quad \text{oder} \quad c'(t) = \Psi_t^{-1} f(t)$$

wählen. Mithin ist

$$p(t) := \Psi_t \int_0^t \Psi_\tau^{-1} f(\tau) d\tau. \quad (3)$$

eine partikuläre Lösung.

Systeme mit konstanten Koeffizienten: Es sei nun $F(t) = A$ eine $n \times n$ -Matrix mit konstanten Koeffizienten. Gl. (2) ist jetzt autonom. Fx genügt einer globalen Lipschitzbedingung in $\mathbb{R} \times X$, der Fluß Φ_t existiert, und man kann ihn explizit angeben. Der einfachste Fall ist der einer diagonalähnlichen Matrix A , das heißt

$$\exists T \quad T^{-1}AT = D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \diagdown & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Dieser Fall tritt auf, wenn A n linear unabhängige Eigenvektoren u^j , $Au^j = \lambda_j u^j$ besitzt; zum Beispiel wenn alle λ_j verschieden sind oder A symmetrisch ist. Dann ist T die Matrix dieser Eigenvektoren,

$$T = (u^1, u^2, \dots, u^n).$$

Die Gleichung $y' = Dy$ wird durch $y(t) = \Psi_t y^0$ mit

$$\Psi_t = \begin{pmatrix} e^{\lambda_1 t} & \cdots & 0 \\ \vdots & \diagdown & \vdots \\ 0 & \cdots & e^{\lambda_n t} \end{pmatrix}$$

gelöst. Das heißt, es ist $\Psi_0 = id$ und $y' = Dy$. Daraus folgt

$$Ty' = AT\Psi_t T^{-1} Ty^0$$

oder mit $y^0 := T^{-1}x^0$ und $x := Ty$

$$x' = A \Phi_t x^0.$$

$\Phi_t := T\Psi_t T^{-1}$ ist daher der Fluß. Auch der Fall allgemeiner Matrizen A läßt sich vollständig durchrechnen. Man benutzt dazu die Jordansche Normalform von A . Ich verzichte hier auf Einzelheiten.

Auch die inhomogene Gleichung läßt sich natürlich lösen. Formel (3) für die partikuläre Lösung vereinfacht sich jetzt zu

$$p(t) = \int_0^t \Phi_{t-\tau} f(\tau) d\tau.$$

Beispiel 8.3.2: Wir diskutieren im \mathbb{R}^2

$$x' = Ax, \quad x(0) = x^0$$

mit $\det A \neq 0$.

Der Punkt $x^0 := 0$ ist Fixpunkt. Für A gibt es drei Grundtypen, nämlich

$$A_1: \quad A \text{ hat die Normalform } \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \text{mit } \lambda_i \in \mathbb{R},$$

$$A_2: \quad A \text{ hat die Normalform } \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad \text{mit } \lambda \in \mathbb{R}$$

und

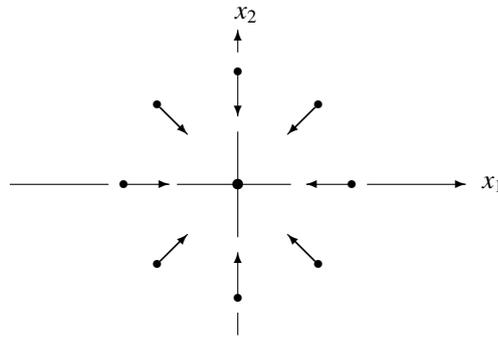
$$A_3: \quad A \text{ hat die Normalform } \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \quad \text{mit } \operatorname{Im} \lambda \neq 0.$$

A_1 und A_3 sind diagonalähnlich. Wir wollen diese Fälle jedoch getrennt diskutieren und beginnen mit A_1 . Es sei

1. $\lambda_1 \cdot \lambda_2 > 0$. Dann ist

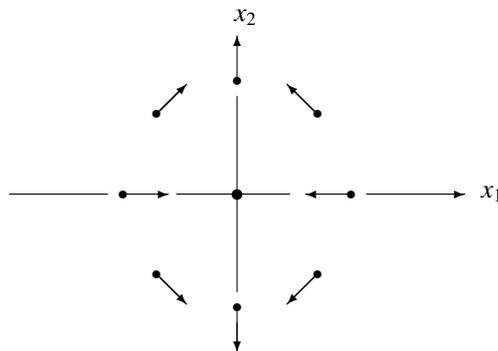
$$\Psi_t = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix},$$

und für $\lambda_i < 0$ und $t \rightarrow \infty$ sehen die Richtungen der Orbits folgendermaßen aus:



x^0 ist in diesem Falle ein *Knoten* oder ein *positiver* bzw. *negativer Attraktor*, auch eine *Senke* bzw. *Quelle*; im Falle $\lambda_1 = \lambda_2$ auch ein *Brennpunkt* oder *Fokus*. Die Orbits sind dann Halbgeraden in den Nullpunkt.

2. Es sei $\lambda_1 \cdot \lambda_2 < 0$. Jetzt erhält man etwa für $\lambda_1 < 0$ und $t \rightarrow \infty$ die Richtungen der Orbits in der folgenden Skizze. x^0 nennt man in diesem Falle *Sattelpunkt*.



Wir diskutieren nun die Bewegung im Falle A_2 . Jetzt ist e^1 Eigen- und e^2 Hauptvektor, das heißt $(A - \lambda id)e^2 = e^1$. Daraus folgt

$$\Psi_t = e^{\lambda t}(e^1, e^2 + t e^1).$$

Auch jetzt ist x^0 ein Knoten.

Im Falle A_3 sei $\lambda = \mu + i\omega$ mit $\omega \neq 0$. Es seien ferner $Au = \lambda u$ mit $u = u^1 + i u^2$ und

$$U := (u^1, u^2).$$

U ist reellwertig und nicht singulär. Dann erhält man $x(t)$ in der Form

$$x(t) = \alpha x^1(t) + \beta x^2(t)$$

mit

$$\begin{aligned} x^1(t) &:= \operatorname{Re} e^{\lambda t} u = e^{\mu t} \{u^1 \cos \omega t - u^2 \sin \omega t\} \\ x^2(t) &:= \operatorname{Im} e^{\lambda t} u = e^{\mu t} \{u^2 \cos \omega t + u^1 \sin \omega t\}. \end{aligned}$$

Daraus folgt

$$x(t) = e^{\mu t} U \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = e^{\mu t} U \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} U^{-1} x^0,$$

also

$$\Phi_t = e^{\mu t} U \begin{pmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{pmatrix} U^{-1}.$$

Im Falle $\mu = 0$ sind die Orbits periodische Bahnen um den Nullpunkt. x^0 heißt dann *Zentrum*. Im Falle $\mu \neq 0$ ist x^0 ein *Strudelpunkt* oder eine *Spirale*. Beispielsweise erhält man für

$$A = \begin{pmatrix} 0 & -1/2 \\ 2 & 0 \end{pmatrix}$$

$\lambda = i$ und $u = (i, 2)^t$, also

$$\Phi_t = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} 0 & 1/2 \\ 1 & 0 \end{pmatrix},$$

und damit die Orbits

$$\left((U^{-1}x)_1\right)^2 + \left((U^{-1}x)_2\right)^2 = \frac{x_2^2}{4} + x_1^2 = \text{const.}$$

Gleichungen höherer Ordnung: Wir betrachten kurz im \mathbb{R}^1

$$y^{(n)} = g(\cdot, y, y', \dots, y^{(n-1)}).$$

$y(0), \dots, y^{(n-1)}(0)$ seien vorgegeben. Solche Gleichungen lassen sich leicht als Systeme erster Ordnung umschreiben. Es seien nämlich

$$x_1 := y, \quad x_2 := y', \quad \dots, \quad x_n := y^{(n-1)}.$$

Dann folgt für x das nichtautonome System

$$x' = Ax + B(x), \quad x(0) = x^0$$

mit $x^0 = (y(0), \dots, y^{(n-1)}(0))$ und

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \searrow & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g \end{pmatrix}.$$

Von besonderem Interesse sind auch hier lineare Gleichungen der Form

$$y^{(n)} - f_{n-1}y^{(n-1)} - \cdots - f_0y = f \tag{4}$$

mit $f, f_i \in C(\mathbb{R}^1, \mathbb{R}^1)$. Ihnen entspricht

$$x' = Fx + \tilde{f}, \quad x(0) = x^0 \tag{5}$$

mit

$$F = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \searrow & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ f_0 & f_1 & f_2 & \cdots & f_{n-1} \end{pmatrix} \quad \text{und} \quad \tilde{f} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f \end{pmatrix}.$$

Damit ist Gl. (4) auf eine bereits behandelte zurückgeführt. Wegen ihrer großen praktischen Bedeutung wollen wir jedoch kurz die erhaltenen Ergebnisse für diesen Spezialfall neu formulieren.

Die Wronskische Determinante hat jetzt eine besonders einfache Form, nämlich

$$\Delta(t) = \Delta(t_0) \exp\left(\int_{t_0}^t f_{n-1}(\tau) d\tau\right).$$

Steht y^i für den Vektor $(y^i, y^{i'}, \dots, y^{i^{(n-1)}})^t$ der i -ten Lösung, dann ist

$$\Delta = \det(x^1, \dots, x^n) = \det(y^1, y^2, \dots, y^n) =: W(y^1, y^2, \dots, y^n).$$

Wenn man $n-1$ Lösungen y^1, \dots, y^{n-1} der homogenen Gleichung (4) kennt, läßt sich aus der Wronskischen Determinante die n -te berechnen. Man muß nur

$$W(y^1, \dots, y^{n-1}, y)(t) = \exp\left(\int_0^t f_{n-1}(\tau) d\tau\right)$$

etwa mit der Methode der Variation der Konstanten lösen. Bezüglich der Einzelheiten der Rechnung sei auf die Übung verwiesen.

Auch jetzt sind Gleichungen mit konstanten Koeffizienten von besonderem Interesse, also

$$y^{(n)} - a_{n-1}y^{(n-1)} - \dots - a_0y = f. \quad (6)$$

Zur Lösung der homogenen Gleichung macht man den Ansatz $\exp(\lambda t)$ und erhält für λ die charakteristische Gleichung

$$0 = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_0 = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_r)^{k_r}$$

mit $k_1 + \dots + k_r = n$.

$$\begin{aligned} y_{11}(t) &= e^{\lambda_1 t}, \quad y_{12}(t) = t e^{\lambda_1 t}, \quad \dots, \quad y_{1k_1}(t) = t^{k_1-1} e^{\lambda_1 t} \\ &\vdots \\ y_{r1}(t) &= e^{\lambda_r t}, \quad \dots \end{aligned}$$

sind dann linear unabhängige Lösungen der homogenen Gleichung. Es sei φ die Lösung der homogenen Gleichung mit

$$\varphi(0) = \varphi'(0) = \dots = \varphi^{(n-2)}(0) = 0, \quad \varphi^{(n-1)}(0) = 1.$$

Dann löst

$$p(t) = \int_0^t \varphi(t-\tau)f(\tau)d\tau$$

die inhomogene Gleichung.

Lineare Gleichungen zweiter Ordnung treten bei vielen Schwingungsproblemen in der Physik auf. So das folgende

Beispiel 8.3.3:

$$\left(Ly'' + Ry' + \frac{1}{C}y \right)(t) = k e^{i\omega t}.$$

Dabei sind y die Stromstärke, L die Induktivität, R der Widerstand und C die Kapazität eines elektrischen Kreises. Auf der rechten Seite steht die Anregung. Die charakteristischen λ -Werte lauten

$$\lambda_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{CL}}.$$

Im Falle positiver R erhält man gedämpfte Schwingungen. Im Falle $R = 0$ ist

$$\omega_0 := \sqrt{\frac{1}{CL}}$$

die Resonanzfrequenz. Dieses Beispiel läßt sich vollständig durchrechnen. Wegen der Einzelheiten sei wieder auf die Übung bzw. §8.7.4 verwiesen.

Zum Abschluß bringe ich noch ein Beispiel, das Sie an die Operatorenrechnung heranführen soll.

Beispiel 8.3.4: Berechnung partikulärer Lösungen.

Es seien

$$D := \frac{d}{dt}, \quad R(0)f := \int f$$

ein unbestimmtes Integral und für $a \neq 0$

$$R(a)f := -\frac{1}{a}\left(1 + \frac{D}{a} + \frac{D^2}{a^2} + \dots\right)f,$$

sofern diese Reihe konvergiert. $R(a)$ ist also zum Beispiel für Polynome erklärt. Formal ist $R(a)$ die geometrische Reihe zu

$$(D - a)^{-1} = -\frac{1}{a}\left(1 - \frac{D}{a}\right)^{-1} = R(a).$$

$R(a)$ läßt sich aber auch für Funktionen wie e^x , $\sin x$, $\cos x$ usw. erklären; das sind alles Funktionen, die in der praktischen Schwingungstheorie häufig auftreten. Zum Beispiel seien

$$\begin{aligned} R(a) e^{\alpha t} &:= \frac{e^{\alpha t}}{\alpha - a} \quad \text{für } a \neq \alpha \\ R(a) e^{\alpha t} &:= t e^{\alpha t} \\ R(a) \sin \alpha t &:= -\frac{a}{a^2 + \alpha^2} \sin \alpha t - \frac{\alpha}{a^2 + \alpha^2} \cos \alpha t \\ R(a) \cos \alpha t &:= \frac{\alpha}{a^2 + \alpha^2} \sin \alpha t - \frac{a}{a^2 + \alpha^2} \cos \alpha t. \end{aligned}$$

In all diesen Fällen gilt also

$$(D - a)R(a)f = f \quad \text{und} \quad R(a)(D - a)f = f.$$

Betrachten wir speziell

$$y'' - 3y' + 2y = t$$

oder

$$(D - 1)(D - 2)y = t.$$

Dem entspricht

$$(D - 2)y = R(1)t = -(1 + D)t = -(t + 1)$$

also

$$y = -R(2)(t + 1) = \frac{1}{2} \left\{ 1 + \frac{D}{2} \right\} (t + 1) = \frac{1}{2} \left(t + \frac{3}{2} \right).$$

y ist in der Tat eine partikuläre Lösung. Im analogen Beispiel

$$y'' - 3y' + 2y = \sin t$$

finden wir

$$y = \frac{1}{10} \{ \sin t + 3 \cos t \}.$$

8.3.2 Spezielle Gleichungen

Separable Gleichungen: Die einfachsten direkt integrierbaren Differentialgleichungen sind wohl die separablen im \mathbb{R}^1 , nämlich

$$x' = f(x)g(t).$$

Ihre Integration erfolgt über

$$\int_{x^0}^x \frac{ds}{f(s)} = \int_{t_0}^t g(s)ds.$$

Beispiel 8.3.5: $y' = e^y \sin t$ mit $y(0) = c$.

Als autonomes System wird daraus mit $x_1 := t$, $x_2 := y$

$$x' = \begin{pmatrix} 1 \\ e^{x_2} \sin x_1 \end{pmatrix}, \quad x(0) = x^0 = \begin{pmatrix} 0 \\ c \end{pmatrix},$$

und man erhält mit $\gamma := e^{-c} - 1$

$$x(t) = \begin{pmatrix} t \\ -\log(\cos t + \gamma) \end{pmatrix}.$$

Ein Phasenbild finden Sie in §8.6, Abbildung 3.

Transformation der Variablen:

Beispiel 8.3.6: $x'(t) = f(ax + bt + c)$ mit $a \neq 0$ und $x(0) = x^0$ im \mathbb{R}^1 .

Es sei $y(t) := ax + bt + c$. Dann ist

$$y' = af(y) + b, \quad y(0) = ax^0.$$

Homogene Differentialgleichung:**Beispiel 8.3.7:** $x'(t) = f(x/t)$ mit $x(1) = x^0$ im \mathbb{R}^1 .Es sei $y(t) := x/t$. Dann ist

$$y'(t) = \frac{1}{t}(f(y) - y), \quad y(1) = x^0.$$

Bernoullische Differentialgleichung:**Beispiel 8.3.8:** $x'(t) = a(t)x + b(t)x^\alpha$ mit $\alpha \neq 0, 1$, $x(t_0) = x^0$ im \mathbb{R}^1 .Es sei $y := x^{1-\alpha}$. Dann ist

$$y'(t) = (1 - \alpha)\{a(t)y + b(t)\}, \quad y(t_0) = (x^0)^{1-\alpha}.$$

Das erste Integral: Das erste Integral (eine Erhaltungsgröße) ist schon in den Beispielen 8.1.2.4 und 8.1.2.5 aufgetreten. Es seien also wieder

$$x' = v(x), \quad x(0) = x^0 \quad (7)$$

und Φ_t der dazugehörige Fluß. Dann nennen wir

$$I: X \rightarrow \mathbb{R} \quad \text{mit} \quad \forall t \quad I(\Phi_t x^0) = I(x^0) \quad (8)$$

erstes Integral von Gl. (7). Solche Erhaltungsgrößen spielen naturgemäß in der Lösungstheorie eine große Rolle. Aus Gl. (8) folgt unmittelbar

$$0 = \nabla I(x) \cdot x' = \nabla I(x) \cdot v(x)$$

 ∇I steht also auf dem Vektorfeld v senkrecht. (∇ ist der Gradient, er wird in §9.4 erklärt.)Es sei nun zur Vereinfachung $X = \mathbb{R}^2$ und

$$w := \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix}.$$

Dann ist $(w, v) = 0$. Um ein I zu finden, müssen wir die Gleichung

$$\nabla I = w$$

nach I auflösen. Dafür gibt es in der Vektoranalysis Integrabilitätsbedingungen. Ein Vektorfeld w heißt *exakt*, wenn es sich als Gradientenfeld darstellen läßt, und die Bedingung dafür lautet

$$dw := -\partial_2 w_1 + \partial_1 w_2 = 0.$$

Für unser w bedeutet das

$$\operatorname{div} v := \partial_1 v_1 + \partial_2 v_2 = 0.$$

Dann erhält man I durch

$$I(x) := I(x^0) + \int_{\gamma(x^0, x)} (w, \tau).$$

Das Integral ist vom Wege $\gamma(x^0, x)$ unabhängig; τ ist der Tangentenvektor von $\gamma(x^0, x)$. Bezüglich der Einzelheiten muß ich auf §12 verweisen.**Beispiel 8.3.9:** Es sei

$$v(x) := \begin{pmatrix} 4x_1 x_2 + 3x_2^2 \\ -3x_1^2 - 2x_2^2 \end{pmatrix}.$$

Dann ist $\operatorname{div} v = 0$. Für γ wählen wir mit $x^0 = 0$ den Streckenzug

$$\gamma = \{(s, t) \mid s \in [0, x_1], t = 0\} \cup \{(s, t) \mid s = x_1, t \in [0, x_2]\}$$

und erhalten

$$\begin{aligned} \int_{\gamma(0, x)} (w, \tau) &= \int_0^{x_1} 3s^2 ds + \int_0^{x_2} (4x_1 s + 3s^2) ds \\ &= (x_1)^3 + 2x_1(x_2)^2 + (x_2)^3 =: I(x). \end{aligned}$$

Offenbar ist $w = \nabla I$, und I ist ein erstes Integral unserer Gleichung (7). Gleichung (7) mit $\operatorname{div} v = 0$ nennt man deshalb auch *exakt*.

Riccatische Differentialgleichung: Unter einer Riccatischen Differentialgleichung versteht man eine Gleichung der Form (im \mathbb{R}^1)

$$x' = f x^2 + g x + h, \quad x(t_0) = x^0$$

mit $f, g, h \in C(\mathbb{R})$. Es handelt sich also um die allgemeine quadratische Gleichung, die naturgemäß nach der linearen von besonderem Interesse ist.

Riccatische Gleichungen sind längst nicht mehr so einfach vollständig zu behandeln wie lineare. Ich erinnere an Beispiel 8.2.5 und gebe noch einige Hinweise.

1. Es sei eine partikuläre Lösung p mit $p(t_0) \neq x^0$ bekannt. Dann kann man die allgemeine Lösung berechnen. Für die Differenz $y := x - p$ zweier Lösungen gilt nämlich

$$y' = f y^2 + (2pf + g)y, \quad y(t_0) = x^0 - p^0.$$

Das ist eine Bernoullische Gleichung. Beispielsweise sei

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

Eine partikuläre Lösung ist $p(t) = -1/t$. Damit folgt

$$y'(t) = y^2 + 2y \left(t - \frac{1}{t} \right),$$

und mit

$$E(t) = \int_0^t e^{s^2} ds$$

berechnet man

$$x(t) = \frac{e^{-t^2} \{x^0 + 2E(t)\}}{1 - t e^{-t^2} \{x^0 + 2E(t)\}}$$

2. Riccatische Gleichungen lassen sich auf lineare Differentialgleichungen zweiter Ordnung zurückführen. Dazu macht man den Ansatz

$$u(t) := \exp \left(- \int_{t_0}^t f(s)x(s) ds \right)$$

und erhält

$$u'' - u' \left(g + \frac{f'}{f} \right) + h f u = 0$$

mit $u(t_0) = 1$ und $u'(t_0) = -f(t_0)x^0$.

3. Man kann versuchen, mit der Picard-Lindelöfschen Iteration oder durch Potenzreihenansatz Lösungen zu erhalten. Beispielsweise folgt aus (vgl. Beispiel 8.2.5)

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

die lineare Gleichung

$$u''(t) + t^2 u(t) = 0, \quad u(0) = 1, \quad u'(0) = 0.$$

Macht man in der Umgebung des Nullpunktes den Potenzreihenansatz

$$u(t) = \sum_{n=0}^{\infty} a_n t^n,$$

dann folgt $a_0 = 1, a_1 = 0$ und

$$a_2 = 0, \quad a_3 = 0, \quad a_{n+2} = -\frac{a_{n-2}}{(n+1)(n+2)}.$$

Man erhält so

$$u(t) = 1 - \frac{t^4}{3 \cdot 4} + \frac{t^8}{3 \cdot 4 \cdot 7 \cdot 8} - \dots = \Gamma(3/4) \cdot \sqrt{t/2} \cdot Z_{-\frac{1}{4}}(t^2/2)$$

mit

$$\Gamma(x) := \int_0^{\infty} e^{-t} t^{x-1} dt.$$

Es ist $\Gamma(1+x) = x\Gamma(x)$, $\Gamma(1+n) = n!$ und $\Gamma(3/4) = 1,22541 \dots$.

8.3.3 Wachstumsmodelle

Es sei $x(t)$ die Population einer Spezies. Dann läßt sich durch

$$x' = f(\cdot, x)x, \quad x(0) = x^0 \quad (9)$$

das Wachstum einer Bevölkerung oder der radioaktive Zerfall beschreiben. f ist beispielsweise die Differenz von Geburten- und Sterberate.

Den einfachsten Spezialfall erhält man durch die Wahl

$$\forall (t, x) \in \mathbb{R} \times \mathbb{R} \quad f(t, x) = a \in \mathbb{R}.$$

Dann ist Gl. (9) linear und wird durch

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

gelöst.

In den meisten Fällen ist die Annahme einer konstanten Wachstumsrate jedoch nicht sehr realistisch. Man geht deshalb von

$$\exists \xi > 0 \quad \forall x \geq \xi \quad f(t, x) \leq 0$$

aus. Eine einfache Situation dieser Art liegt vor, wenn

$$f(t, x) = b(\xi - x)$$

ist mit $b > 0$. Mit diesem f und $a := b\xi$ lautet Gl. (9)

$$x' = a x - b x^2 \quad (10)$$

Dies ist die *Gleichung des beschränkten Wachstums* oder die *logistische Gleichung*. Sie wurde 1837 von dem holländischen Biomathematiker P. VERHULST (1804–1849) aufgestellt. Mit einer einfachen Variablentransformation wird aus Gl. (10)

$$x' = x(1 - x). \quad (11)$$

Diese Gleichung weist schon interessante Phänomene auf. Sie läßt sich explizit integrieren (Separation der Variablen). Die Punkte $x^1 = 0$ und $x^2 = 1$ sind Fixpunkte der Bewegung. x^1 ist abstoßend, x^2 ist anziehend. Man vergleiche das Phasenbild 4 in §8.6.

Räuber-Beute Modelle: Wir betrachten nun noch ein etwas komplizierteres Modell für zwei Populationen $x_1(t)$ und $x_2(t)$. Dabei seien $x_1(t)$ die Beutespezies und $x_2(t)$ die Räuberspezies. Es gelte

$$\begin{aligned} x_1' &= r_1(\cdot, x)x_1 \\ x_2' &= r_2(\cdot, x)x_2. \end{aligned} \quad (12)$$

In dem Modell wurde angenommen, daß die Räuberspezies sich ausschließlich von den Beutespezies ernähren, während für die Beutespezies unbegrenzt Nahrung vorhanden ist.

Als erstes betrachten wir wieder den Fall des unbegrenzten Wachstums. Mit positiven Konstanten a, b, c, d wählt man dann

$$\begin{aligned} r_1(t, x) &= a - b x_2 \\ r_2(t, x) &= -d + c x_1. \end{aligned}$$

Diese Gleichungen sind nach VITO VOLTERRA, 1860–1940, und ALFRED JAMES LOTKA, 1880–1949, benannt. Es gibt wieder zwei Fixpunkte der Bewegung, nämlich $x^1 = 0$ und $x^2 = (d/c, a/b)$. x^1 ist ein Sattelpunkt, um x^2 gibt es periodische Orbits. Man vergleiche das Phasenbild 5 in §8.6.

Beim Räuber-Beutemodell mit begrenztem Wachstum modifiziert man das Volterra-Lotka-System und fügt *soziale Reibungsterme* hinzu, nämlich mit positiven Konstanten m, n

$$\begin{aligned} r_1(t, x) &= a - b x_2 - m x_1 \\ r_2(t, x) &= -d + c x_1 - n x_2. \end{aligned}$$

Jetzt gibt es vier Fixpunkte, nämlich wieder $x^1 = 0$, sowie mit $k := mn + bc$

$$x^2 = \begin{pmatrix} 0 \\ -d/n \end{pmatrix}, \quad x^3 = \begin{pmatrix} a/m \\ 0 \end{pmatrix} \quad \text{und} \quad x^4 = \frac{1}{k} \begin{pmatrix} na + bd \\ ac - md \end{pmatrix}.$$

Man beachte, daß längs der Geraden

$$g_1 : a - b x_2 - m x_1 = 0$$

das Vektorfeld parallel zur x_2 -Achse verläuft ($x_1' = 0$) und daß es längs

$$g_2 : -d + c x_1 - n x_2 = 0$$

parallel zur x_1 -Achse verläuft ($x_2' = 0$). x^4 ist der Schnittpunkt beider Geraden.

Je nachdem x^4 im ersten Quadranten liegt oder nicht, verlaufen die Orbits verschieden. Sie können sich einem Fixpunkt oder einem Grenzzyklus annähern. Die Achsen selbst sind Orbits. In jedem Fall gibt es obere Grenzen für beide Spezies. Man vergleiche die Phasenbilder 6-8 in §8.6 und Beispiel 8.4.12.

8.4 Qualitative Aspekte

In diesem Abschnitt wollen wir qualitative Aspekte bei der Diskussion des Lösungsverhaltens in den Vordergrund stellen und zum Beispiel die Frage nach dem asymptotischen Verhalten der Lösungen aufwerfen. Hier gibt es noch viel zu tun. Denken Sie zum Beispiel an das Mehrkörperproblem in der Himmelsmechanik. Man kann zwar lokal die Lösungen berechnen, aber die Frage, ob eines Tages die Erde mit einem anderen Planeten zusammenstößt, läßt sich noch nicht beantworten. Beim Dreikörperproblem etwa kann man nur Aussagen der folgenden Art machen: Stoßen zwei Körper zusammen, dann hat der dritte einen endlichen Abstand. Oder: Die Zusammenstöße häufen sich im Endlichen nicht.

Leider kann ich in dieser Anfängervorlesung auf solche interessante und moderne Fragen nicht tiefer eingehen. Ich werde also in diesem Abschnitt nur etwas erzählen und ein paar Beispiele bringen.

8.4.1 Maximale Lösungen

Wir gehen wieder von von

$$x' = f(\cdot, x), \quad x(0) = x^0 \quad (1)$$

in $J \times X, X = \mathbb{R}^n$, aus. Das Vektorfeld f sei in $I \times X$ stetig mit $J \subset I$. Dann ist Gl. (1) lokal lösbar. Es sei $x : [0, b) \rightarrow X$ eine solche Lösung. Dann interessiert uns natürlich das Verhalten von $x(t)$ für $t \rightarrow b$. Wenn dieser Limes existiert, dann können wir in $t = b$ den Existenzsatz wieder anwenden und die Lösung über $t = b$ hinaus fortsetzen. Analog geht man für negative t vor.

Definition 8.4.1: $x : (a, b) \subset I \rightarrow X$ heißt maximale Lösung von Gl. (1), wenn x nicht aus einer anderen Lösung durch Einschränkung des Definitionsintervalles gewonnen werden kann.

Es gilt

Satz 8.4.2: Jede lokale Lösung von Gl. (1) läßt sich zu einer maximalen fortsetzen.

Eine maximale Lösung

$$x : (a, b) \rightarrow X$$

wird also für $t \downarrow a$ bzw. $t \uparrow b$ singular. Weiß man, daß f beschränkt ist, dann existieren $x(a)$ und $x(b)$ jedoch, genauer

Satz 8.4.3: f sei in $(a, b) \times X$ beschränkt. Dann existieren

$$\lim_{t \downarrow a} x(t) \quad \text{und} \quad \lim_{t \uparrow b} x(t).$$

f sei in $[a, b] \times X$ beschränkt und stetig. Dann existieren die rechts- bzw. linksseitigen Ableitungen $x'_+(a)$ und $x'_-(b)$.

Mit $t_1 < \tau < t_2$ ist nämlich

$$|x(t_2) - x(t_1)| = (t_2 - t_1)|x'(\tau)| \leq c|t_2 - t_1|.$$

Mithin konvergiert $x(t_n)$ für $t_n \downarrow a$. Dann ist auch

$$\frac{x(a+h) - x(a)}{h} = \frac{1}{h} \int_a^{a+h} f(s, x(s)) ds,$$

und der Limes existiert für $h \downarrow 0$.

Eine Lösung kann also bei beschränktem f am Rande nicht oszillieren. In $(0, 1/\pi)$ ist zum Beispiel

$$x(t) = \begin{pmatrix} t^2 \sin \frac{1}{t} \\ 2t \sin \frac{1}{t} - \cos \frac{1}{t} \end{pmatrix}$$

beschränkt und löst die Differentialgleichung

$$x'(t) = \begin{pmatrix} x^2 \\ -(\frac{2}{t^2} + \frac{1}{t})x_1 + \frac{2}{t}x_2 \end{pmatrix}, \quad x(1/\pi) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Genügt f nun zusätzlich beispielsweise einer Lipschitzbedingung, dann sind die erhaltenen Lösungen eindeutig. Man definiert

Definition 8.4.4: $(\tau, \xi) \in \mathbb{R} \times X$ heißt globaler Eindeigkeitspunkt für $x' = f(\cdot, x)$, wenn durch (τ, ξ) genau eine maximale Lösung verläuft. (τ, ξ) heißt lokaler Eindeigkeitspunkt, wenn es eine Umgebung $U(\tau, \xi)$ gibt, in der (τ, ξ) globaler Eindeigkeitspunkt ist.

Es folgt

Satz 8.4.5: Jeder Punkt $(\tau, \xi) \in G \subset \mathbb{R} \times X$ sei lokaler Eindeigkeitspunkt für $x' = f(\cdot, x)$. Dann sind die (τ, ξ) sogar globale Eindeigkeitspunkte.

Es sei nun x eine maximale Lösung von $x' = f(\cdot, x)$ mit $\mathcal{D}(x) = (a, b)$. Dann nennen wir $\xi \in X$ Grenzpunkt von x bzgl. $t \rightarrow b$, wenn es eine Folge $t_n \uparrow b$ gibt mit

$$\xi = \lim_{t_n \uparrow b} x(t_n).$$

Definition 8.4.6: Die Grenzpunktmenge (ω -Limesmenge) M_b von x bzgl. b ist die Menge aller Grenzpunkte von x bzgl. $t \uparrow b$. Entsprechend heißt M_a auch α -Limesmenge.

Es gilt

Satz 8.4.7: f sei in $G \subset \mathbb{R} \times X$ stetig und entweder beschränkt oder jeder Punkt von G sei lokaler Eindeigkeitspunkt. x sei maximale Lösung mit $\mathcal{D}(x) = (a, b)$. Dann gilt mit $a < \alpha < b$:

1. Ist $M_b = \emptyset$, dann gilt $|x(t)| \rightarrow \infty$ für $t \rightarrow b$.
2. Es sei $M_b \neq \emptyset$. Dann ist M_b abgeschlossen.
 M_b ist beschränkt $\iff x$ ist in $[\alpha, b)$ beschränkt.
Ist b endlich, dann ist $\{b\} \times M_b \subset \partial G$.

8.4.2 Stabilität

Wir wollen nun wieder Gl. (1) betrachten und annehmen, daß sie eindeutig lösbar ist und daß eine Lösung x im ganzen \mathbb{R} definiert ist. Dementsprechend bezeichnen wir mit M_+ und M_- die Grenzpunktmenge bzgl. $t \rightarrow \infty$ und $t \rightarrow -\infty$. Es interessieren nun folgende Fragen:

1. Gibt es Fixpunkte x^0 , für die also $x \equiv x_0$ Lösungen sind?
2. Es sei y eine Lösung (zum Beispiel ein Fixpunkt) und x eine zweite Lösung, die zur Zeit $t = t_0$ nahe bei y liegt. Bleibt x dann für alle $t \geq t_0$ nahe bei y ?
3. Es sei x eine Lösung. Was sind M_+ und M_- ?

Die erste Frage läßt sich leicht beantworten. Fixpunkte sind gerade die singulären Punkte des Vektorfeldes, also $f(\cdot, x^0) \equiv 0$. Bei der Behandlung der beiden anderen Fragen spielt der Begriff der Stabilität eine große Rolle. Er geht auf A. LYAPUNOV (1857–1918) zurück. Es gibt verschiedene Konzepte dafür. Zur Vereinfachung wählen wir wieder $t_0 = 0$ und für y einen Fixpunkt, o.B.d.A. $y = 0$. Dann verwenden wir für $t \rightarrow \infty$ (und entsprechend für $t \rightarrow -\infty$)

Definition 8.4.8: y_0 heißt stabil \iff

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, |x(0) - y_0| < \delta \quad \forall t \geq 0 \quad |x(t) - y_0| < \varepsilon.$$

Definition 8.4.9: y_0 heißt asymptotisch stabil $\iff y$ ist stabil und

$$\lim_{t \rightarrow \infty} |x(t) - y_0| = 0.$$

Mit instabil bezeichnet man nicht das logische Gegenteil von stabil sondern

Definition 8.4.10: y_0 heißt instabil \iff

$$\forall \varepsilon > 0 \quad \forall x, 0 < |x(0) - y_0| < \varepsilon \quad \exists t > 0 \quad |x(t) - y_0| > \varepsilon.$$

Die andere Möglichkeit nennt man *labil*.

Die Frage nach der Stabilität der Lösungen läßt sich bei den linearen Systemen

$$x' = Ax \tag{2}$$

durch Ausrechnen explizit beantworten. Man vergleiche §8.3.1. Ist der Realteil aller Eigenwerte nicht positiv (Knoten, Strudel oder Zentrum), dann ist $y = 0$ für $t \rightarrow \infty$ stabil. Sind diese Realteile sogar negativ (Knoten oder Strudel), dann ist $y = 0$ asymptotisch stabil. Sind sie positiv, dann ist $y = 0$ instabil. Ein Sattelpunkt ist labil.

Interessant ist nun die Behandlung gestörter linearer Systeme, nämlich

$$x' = Ax + g(\cdot, x) \tag{3}$$

mit $\det A \neq 0$, $g \in C_1(\mathbb{R} \times \mathcal{X}, \mathcal{X})$, $\mathcal{X} = \mathbb{R}^n$ und

$$g(t, x) = o(|x|) \tag{4}$$

für $|x| \rightarrow 0$, und zwar gleichmäßig bzgl. t . Man kann zeigen, daß sich die Lösungen von Gl. (3) in der Umgebung des Nullpunktes wie die Lösungen des linearen Systems (2) verhalten. Jedoch kann aus einem Zentrum ein Strudel werden. Ich beweise hier nur

Satz 8.4.11: Die Realteile der Eigenwerte von A seien negativ. Dann ist $y = 0$ bzgl. Gl. (3) und $t \rightarrow \infty$ asymptotisch stabil.

Zum Beweis verwenden wir

$$\exists a > 1, b > 0 \quad \forall t \geq 0 \quad \|e^{At}\| \leq a e^{-bt}.$$

Dabei ist

$$e^{At} := \sum_{n=0}^{\infty} \frac{1}{n!} (At)^n$$

der Fluß zur linearen Gleichung (2), und im Falle diagonalähnlicher Matrizen folgt die Abschätzung leicht aus

$$T^{-1} e^{At} T = e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & \dots & 0 \\ \vdots & \diagdown & \vdots \\ 0 & \dots & e^{\lambda_n t} \end{pmatrix}.$$

Die Lösung x mit $x(0) = x^0$ von Gl. (3) hat die Form

$$x(t) = e^{At} x^0 + \int_0^t e^{A(t-\tau)} g(\tau, x(\tau)) d\tau.$$

Daraus folgt

$$|x(t)| \leq a \left\{ e^{-bt} |x^0| + \int_0^t e^{-b(t-\tau)} |g(\tau, x(\tau))| d\tau \right\}.$$

Nun gilt wegen Gl. (4)

$$\forall \varepsilon \in (0, b) \quad \exists \delta \in (0, \varepsilon) \quad \forall x, |x| < \delta \quad \forall t \geq 0 \quad |g(t, x)| < \frac{\varepsilon}{a} |x|.$$

Wir behaupten nun

$$\forall t \geq 0 \quad \forall x^0, |x^0| < \frac{\delta}{a} \quad |x(t)| < \delta < \varepsilon. \tag{5}$$

Es sei nämlich $x^0 \in \mathcal{X}$ mit $|x^0| < \delta/a$ und

$$t^0 := \inf \{ t \in \mathbb{R}_0^+ \mid |x(t)| = \delta \}.$$

Dann würde für $0 \leq t \leq t^0$

$$|x(t)| \leq \delta e^{-bt} + \varepsilon \int_0^t e^{-b(t-\tau)} |x(\tau)| d\tau$$

folgen, und aus dem Gronwallschen Lemma würden wir

$$|x(t)| \leq \delta e^{-bt} \exp\left(\varepsilon \int_0^t e^{-b(t-\tau)} d\tau\right) \leq \delta e^{-bt} e^{\varepsilon t} \quad (6)$$

erhalten, insbesondere also

$$\delta = |x(t^0)| \leq \delta e^{-(b-\varepsilon)t^0} < \delta.$$

Damit ist Abschätzung (5), also die Stabilität, bewiesen. Abschätzung (6) gilt dann aber für alle $t \geq 0$. Mithin ist $y = 0$ asymptotisch stabil. Das war zu zeigen.

Beispiel 8.4.12: Das Räuber-Beute Modell in §8.3.3.

In diesem Falle ist

$$f(x) = \begin{pmatrix} x_1(a - bx_2 - mx_1) \\ x_2(-d + cx_1 - nx_2) \end{pmatrix},$$

also mit $f(x^i) = 0$, $i = 1, 2, 3, 4$,

$$f(x) = \partial f(x^i)(x - x^i) + \dots$$

mit

$$\partial f(x) = \begin{pmatrix} a - 2mx_1 - bx_2 & -bx_1 \\ cx_2 & -d + cx_1 - 2nx_2 \end{pmatrix}.$$

Mithin gilt für die Linearisierung:

- $x^1 = 0$ ist ein Sattelpunkt (labil),
- x^2 ist instabil,
- x^3 ist für $ac < md$ asymptotisch stabil, sonst labil,
- x^4 ist für $ac > md$ asymptotisch stabil.

Unsere Abbildungen bestätigen dieses Verhalten.

Beispiel 8.4.13: Das ungedämpfte Pendel, Beispiel 8.1.2.5.

Nun ist

$$\partial f(x) = \begin{pmatrix} 0 & 1 \\ -\cos x_1 & 0 \end{pmatrix}.$$

Mithin gilt für die Linearisierung $\lambda^2 = -\cos x_1$ und

$$x^{2n} = \begin{pmatrix} 2n\pi \\ 0 \end{pmatrix} \text{ sind Zentren (stabil)}$$

$$x^{2n+1} = \begin{pmatrix} (2n+1)\pi \\ 0 \end{pmatrix} \text{ sind Sattelpunkte (instabil).}$$

Als letztes möchte ich noch den Begriff der Lyapunov-Funktion einführen. Zur Motivation erinnere ich an die Gleichung für gedämpfte Schwingungen, nämlich mit $r \geq 0$

$$x' = v(x) := \begin{pmatrix} x_2 \\ -x_1 - rx_2 \end{pmatrix}.$$

Es ist $\operatorname{div} v = -r$. Für $r > 0$ existiert also kein erstes Integral, die Summe von potentieller und kinetischer Energie ist ja nicht mehr konstant. Im Falle $r = 0$ ist

$$E(x) = |x|^2$$

die Energie, und es gilt $E' = (\nabla E) \cdot x' = 2x \cdot v(x) = 0$. Im Falle $r > 0$ haben wir aber noch

$$E' = (\nabla E)x' = 2x \cdot v(x) = -2r(x_2)^2 \leq 0.$$

Wenn man sich auf einem Orbit bewegt, nimmt die Energie wegen der Dämpfung ab. Im Fixpunkt $x^0 = 0$ verschwindet sie.

Obwohl E im Falle gedämpfter Schwingungen nicht mehr konstant ist, kann man wegen der Monotonie aus der Kenntnis von E wichtige Schlüsse ziehen. Deshalb definiert man

Definition 8.4.14: Es sei x^0 ein isolierter singulärer Punkt des Vektorfeldes $v(x)$. Dann heißt $L \in C_1(X, \mathbb{R}_0^+)$ Lyapunov-Funktion zu v, x^0 : \iff L besitzt in einer Umgebung von x^0 folgende Eigenschaften:

1. $L(x) = 0$ nur für $x = x^0$.
2. $(\nabla L)v \leq 0$.

Gilt statt 2 sogar

- 2'. $(\nabla L)v < 0$ für $x \neq x^0$,

dann heißt L strenge Lyapunov-Funktion.

Es gilt das

Stabilitätskriterium: $x^0 = 0$ sei isolierter singulärer Punkt von v , und es existiere eine Lyapunov-Funktion. Dann ist x^0 bzgl. $x' = v(x)$ stabil. Existiert sogar eine strenge Lyapunov-Funktion, dann ist x^0 asymptotisch stabil.

Den Beweis führt man in zwei Schritten:

1. L existiere in $B(x^0, \varepsilon_0)$, es sei $\varepsilon > 0$ gegeben, $\varepsilon < \varepsilon_0$ und

$$m(\varepsilon) := \min_{|x|=\varepsilon} L(x) > 0.$$

Wegen $L(x^0) = 0$ und der Stetigkeit von L gibt es ein $\delta(\varepsilon)$, etwa $\delta < \varepsilon$ mit

$$\forall x, |x| < \delta \quad L(x) < m(\varepsilon).$$

Es sei nun x ein Orbit und $t_1 \in \mathbb{R}$ mit $|x(t_1)| < \delta$. Es sei ferner $E := L(x)$. Dann ist

$$E(t_1) < m(\varepsilon)$$

und

$$E'(t) = (\nabla L)v \Big|_{x=x(t)} \leq 0.$$

Daraus folgt

$$\forall t \geq t_1 \quad E(t) < m(\varepsilon).$$

Der Orbit kann daher die Kreislinie $|x| = \varepsilon$ nicht erreichen, denn dort wäre $E \geq m$. Mithin ist x^0 stabil.

2. Es sei L sogar eine strenge Lyapunov-Funktion. Dann fällt $E(t)$ streng monoton. Es ist also

$$E(t) \geq \lambda := \lim_{t \rightarrow \infty} E(t) \geq 0,$$

und wir müssen nur noch $\lambda = 0$ zeigen.

Dazu nehmen wir $\lambda > 0$ an. Dann existiert wegen $L(x^0) = 0$ ein ρ mit $0 < \rho < \delta < \varepsilon$, so daß

$$\forall x, |x| < \rho \quad L(x) < \lambda$$

ist. Ferner gilt nach Voraussetzung

$$\forall x, \rho \leq |x| \leq \varepsilon \quad ((\nabla L)v)(x) \leq M < 0.$$

Für $t \geq t_1$ ist $|x(t)| < \varepsilon$ (erster Beweisschritt) und wegen $E(t) \geq \lambda$ auch $|x(t)| \geq \rho$, also

$$E(t) = E(t_1) + \int_{t_1}^t E'(\tau) d\tau \leq E(t_1) + M(t - t_1).$$

Mithin würde $E(t) \rightarrow -\infty$ folgen im Widerspruch zu $E \geq 0$. Also gilt $\lambda = 0$ und damit die asymptotische Stabilität.

Es ist nicht immer leicht, zu einem vorgelegten System eine Lyapunov-Funktion zu finden. Beim *Gradientensystem*

$$x' = v(x) := -(\nabla U)(x)$$

mit $U \in C_1(X, \mathbb{R})$ ist

$$(\nabla U)v = -(\nabla U)^2 \leq 0.$$

Das *Potential* U ist also ein natürlicher Kandidat für eine Lyapunov-Funktion.

Betrachten wir noch einmal ein gestörtes System in Normalform, etwa

$$x' = v(x) := -x + a(x) \quad \text{mit} \quad a(x) = O(|x|^2).$$

Wir versuchen L als quadratisches Energiefunktional anzusetzen, in diesem einfachen Falle als

$$L(x) = |x|^2.$$

Dann ist

$$((\nabla L)v)(x) = -2|x|^2 + \mathcal{O}(|x|^3).$$

In einer Umgebung des Nullpunktes ist L mithin eine strenge Lyapunov-Funktion, und x^0 ist asymptotisch stabil.

Nichtlineare Terme können das Verhalten der Lösungen natürlich empfindlich stören. Betrachten wir etwa im \mathbb{R}^2 die Gleichung

$$x' = v(x) := \begin{pmatrix} x_2 \\ \varepsilon x_1^3 \end{pmatrix}$$

mit $x(0) = x^0$ und $\varepsilon \in \{-1, 0, 1\}$. Dann ist

$$I(x) := 2(x_2)^2 - \varepsilon(x_1)^4 = c$$

ein erstes Integral. Für $\varepsilon = 0$ ist x_2 konstant und x_1 eine lineare Funktion von t . Für $\varepsilon = -1$ existiert $x(t)$ für jedes x^0 und ist für alle t beschränkt. Für $\varepsilon = 1$ erhält man aus $x_1' = x_2 = \sqrt{c + (x_1)^4} / \sqrt{2}$

$$t = \sqrt{2} \int_{x_1(0)}^{x_1(t)} \frac{ds}{\sqrt{c + s^4}}.$$

Wenn c und $x_1(0)$ positiv sind, wächst deshalb $x_1(t)$ und damit auch $x_2(t)$ für

$$t \rightarrow \sqrt{2} \int_{x_1(0)}^{\infty} \frac{ds}{\sqrt{c + s^4}}$$

unbeschränkt.

8.4.3 Normalformen

Bei der Diskussion von Abbildungen spielt das Zurückführen auf Normalfunktionen eine große Rolle. Denken Sie an die Diagonalgestalt oder die Jordansche Normalform bei Matrizen und ihre Anwendung bei den linearen Differentialgleichungssystemen.

Eine solche Typeneinteilung möchte man auch bei den Differentialgleichungen finden. Betrachten wir wieder

$$x' = v(x).$$

Hier spielen die Fixpunkte eine große Rolle. Das haben wir ja bereits an den Beispielen bemerkt. Ist x^0 kein Fixpunkt, gilt also in einer Umgebung $U(x^0)$

$$v(x) \neq 0,$$

dann folgt aus unseren Existenzsätzen, daß man v in dieser Umgebung *geradebiegen* kann. Das heißt, es gibt eine bijektive stetig differenzierbare Abbildung (einen Diffeomorphismus), die $x' = v(x)$ in $y' = e^1$ überführt.

Betrachten wir nun wieder ein lineares System im \mathbb{R}^2 in einer Umgebung eines singulären Punktes. Dann ordnet man den singulären Punkten Zahlen zu, um sie dadurch zu charakterisieren. Ich kann das jetzt nur kurz skizzieren, weil wir Wegintegrale noch nicht besprochen haben. Wir wollen den *Index eines Vektorfeldes* verwenden.

Es sei γ eine geschlossene Jordankurve (eine bijektive stetige Abbildung der Kreislinie) und $v \in C_1$ ein Vektorfeld mit $v \neq 0$ auf γ . Dann definiert man

$$i(v, \gamma) := \frac{1}{2\pi} \int_{\gamma} d\varphi \quad \text{mit} \quad \varphi := \arctan \frac{v_2}{v_1}.$$

Offenbar ist i eine ganze Zahl, und es gilt

$$i(v, \gamma) = 0,$$

wenn v auch im Inneren von γ nicht verschwindet. Denn i hängt dann stetig von γ ab und verschwindet für kleine γ .

Es sei nun x^0 ein isolierter singulärer Punkt von v . γ_1 und γ_2 seien geschlossene Jordankurven; x^0 liege im Inneren von γ_1 und γ_2 . Dann gilt auch

$$i(v, \gamma_1) = i(v, \gamma_2).$$

Der Index hängt dann also auch nicht von der speziellen Wahl der Jordankurve ab.

Auf diese Weise kann man jedem isolierten singulären Punkt x^0 von v eine ganze Zahl zuordnen, nämlich den *Index*

$$\text{ind}(v, x^0) := i(v, \gamma).$$

Dabei ist γ irgendeine geschlossene Jordankurve, die x^0 im Inneren enthält, aber keine anderen singulären Punkte von v .

Durch die Berechnung der Indizes lassen sich wichtige Schlüsse ziehen. Auf diese Weise kann man beispielsweise den Fundamentalsatz der Algebra beweisen. Es seien $z = x_1 + i x_2$, $P(z) = z^n + \dots + a_0$ und mit $x := (x_1, x_2)$

$$v(x) := \begin{pmatrix} \text{Re } P(z) \\ \text{Im } P(z) \end{pmatrix}.$$

Wäre überall $v \neq 0$, dann würde der Index verschwinden. Er läßt sich aber für einen Kreis mit großem Radius leicht berechnen, es ist nämlich $i(v, B(r, 0)) = n$ für große r .

Man kann nun die Indizes für die in §8.3.1.1 angegebenen Normalformen ausrechnen und erhält für $v(x) = Ax$ mit $\det A \neq 0$

$$\text{ind}(v, 0) = \text{sign}(\det A),$$

also $\text{ind}(v, x^0) = -1$ für Sattelpunkte und $= +1$ sonst. Man zeigt auch, daß sich der Index bei Störungen der Form (3), (4) nicht ändert. $\text{ind}(v, x^0)$ ist der *lokale Index* und läßt sich auch im \mathbb{R}^n erklären, nämlich wieder als $\text{sign}(\det A)$. Im \mathbb{R}^3 haben Sattelpunkte deshalb den Index -1 , aber auch Senken (alle Eigenwerte sind negativ). Hingegen hat eine Quelle (alle Eigenwerte sind positiv) den Index $+1$.

8.4.4 Periodische Lösungen im \mathbb{R}^2

In den Anwendungen ist es oft wichtig, geschlossene Orbits zu finden. Ich muß mich jetzt wirklich kurz fassen und teile nur einige Ergebnisse mit:

Umlaufsatz: *Ein periodischer Orbit muß singuläre Punkte umlaufen, deren Indexsumme $+1$ ist.*

Satz von Poincaré-Bendixson: *Es seien K ein kompaktes Gebiet in der Ebene und x Lösung von $x' = v(x)$ mit $\forall t \ x(t) \in K$. K enthalte keine singulären Punkte von v . Dann ist M_+ ein periodischer Orbit, und $x(t)$ strebt spiralförmig dagegen.*

Beispiel 8.4.15: *Wir zeigen, daß*

$$y'' + (y^2 + 2(y')^2 - 1)y' + y = 0$$

eine nichttriviale periodische Lösung besitzt.

Das entsprechende System lautet

$$x' = v(x) := \begin{pmatrix} x_2 \\ -x_1 + (1 - (x_1)^2 - 2(x_2)^2)x_2 \end{pmatrix}.$$

Nur der Nullpunkt ist singulär, er hat den Index 1. Es sind also periodische Orbits um den Nullpunkt herum möglich. Nun ist

$$\left(\frac{|x|^2}{2}\right)' = x_1 x_1' + x_2 x_2' = (1 - (x_1)^2 - 2(x_2)^2)(x_2)^2$$

und

$$1 - (x_1)^2 - 2(x_2)^2 \text{ ist } \begin{cases} \text{positiv für } 2|x|^2 < 1 \\ \text{negativ für } |x|^2 > 1. \end{cases}$$

Daher folgt für jede Lösung, die mit

$$\sqrt{1/2} < |x(t_0)| < 1$$

startet,

$$\forall t \quad \sqrt{1/2} \leq |x(t)| \leq 1.$$

Aufgrund des Satzes von Poincaré-Bendixson gibt es deshalb in $\sqrt{1/2} \leq |x| \leq 1$ einen periodischen Orbit. Man vergleiche das Phasenbild 9 in §8.6.

Beispiel 8.4.16: Es sei

$$x' = v(x) := \begin{pmatrix} x_2 \\ -x_1 - x_1^3 \end{pmatrix}.$$

Dann ist der Nullpunkt die einzige singuläre Stelle von v mit dem Index 1. Mithin sind periodische Lösungen um den Nullpunkt herum möglich. Weil das Vektorfeld divergenzfrei ist, existiert auch ein erstes Integral, und man findet leicht

$$I(x) = 2|x|^2 + (x_1)^4 = c^2.$$

Es sei $x_c(0) = x_c^0$ mit $I(x_c^0) = c$. Dann folgt sofort $|x_c| \leq c/\sqrt{2}$. $x_c(t)$ ist aber auch von unten beschränkt. Es sei nämlich $|x(t)| \leq 1$. Dann ist

$$|x(t)|^2 \geq \frac{1}{3}(3(x_2)^2 + 2(x_1)^2 + (x_1)^4)(t) \geq \frac{c^2}{3},$$

also für alle t

$$\min\left(1, \frac{c}{\sqrt{3}}\right) \leq |x_c(t)| \leq \frac{c}{\sqrt{2}}.$$

x_c ist deshalb periodisch ($x = M_+$), oder $x(t)$ nähert sich spiralförmig einem periodischen Orbit. Wegen $x_1 = 0 \implies x_2 = \pm c/\sqrt{2}$ scheidet letzteres aus. Jeder Orbit ist also periodisch. Man vergleiche das Phasenbild 10 in §8.6. Wir berechnen noch die Periode $T(c)$. Wegen

$$x_2 = 0 \implies (x_1)^2 = -1 \pm \sqrt{1+c^2}$$

folgt mit $a(c) := \sqrt{\sqrt{1+c^2}-1}$

$$T(c) = 4\sqrt{2} \int_0^{a(c)} \frac{ds}{\sqrt{c^2 - 2s^2 - s^4}}.$$

8.4.5 Verzweigungen

Als letztes möchte ich noch kurz Verzweigungsprobleme anreißen. Solche Probleme treten in den Anwendungen oft auf und sind sehr wichtig.

Betrachten wir die parameterabhängige Gleichung

$$x'(\lambda, t) = f(\lambda, t, x). \quad (7)$$

Dabei seien $\Lambda \ni \lambda$ die Parametermenge, $T \in C_1(\Lambda, \mathbb{R})$ die Periode und $f \in C_1(\Lambda \times \mathbb{R} \times \mathcal{X}, \mathcal{X})$. Nehmen wir an, Gl. (7) besäße für ein $\lambda_0 \in \Lambda$ eine $T(\lambda_0)$ -periodische Lösung $x(\lambda_0, t)$. Dann stellt sich die Frage, ob Gl. (7) für λ nahe λ_0 auch Frage, ob Gl. (7) für λ nahe λ_0 auch $T(\lambda)$ -periodische Lösungen $x(\lambda, t)$ nahe $x(\lambda_0, t)$ besitzt.

Diese Frage ist im allgemeinen zu verneinen. Denken Sie nur an das gedämpfte mathematische Pendel

$$x' = \begin{pmatrix} x_2 \\ -\sin x_1 - r x_2 \end{pmatrix}.$$

Dabei ist $r \geq 0$ der Dämpfungsparameter. Im ungedämpften Fall gibt es in einer Umgebung $U(0)$ des Ursprungs durch jeden Punkt genau einen periodischen Orbit. Es sei

$$L(x) := \frac{1}{2}(x_2)^2 + (1 - \cos x_1)$$

die Energie im ungedämpften Fall. Dann ist L für $r > 0$ eine Lyapunov-Funktion. Es ist nämlich

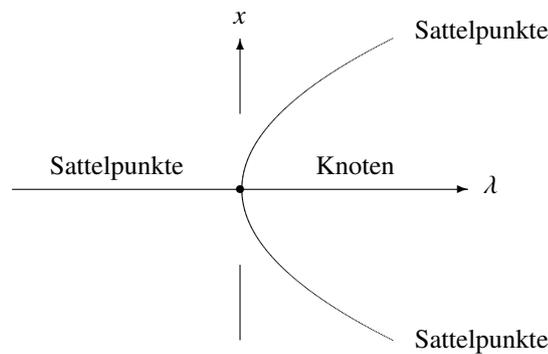
$$(\nabla L \cdot v)(x) = -r(x_2)^2.$$

Mithin ist der Nullpunkt im Falle $r > 0$ stabil, und aus Satz 8.4.11 folgt sogar $x(t) \rightarrow 0$.

Auf weitere Einzelheiten kann ich aber jetzt wirklich nicht mehr eingehen. Ich schließe mit folgendem

Beispiel 8.4.17:

$$x' = v(\lambda, x) := \begin{pmatrix} \lambda x_1 - (x_1)^3 \\ x_2 \end{pmatrix}.$$



Der Punkt $x^0 = 0$ ist dann für jedes λ singulärer Punkt von v . Es sei $\lambda_0 = 0$. Wegen

$$(\partial_x v)(\lambda, 0) = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$$

wechselt $\det(\partial_x v)$ ihr Vorzeichen für $\lambda \in U(0)$. Man kann zeigen, daß dann $(0, 0)$ Verzweigungspunkt für singuläre Punkte von v ist. Oder anders ausgedrückt, $(0, 0)$ ist Bifurkationspunkt des von $x' = v$ erzeugten Flusses.

Das kann man in diesem Beispiel durch Nachrechnen bestätigen. Für negative λ ist nur $x^0(\lambda) = 0$ singulärer Punkt von v mit dem Index -1 . Das ändert sich bei $\lambda = 0$. Für $\lambda > 0$ bleibt x^0 singulär, jetzt aber mit dem Index 1 . Es spalten sich jedoch zwei weitere singuläre Punkte x^1, x^2 ab, beide mit dem Index -1 , nämlich

$$x^1, x^2(\lambda) = \begin{pmatrix} \pm \sqrt{\lambda} \\ 0 \end{pmatrix}.$$

Wir erhalten also folgendes Bifurkationsdiagramm

Für $\lambda > 0$ sind $x^0(\lambda)$ Senken und $x^1(\lambda)$ sowie $x^2(\lambda)$ Sattelpunkte. Man vergleiche die Phasenbilder 11 und 12 in §8.6. Man beachte auch, daß die vorgelegte Gleichung ein Gradientensystem ist. Es gilt nämlich $v = \nabla U$ mit

$$U(\lambda, x) = \frac{1}{4}(2\lambda(x_1)^2 - (x_1)^4 + 2(x_2)^2).$$

Für $\lambda > 0$ ist $x^0(\lambda) = 0$ isolierte Nullstelle von $U(\lambda, x)$, und U ist in einer Umgebung von x^0 Lyapunov-Funktion. Für $\lambda < 0$ gilt das nicht. Es folgt, daß $x^0(\lambda) = 0$ für $\lambda > 0$ instabil ist (eine Quelle).

8.5 Rand- und Eigenwertaufgaben

Bisher haben wir uns mit Anfangswertaufgaben beschäftigt. Es waren eine Differentialgleichung und Anfangswerte, etwa zur Zeit $t = 0$, gegeben. Gesucht war die Lösung. Es gibt aber noch eine zweite, völlig andere Problemstellung, die sowohl aus mathematischer Sicht als auch vom Anwendungsbezug her besonders wichtig und interessant ist. Es handelt sich um Randwertaufgaben und damit zusammenhängend um Eigenwertaufgaben und Schwingungsprobleme.

Ein typisches Beispiel ist die Beschreibung der Auslenkung einer Saite. Es seien ein Intervall $I := (a, b)$ auf der x -Achse und eine Funktion $f \in C(\bar{I}, \mathbb{R})$ gegeben. Dann möchte man etwa ein $u \in C_2(\bar{I}, \mathbb{R})$ bestimmen mit

1. $u'' + u = f$
2. $u(a) = u(b) = 0$.

Das ist eine *Randwertaufgabe*, denn u wird am Rande ∂I vorgegeben (etwa die Saite dort fest eingespannt). Gesucht wird u für alle $x \in I$.

Die Randwertaufgaben hängen eng mit Schwingungsproblemen zusammen. Denken Sie wieder an die Saite. Ihre Auslenkung wird durch die Wellengleichung $\partial_t^2 u = \partial_x^2 u$ beschrieben. Man fragt dann zunächst nach *stehenden Wellen* $a(t) \cdot v(x)$ mit der Amplitude a und der Schwingung v ; man macht also einen Separationsansatz und erhält für v die *Helmholtzsche Schwingungsgleichung*

$$v'' + \lambda v = 0 \quad \text{mit } v(a) = v(b) = 0.$$

Diese Gleichung hat nur für bestimmte *Eigenwerte* $\lambda \in \mathbb{C}$ nichttriviale Lösungen, die *Eigenfunktionen*, und man möchte diese Eigenwerte und damit die Eigenfrequenzen der Saite bestimmen.

Solche Probleme haben zu sehr interessanten Entwicklungssätzen für Differential- und allgemeinere Operatoren geführt. Wie in der linearen Algebra läßt sich eine Normalform für den zugrundeliegenden Differentialoperator angeben, analog zur Transformation auf Diagonalfarm bei symmetrischen Matrizen. Im folgenden möchte ich

solche Fragen für eine Klasse gewöhnlicher Differentialgleichungen zweiter Ordnung behandeln. Die meisten Resultate lassen sich wesentlich verallgemeinern, und zwar sowohl bezüglich der Ordnung des Differentialoperators als auch bezüglich der Dimension des zugrundeliegenden Raumes. Die in §7.5,6 erhaltenen Entwicklungssätze (Fourierreihen) werden dabei als Spezialfälle wiederkehren.

8.5.1 Sturm-Liouvillesche Randwertaufgaben

Im folgenden wollen wir eine Klasse von Differentialgleichungen untersuchen, die nach CH. STURM (1803–1855) und J. LIOUVILLE (1809–1882) benannt sind, nämlich in $I := (a, b)$

$$-(pu')' + qu = \lambda ru + rf. \quad (1)$$

Dabei seien $u \in C_2(J, \mathbb{C})$, $\lambda \in \mathbb{C}$, $p \in C_1(J, \mathbb{R})$, $p > 0$, $f, q, r \in C(J, \mathbb{R})$ und $r > 0$ mit $J := [a, b]$.

Zur Differentialgleichung (1) kommt eine Randbedingung hinzu. Es gibt viele mögliche Randbedingungen, ich nenne hier nur

1. $u(a) = u(b) = 0$, die Dirichletsche Randbedingung.
2. $u'(a) = u'(b) = 0$, die Neumannsche Randbedingung.
3. $(u' + hu)(a) = (u' + hu)(b) = 0$.
4. Man kann zum Beispiel auch $p(a) = 0$ zulassen. Dann genügt $u \in C(J, \mathbb{C})$ bereits als Randbedingung; u soll also in a endlich sein.

Benannt sind diese Randbedingungen nach P. LEJEUNE-DIRICHLET (1805–1859) und C. NEUMANN (1832–1925). Auf die feineren Unterschiede bei den einzelnen Randwertaufgaben möchte ich hier nicht weiter eingehen, sondern nur die wesentlichen Ideen herausstellen. Ich beschränke mich daher im folgenden auf die Dirichletsche Randwertaufgabe.

Als zugrundeliegende Funktionenmenge wählen wir deshalb

$$\mathcal{D} := \{u \in C_2(J, \mathbb{C}) \mid u(a) = u(b) = 0\}.$$

Wir suchen also nach Lösungen $u \in \mathcal{D}$ von Gl. (1). Es sei aber noch einmal ausdrücklich betont, daß es sich hierbei um eine spezielle Randwertaufgabe handelt. Auch Differentialgleichungen erster oder höherer Ordnung und unbeschränkte Gebiete lassen sich behandeln.

Die folgenden Überlegungen sind wie die in §7.5,6 dadurch gekennzeichnet, daß wir in $C(J)$ zwei Normen nebeneinander verwenden: einmal die übliche Supremumsnorm

$$\|u\| := \sup_{x \in J} |u(x)|,$$

und zum anderen die gewichtete $\mathcal{L}^2(I)$ -Norm

$$\|u\|_{\mathcal{L}^2} := \sqrt{\int_I r(x) |u(x)|^2 dx}$$

mit dem Skalarprodukt

$$(u, v) := \int_I r(x) u(x) \overline{v(x)} dx.$$

Gerade dieses Skalarprodukt möchten wir benutzen. Natürlich ist $C(J)$ unter der \mathcal{L}^2 -Norm nicht vollständig, es gilt nur

$$\|f\|_{\mathcal{L}^2} \leq c \|f\| \quad \text{mit} \quad c := \sqrt{(b-a)} \|r\|.$$

Wir definieren nun

$$L : \mathcal{D} \longrightarrow \mathcal{L}^2(I, \mathbb{C}), \\ u \longmapsto Lu$$

mit

$$Lu := \frac{1}{r} \{-(pu')' + qu\}.$$

Dann lautet Gl. (1)

$$Lu = \lambda u + f, \quad (2)$$

das heißt, wir suchen ein $u \in \mathcal{D}$ mit $Lu = \lambda u + f$.

Wir definieren ferner

Definition 8.5.1: Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert zum Operator L , wenn es ein $u \in \mathcal{D}$, $u \neq 0$, gibt mit

$$Lu = \lambda u.$$

Die Funktion u heißt dann Eigenfunktion zum Eigenwert λ und wird durch $\|u\|_{\mathcal{L}}^2 = 1$ normiert.

Es ist unser Ziel, in diesem Abschnitt folgendes zu zeigen:

1. L besitzt abzählbar unendlich viele Eigenwerte $\lambda_n \rightarrow \infty$ mit zugehörigen orthonormalen Eigenfunktionen.
2. Jede genügend glatte Funktion g , die der Randbedingung genügt, kann in eine gleichmäßig konvergente Fourierreihe nach diesen Eigenfunktionen entwickelt werden.
3. Aus $h \in \mathcal{L}^2(I)$ und $(h, u_n) = 0$ für alle n folgt $\|h\|_{\mathcal{L}^2} = 0$. Das System der $\{u_n\}$ ist also vollständig.
4. Lösung der Randwertaufgabe.

Wir beginnen mit einigen vorbereitenden Hilfssätzen.

Lemma 8.5.2: $f, g \in \mathcal{D} \implies (Lf, g) = (f, Lg)$.

Der Beweis folgt durch partielles Integrieren

$$(Lf, g) = \int_I \{(-pf')' + qf\} \bar{g} = B(f, g) = (f, Lg). \quad (3)$$

Dabei haben wir zur Abkürzung

$$B(f, g) := \int_I \{p f' \bar{g}' + q f \bar{g}\}$$

gesetzt. $B(\cdot, \cdot)$ heißt Dirichletsche Form zu L .

Lemma 8.5.3: Alle Eigenwerte sind reell.

Es sei nämlich $Lu = \lambda u$ mit $\|u\|_{\mathcal{L}}^2 = 1$. Dann folgt aus Lemma 8.5.2

$$\lambda = (Lu, u) = (u, Lu) = \bar{\lambda}.$$

Wir wollen uns daher von nun an auf reellwertige Funktionen beschränken.

Lemma 8.5.4: $\lambda_m \neq \lambda_n \implies (u_m, u_n) = 0$.

Auch das folgt aus Lemma 8.5.2. Es gilt nämlich

$$\lambda_m (u_m, u_n) = (Lu_m, u_n) = (u_m, Lu_n) = \lambda_n (u_m, u_n).$$

Satz 8.5.5: Die Eigenwerte sind von unten beschränkt. Es sei für $x \in J$

$$q_0 := \min_{x \in J} \frac{q(x)}{r(x)}.$$

Dann ist $\lambda \geq q_0$.

Es sei nämlich u die Eigenfunktion zu λ . Dann folgt

$$\lambda = (Lu, u) = B(u, u) \geq \int_I q|u|^2 \geq q_0 \int_I r|u|^2.$$

Wegen $\|u\|_{\mathcal{L}}^2 = 1$ ist daher $\lambda \geq q_0$.

Korollar 8.5.6: Es sei $q \geq 0$. Dann sind alle Eigenwerte positiv.

Aus Satz 8.5.5 folgt zunächst nur $\lambda \geq 0$. Es sei also $\lambda = 0$ und u Eigenfunktion. Dann ist $B(u, u) = 0$, das heißt

$$\int_I p|u'|^2 = 0.$$

Wegen $p > 0$ folgt daraus $u' = 0$, also $u = \text{const}$ und wegen der Randbedingung schließlich $u = 0$ im Widerspruch zu $\|u\|_{\mathcal{L}}^2 = 1$.

Korollar 8.5.6 gilt beispielsweise bei der Neumannschen Randwertaufgabe nicht. Hier ist $\lambda = 0$ möglich; die Konstante ist Eigenfunktion, wenn q verschwindet.

Es ist oft bequemer, mit positiven Eigenwerten rechnen zu können. Wir verschieben deshalb die Eigenwerte bei negativem q_0 um $-q_0$ nach rechts. Es seien

$$L_0 := L - q_0$$

$$B_0(u, u) := (L_0 u, u) = B(u, u) - q_0(u, u) \geq 0.$$

Aus $Lu_n = \lambda_n u_n$ folgt dann

$$L_0 u_n = (\lambda_n - q_0) u_n =: \lambda_n^0 u_n,$$

und die Eigenwerte λ_n^0 von L_0 sind analog Korollar 8.5.6 positiv.

8.5.2 Die Greensche Funktion

Wir haben schon im zweiten Abschnitt die Äquivalenz von Anfangswertaufgaben zu Integralgleichungen herausgestellt. Wir wollen jetzt ähnlich vorgehen und zeigen (wenn $\lambda = 0$ kein Eigenwert ist)

$$Lu = f \iff u = Mf$$

mit

$$(Mf)(x) = \int_I f(y)G(x, y)r(y) dy.$$

Dabei ist M ein Integraloperator; er wird von der *Greenschen Funktion* $G : J \times J \rightarrow \mathbb{R}$ erzeugt, benannt nach G. GREEN (1793–1841).

Als erstes stellen wir die Eigenschaften einer Greenschen Funktion heraus und zeigen dann ihre Existenz. Man überlegt sich, daß G eine *singuläre Lösung* der Differentialgleichung sein muß. In Distributionenschreibweise ist

$$\forall y \in I \quad LG(\cdot, y) = \delta(\cdot, y).$$

Klassisch verwenden wir

Definition 8.5.7: $G \in C(J \times J, \mathbb{R})$ heißt *Greensche Funktion* zu L , wenn für alle $y \in I$ folgendes gilt:

1. $G(\cdot, y) \in C_2(J \setminus \{y\}, \mathbb{R})$
2. $G(a, y) = G(b, y) = 0$
3. $(LG)(x, y) = 0$ für $x \neq y$
4. $p(y)\{\partial_1 G(y+0, y) - \partial_1 G(y-0, y)\} = -1$.

Das Wort *singuläre Lösung* bedeutet also, daß die erste Ableitung von G an der Stelle y nicht mehr stetig ist. ∂_1 bezeichnet die Ableitung von G nach der ersten Variablen (die zweite, nämlich y , ist fest).

Wenn $\lambda = 0$ kein Eigenwert von L ist, läßt sich die Greensche Funktion relativ leicht angeben, also zum Beispiel für L_0 . Wenn $\lambda = 0$ Eigenwert ist, weiß man schon aus der linearen Algebra, daß dort die zu $Lu = f$ analoge Gleichung nicht mehr für alle f gelöst werden kann. Dann existiert nur eine *Greensche Funktion im erweiterten Sinne*. Dieser Fall ist für die Eigenwerttheorie nicht sehr interessant, weil man sich wegen der Linearität auf L_0 zurückziehen kann. Ich möchte auf die Greensche Funktion im erweiterten Sinne deshalb in dieser Vorlesung nicht weiter eingehen.

Es sei also für den Rest dieses Abschnitts §8.5.2 $\lambda = 0$ nicht Eigenwert von L . Dann konstruieren wir eine Greensche Funktion zu L folgendermaßen:

1. $Lu = 0$, $u(a) = 0$ und $u'(a) = 1$ ist eindeutig lösbar.
2. $Lv = 0$, $v(b) = 0$ und $v'(b) = 1$ ist eindeutig lösbar.

Beide Funktionen, u und v , verschwinden nicht identisch. Sie sind aus $C_2(J)$. Es sei W ihre Wronskische Determinante,

$$W = \det \begin{pmatrix} u & v \\ u' & v' \end{pmatrix}.$$

Dann gilt

$$(pW)' = \det \begin{pmatrix} u & v \\ pu' & pv' \end{pmatrix}' = \det \begin{pmatrix} u & v \\ qu & qv \end{pmatrix} = 0.$$

Die Funktionen u und v sind linear unabhängig, weil nach Voraussetzung $\lambda = 0$ kein Eigenwert ist. Mithin verschwindet W nicht, es gilt also

$$pW = c \quad \text{mit} \quad c \in \mathbb{R} \setminus \{0\}.$$

Es sei nun

$$G(x, y) := -\frac{1}{c} \begin{cases} u(x) \cdot v(y) & \text{für } x \leq y \\ u(y) \cdot v(x) & \text{für } x \geq y. \end{cases} \quad (4)$$

Dann ist G Greensche Funktion. G ist nämlich in $J \times J$ stetig, und auch die Bedingungen 1–3 sind offensichtlich erfüllt. Es bleibt der Nachweis von 4:

$$p(y)\{\partial_1 G(y+0, y) - \partial_1 G(y-0, y)\} = -\frac{1}{c}p(y)\{u(y)v'(y) - u'(y)v(y)\} = -\frac{1}{c}(pW) = -1.$$

Aus der Darstellung (4) der Greenschen Funktion folgt unmittelbar ihre Symmetrie, also

$$G(x, y) = G(y, x). \quad (5)$$

Beispiel 8.5.8: Es seien $a = 0$, $b = 1$ und $L = -d^2/dx^2$.

Dann ist $u(x) = x$, $v(x) = x - 1$ und

$$W = \det \begin{pmatrix} x & (x-1) \\ 1 & 1 \end{pmatrix} = 1,$$

also

$$G(x, y) = \begin{cases} x(1-y) & \text{für } x \leq y \\ y(1-x) & \text{für } x \geq y. \end{cases}$$

8.5.3 Randwertaufgaben

In diesem Abschnitt wollen wir

$$Lu = f \quad (6)$$

lösen, und zwar unter der Voraussetzung, daß $\lambda = 0$ nicht Eigenwert von L ist. Den allgemeinen Fall behandeln wir im Anschluß an die Eigenwerttheorie.

Es sei

$$(Mf)(x) := \int_a^b f(y)G(x, y)r(y)dy. \quad (7)$$

Wir wollen zeigen, daß $M : C(J) \rightarrow \mathcal{D}$ abbildet und daß $L(Mf) = f$ ist. Dann gilt

Satz 8.5.9: $\lambda = 0$ sei nicht Eigenwert von L . Dann ist

$$Lu = f$$

für alle $f \in C(J)$ eindeutig lösbar; $u = Mf$ ist die Lösung.

Wegen der Stetigkeit von G gilt nämlich $Mf \in C(J, \mathbb{R})$ und $(Mf)(a) = (Mf)(b) = 0$. Weiter ist

$$(Mf)'(x) = \left(\int_a^x \dots + \int_x^b \dots \right)'(x) = \int_a^b f(y)\partial_1 G(x, y)r(y)dy,$$

also

$$(p(Mf)')'(x) = \int_a^b f(y)q(y)G(x, y)r(y)dy + p(x)f(x)\{\partial_1 G(x, x-0) - \partial_1 G(x, x+0)\}.$$

Wegen der speziellen Gestalt (4) von G ist dann auch $Mf \in \mathcal{D}$, und es folgt

$$L(Mf)(x) = -p(x)f(x)\{\partial_1 G(x+0, x) - \partial_1 G(x-0, x)\} = f(x).$$

Damit ist Satz 8.5.9 bewiesen.

8.5.4 Eigenwertaufgaben

In diesem Abschnitt wollen wir Existenz und Eigenschaften der Eigenwerte von L herleiten. Wir suchen also Zahlen $\lambda \in \mathbb{R}$, für die

$$Lu = \lambda u \quad \text{mit } \|u\|_{\mathcal{L}}^2 = 1 \quad (8)$$

gilt.

Zu festem λ kann es nur eine linear unabhängige Eigenfunktion geben. $pW(u_1, u_2)$ ist nämlich für zwei Lösungen u_1, u_2 konstant und verschwindet wegen $u_1(a) = u_2(a) = 0$. Mithin sind u_1, u_2 linear abhängig.

Wir wollen die Eigenwertaufgabe noch etwas vereinfachen. Weil wir die Eigenwerte verschieben können, nehmen wir in diesem Abschnitt 8.5.4 an, daß alle Eigenwerte positiv sind (Übergang von L zu L_0). Es gibt noch andere Transformationen, die unsere Gleichung weiter vereinfachen. So kann man für differenzierbare r von u zu $v := \sqrt{r}u$ übergehen und sich so von r befreien. Das wollen wir aber hier nicht weiter verfolgen.

Es sei also für den Rest dieses Abschnitts und in §8.5.5 $\lambda \in \mathbb{R}^+$. Dann existiert nach §8.5.3 ein

$$M : C(J) \longrightarrow \mathcal{D}$$

mit $LM = id$. Es sei $u \in \mathcal{D}$ und $f := Lu$. Dann folgt

$$LMf = f = Lu,$$

also auch für alle $u \in \mathcal{D}$

$$u = Mf = MLu.$$

Damit erhalten wir aus Gl. (8) die äquivalente Integralgleichung

$$Mu = \mu u, \quad \|u\|_{\mathcal{L}}^2 = 1 \quad (9)$$

mit $\mu \in \mathbb{R}^+$ und $u \in C(J)$.

Es sei nämlich u eine Lösung von Gl. (9). Dann erhalten wir mit $\lambda := 1/\mu$

$$Lu = \lambda u, \quad \|u\|_{\mathcal{L}}^2 = 1.$$

Zum Nachweis der Existenz von Eigenwerten benötigen wir nun einige Vorbereitungen.

Lemma 8.5.10: *Es gilt auch $M : \mathcal{L}^2(I) \rightarrow C(J)$ mit*

$$\|Mf\| \leq k_1 \|f\|_{\mathcal{L}}^2 \quad \text{und} \quad \|Mf\|_{\mathcal{L}}^2 \leq k_2 \|f\|_{\mathcal{L}}^2.$$

Dabei sind

$$k_1 := \sup_{x,y \in J} |G(x,y)| \cdot \|1\|_{\mathcal{L}}^2$$

$$k_2 := \sqrt{\int_{I \times I} G^2(x,y)r(x)r(y) dx dy}.$$

Der Beweis folgt aus

$$(Mf)(x) = \int_I f(y)G(x,y)r(y) dy,$$

also

$$\begin{aligned} |(Mf)(x) - (Mf)(x_0)| &\leq \sup_{y \in J} |G(x,y) - G(x_0,y)| \cdot \int_I |f(y)r(y)| dy \\ &\leq \sup_{y \in J} |G(x,y) - G(x_0,y)| \cdot \|f\|_{\mathcal{L}}^2 \cdot \|1\|_{\mathcal{L}}^2 \end{aligned}$$

und

$$\begin{aligned} \|Mf\|_{\mathcal{L}^2}^2 &= \int_I \left| \int_I f(y)G(x,y)r(y) dy \right|^2 r(x) dx \\ &\leq \|f\|_{\mathcal{L}^2}^2 \int_{I \times I} G^2(x,y)r(y)r(x) dx dy. \end{aligned}$$

Lemma 8.5.11: (Mf, f) verschwindet nicht identisch.

Es sei nämlich $(Mf, f) = 0$ für alle $f \in \mathcal{L}^2(I)$. Dann gilt

$$\forall f, g \in \mathcal{L}^2(I) \quad 0 = (M(f+g), f+g) = 2(Mf, g).$$

Dabei haben wir die Symmetrie von M benutzt, die aus der Symmetrie der Greenschen Funktion folgt. Mit $g := Mf$ ist also

$$\forall f \in \mathcal{L}^2(I) \quad 0 = \|Mf\|_{\mathcal{L}^2}^2.$$

Daraus folgt $\|f\|_{\mathcal{L}^2} = 0$, und für $f \in C(J)$ ist $f = LMf = 0$.

Lemma 8.5.12: Es gilt

$$\sup_{\|f\|_{\mathcal{L}^2}^2=1} |(Mf, f)| = \sup_{\|f\|_{\mathcal{L}^2}^2=1} \|Mf\|_{\mathcal{L}^2}^2 =: \|M\|.$$

$\|M\|$ ist die *Operatornorm*, vgl. §9.2. Wir beweisen Lemma 8.5.12: Zunächst ist

$$|(Mf, f)| \leq \|Mf\|_{\mathcal{L}^2}^2 \cdot \|f\|_{\mathcal{L}^2}^2.$$

Wir müssen also

$$\|M\| \leq \sigma := \sup_{\|f\|_{\mathcal{L}^2}^2=1} |(Mf, f)|$$

zeigen. Dazu bilden wir mit $\alpha > 0$

$$\begin{aligned} \|Mf\|_{\mathcal{L}^2}^2 &= \frac{1}{4} \left\{ (M(\alpha f + \frac{1}{\alpha} Mf), \alpha f + \frac{1}{\alpha} Mf) - (M(\alpha f - \frac{1}{\alpha} Mf), \alpha f - \frac{1}{\alpha} Mf) \right\} \\ &\leq \frac{\sigma}{4} \left\{ \|\alpha f + \frac{1}{\alpha} Mf\|_{\mathcal{L}^2}^2 + \|\alpha f - \frac{1}{\alpha} Mf\|_{\mathcal{L}^2}^2 \right\} \\ &= \frac{\sigma}{2} \left\{ \alpha^2 \|f\|_{\mathcal{L}^2}^2 + \frac{1}{\alpha^2} \|Mf\|_{\mathcal{L}^2}^2 \right\}. \end{aligned}$$

Es seien nun $f \neq 0$ und $Mf \neq 0$. Dann folgt mit $\alpha^2 := \|Mf\|_{\mathcal{L}^2} / \|f\|_{\mathcal{L}^2}$

$$\|Mf\|_{\mathcal{L}^2}^2 \leq \sigma \|Mf\|_{\mathcal{L}^2} \cdot \|f\|_{\mathcal{L}^2}$$

oder

$$\forall f \in \mathcal{L}^2(I) \quad \|Mf\|_{\mathcal{L}^2} \leq \sigma \|f\|_{\mathcal{L}^2}.$$

Das war zu zeigen.

Lemma 8.5.13: Es seien $f_n \in \mathcal{L}^2(I)$ mit $\|f_n\|_{\mathcal{L}^2} \leq 1$. Dann enthält (Mf_n) eine bezüglich der Supremumsnorm konvergente Teilfolge.

Beweis: Nach Lemma 8.5.10 bildet M in $C(J)$ ab. Es sei $g_n := Mf_n$. Dann ist

$$\|g_n\| \leq k_1 \|f_n\|_{\mathcal{L}^2} \leq k_1$$

und

$$|g_n(x) - g_n(x_0)|^2 \leq \sup_{y \in J} |G(x, y) - G(x_0, y)| \cdot \|id\|_{\mathcal{L}^2}.$$

Mithin ist (g_n) bezüglich $\|\cdot\|$ gleichmäßig beschränkt und gleichartig stetig. Auf Grund des Satzes von Arzelà-Ascoli enthält (g_n) deshalb eine in $C(J)$ konvergente Teilfolge.

Nach diesen Vorbereitungen beweisen wir

Satz 8.5.14: Gl. (9) besitzt einen Eigenwert μ mit $\mu = \|M\|$.

Es ist klar, daß es größere Eigenwerte nicht geben kann. Denn aus

$$M - \mu = -\mu \left(id - \frac{M}{\mu} \right)$$

folgt für $\mu > \|M\|$ die Existenz von

$$\left(id - \frac{M}{\mu} \right)^{-1} : \mathcal{L}^2(I) \longrightarrow \mathcal{L}^2(I),$$

nämlich die Neumannsche Reihe

$$\left(id - \frac{M}{\mu} \right)^{-1} = \sum_{n=0}^{\infty} \left(\frac{M}{\mu} \right)^n.$$

Zum Beweis wählen wir nun eine Folge (f_n) , $f_n \in \mathcal{L}^2(I)$ mit $\|f_n\|_{\mathcal{L}^2} = 1$ und

$$|(Mf_n, f_n)| \rightarrow \|M\|.$$

Das ist nach Lemma 8.5.12 möglich. Die (f_n) seien so gewählt, daß

$$(Mf_n, f_n) \rightarrow \mu \in \mathbb{R} \quad \text{mit} \quad |\mu| = \|M\| > 0$$

gilt. Wegen Lemma 8.5.13 enthält (Mf_n) eine bezüglich der Supremumsnorm konvergente Teilfolge. Es sei bereits (Mf_n) diese Folge und $f \in C(J)$ mit

$$\|Mf_n - \mu f\| \rightarrow 0.$$

Aus

$$0 \leq \|Mf_n - \mu f_n\|_{\mathcal{L}^2}^2 = \|Mf_n\|_{\mathcal{L}^2}^2 - 2\mu(Mf_n, f_n) + \mu^2 \leq 2\mu(\mu - (Mf_n, f_n)) \rightarrow 0$$

folgt

$$\|Mf_n - \mu f_n\|_{\mathcal{L}^2} \rightarrow 0$$

und

$$\|\mu(f_n - f)\|_{\mathcal{L}^2} \leq \|\mu f_n - Mf_n\|_{\mathcal{L}^2} + \|Mf_n - \mu f\|_{\mathcal{L}^2} \rightarrow 0,$$

also

$$\|f_n - f\|_{\mathcal{L}^2} \rightarrow 0 \quad \text{und} \quad \|f\|_{\mathcal{L}^2} = 1.$$

Damit erhalten wir

$$\|Mf - \mu f\|_{\mathcal{L}^2} = \lim_{n \rightarrow \infty} \|Mf_n - \mu f_n\|_{\mathcal{L}^2} = 0,$$

oder

$$Mf = \mu f, \quad \|f\|_{\mathcal{L}^2} = 1.$$

Wir wissen bereits, daß alle Eigenwerte positiv sind. Deshalb gilt $\mu = \|M\|$, und Satz 8.5.14 ist bewiesen.

Damit haben wir den größten Eigenwert $\mu_1 := \mu$ von M und die zugehörige Eigenfunktion $u_1 := f$ gefunden. Es ist $u_1 \in C(J)$,

$$Mu_1 = \mu_1 u_1,$$

und μ_1 ist durch

$$|\mu_1| = \sup_{\|f\|_{\mathcal{L}^2}=1} |(Mf, f)|$$

charakterisiert. Wegen $M : C(J) \rightarrow \mathcal{D}$ gilt auch $u_1 \in \mathcal{D}$ und mit $\lambda_1 := 1/\mu_1$

$$Lu_1 = \lambda_1 u_1, \quad \|u_1\|_{\mathcal{L}^2} = 1.$$

Damit haben wir auch den ersten Eigenwert von L gefunden.

Zur Bestimmung weiterer Eigenwerte setzen wir dieses Verfahren fort und bilden für $n \geq 2$

$$|\mu_n| := \sup_{\substack{\|f\|_{\mathcal{L}^2}=1 \\ (f, u_i)=0, i=1, \dots, n-1}} |(Mf, f)|.$$

Wir erhalten auf diese Weise eine Folge (λ_n) von Eigenwerten $\lambda_n > 0$ und zugehörige Eigenfunktionen $u_n \in \mathcal{D}$ mit

$$Lu_n = \lambda_n u_n$$

$$\|M\|^{-1} = \lambda_1 < \lambda_2 < \dots$$

$$(u_m, u_n) = \delta_{mn}.$$

Ferner gilt

Lemma 8.5.15: Die Eigenwerte häufen sich im Endlichen nicht.

Der Beweis folgt aus

$$\begin{aligned} 0 &\leq \int_I \left(G(x, y) - \sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i} \right)^2 r(y) dy \\ &= \int_I G^2(x, y) r(y) dy - \sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i^2} \end{aligned}$$

oder

$$\sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i^2} \leq \int_I G^2(x, y) r(y) dy \leq k_3 < \infty, \quad (10)$$

also

$$\sum_{i=1}^n \frac{1}{\lambda_i^2} \leq \int_{I \times I} G^2(x, y) r(y) r(x) dy dx = k_2^2. \quad (11)$$

Mithin konvergiert die Reihe links.

Lemma 8.5.16: $\forall f \in \mathcal{L}^2(I)$ ist $(Mf, f) \geq 0$.

Denn sonst besäße L negative Eigenwerte.

Lemma 8.5.17: Es gibt unendlich viele Eigenwerte.

Es seien nämlich $\lambda_1, \dots, \lambda_k$ die Eigenwerte,

$$H(x, y) := G(x, y) - \sum_{i=1}^k \frac{u_i(x)u_i(y)}{\lambda_i}$$

und N der analog M von H erzeugte Operator, also

$$N = M - \sum_{i=1}^k \frac{(\cdot, u_i)}{\lambda_i} u_i.$$

Dann folgt aus der Konstruktion der Eigenwerte

$$\sup_{\|f\|=1} |(Nf, f)| = \sup_{\substack{\|f\|_{\mathcal{L}^2}=1 \\ (f, u_i)=0, i=1, \dots, k}} |(Mf, f)| = 0.$$

Weil N ebenfalls symmetrisch ist, können wir Lemma 8.5.12 mit N verwenden und finden $N = 0$; das heißt

$$\forall f \in \mathcal{L}^2(I) \quad \int_I f(y) H(x, y) r(y) dy = 0.$$

Wegen $H \in C(J \times J)$ folgt daraus $H(x, y) = 0$ oder

$$G(x, y) = \sum_{i=1}^k \frac{u_i(x)u_i(y)}{\lambda_i}.$$

Dann wäre aber $G \in C_2(J \times J)$. Wir wissen jedoch, daß die erste Ableitung von G in $x = y$ springt. (Oder: Es wäre M endlich dimensional, das heißt $\dim R(M) < \infty$, im Widerspruch zu $R(M) = \mathcal{D}$.)

Mit dieser Konstruktion haben wir auch wirklich alle Eigenwerte erfaßt. Es sei nämlich $0 < \sigma < \infty$ ein weiterer Eigenwert zur Eigenfunktion v , also

$$Lv = \sigma v \quad \text{mit } \sigma \notin \{\lambda_1, \lambda_2, \dots\}.$$

Dann gilt nach Lemma 8.5.4

$$\forall n \quad (v, u_n) = 0$$

oder

$$\forall n \quad \frac{1}{\sigma} = (Mv, v) \leq \mu_n = \frac{1}{\lambda_n} \rightarrow 0.$$

Damit haben wir bewiesen:

Satz 8.5.18: L besitzt abzählbar unendlich viele Eigenwerte und zugehörige orthonormale Eigenfunktionen. Die Eigenwerte sind von unten beschränkt und häufen sich im Endlichen nicht.

Für praktische Zwecke hat das Konstruktionsverfahren allerdings den Nachteil, daß man zur Berechnung des n -ten Eigenwertes alle vorherigen kennen muß. Deshalb ist die folgende auf R. COURANT (1888–1972) zurückgehende Charakterisierung von Interesse:

Courantsches Mini-Max-Prinzip: Es seien $h_1, \dots, h_{n-1} \in \mathcal{L}^2(I)$ beliebige Funktionen und

$$\nu_n(h_1, \dots, h_{n-1}) := \sup_{\substack{\|f\|_{\mathcal{L}^2} = 1 \\ (f, h_i) = 0, i=1, \dots, n-1}} (Mf, f).$$

Dann gilt

$$\frac{1}{\lambda_n} = \mu_n = \inf_{h_1, \dots, h_{n-1} \in \mathcal{L}^2(I)} \nu_n(h_1, \dots, h_{n-1}).$$

Der Beweis benutzt das Ergebnis von Satz 8.5.18. Demnach ist

$$\nu_n(u_1, \dots, u_{n-1}) = \mu_n.$$

Wir müssen also $\nu_n(h_1, \dots, h_{n-1}) \geq \mu_n$ zeigen. Dazu wählen wir

$$u = \sum_{i=1}^n c_i u_i$$

mit $\|u\|_{\mathcal{L}^2} = 1$ und $(u, h_i) = 0$ für $i = 1, \dots, n-1$. Offenbar lassen sich geeignete Konstanten c_i immer finden. Dann folgt

$$\nu_n(h_1, \dots, h_{n-1}) \geq (Mu, u) = \sum_{i,k=1}^n c_i c_k (Mu_i, u_k) = \sum_{i=1}^n \frac{c_i^2}{\lambda_i} \geq \frac{1}{\lambda_n} = \mu_n.$$

8.5.5 Entwicklungssätze

In diesem Abschnitt zeigen wir die Vollständigkeit des Systems der Eigenfunktionen und Entwicklungen danach. Es seien $h \in \mathcal{L}^2(I)$ und

$$h_i := (h, u_i),$$

der i -te Fourierkoeffizient von h . Es seien ferner

$$g_n := h - \sum_{i=1}^n h_i u_i.$$

Dann gilt $(g_n, u_i) = 0$ für $i = 1, 2, \dots, n$ und

$$\|g_n\|_{\mathcal{L}^2}^2 = \|h\|_{\mathcal{L}^2}^2 - \sum_{i=1}^n h_i^2,$$

also

$$\sum_{i=1}^n h_i^2 \leq \|h\|_{\mathcal{L}^2}^2. \quad (12)$$

Das ist wieder die Besselsche Ungleichung. Mithin konvergiert

$$\sum_{i=1}^n h_i u_i \rightarrow h^* \in \mathcal{L}^2(I).$$

Wir möchten $h = h^*$ beweisen. Dazu holen wir etwas aus. Aus der Definition der Eigenwerte folgt wegen $(g_n, u_i) = 0$ für $i = 1, 2, \dots, n$

$$\|Mg_n\|_{\mathcal{L}^2} \leq \mu_n \|g_n\|_{\mathcal{L}^2} \leq \mu_n \|h\|_{\mathcal{L}^2} \rightarrow 0,$$

also

Lemma 8.5.19: Es seien $h \in \mathcal{L}^2(I)$ und $f := Mh$. Dann gilt

$$\left\| f - \sum_{i=1}^n f_i u_i \right\|_{\mathcal{L}^2} \rightarrow 0.$$

Es ist nämlich $f_i = h_i/\lambda_i$. Für solche *quellenmäßig darstellbare* Funktionen gilt also der Entwicklungssatz. Zum Beispiel sei $f \in C_2(J)$ mit $f(a) = f(b) = 0$ und $h := Lf$. Dann ist $f = Mh$.

Das allgemeine Resultat folgt dann durch Approximieren. Testfunktionen sind in $\mathcal{L}^2(I)$ dicht, und Testfunktionen sind *quellenmäßig darstellbar*. Folglich gilt für $h \in \mathcal{L}^2(I)$

$$\forall \varepsilon > 0 \quad \exists \varphi \in \mathring{C}_\infty(I) \quad \|h - \varphi\|_{\mathcal{L}^2} < \frac{\varepsilon}{2}$$

$$\forall \varepsilon > 0 \quad \forall \varphi \in \mathring{C}_\infty(I) \quad \exists n \in \mathbb{N} \quad \exists a_i \in \mathbb{R} \quad \left\| \varphi - \sum_{i=1}^n a_i u_i \right\|_{\mathcal{L}^2} < \frac{\varepsilon}{2},$$

also

$$\forall \varepsilon > 0 \quad \exists n \in \mathbb{N} \quad \exists a_i \in \mathbb{R} \quad \left\| h - \sum_{i=1}^n a_i u_i \right\|_{\mathcal{L}^2} < \varepsilon.$$

Es sei betont, daß die a_i von ε abhängen. Die a_i sind nicht die Fourierkoeffizienten von h . Es folgt

Lemma 8.5.20: Es seien $h \in \mathcal{L}^2(I)$ und $(h, u_i) = 0$ für alle i . Dann ist $\|h\|_{\mathcal{L}^2} = 0$.

Aus der letzten Abschätzung erhält man nämlich durch Quadrieren

$$\forall \varepsilon > 0 \quad \|h\|_{\mathcal{L}^2}^2 + \sum_{i=1}^n a_i^2 < \varepsilon^2,$$

also verschwindet h .

Es gilt

Satz 8.5.21: Es sei $h \in \mathcal{L}^2(I)$. Dann ist

$$\left\| h - \sum_{i=1}^n h_i u_i \right\|_{\mathcal{L}^2} \rightarrow 0.$$

Wir wissen nämlich bereits $\sum_{i=1}^n h_i u_i \rightarrow h^* \in \mathcal{L}^2(I)$, und $g := h - h^*$ hat die Fourierkoeffizienten Null.

Aus Satz 8.5.21 folgt unmittelbar die Parsevalsche Gleichung

$$\|h\|_{\mathcal{L}^2}^2 = \sum_{i=1}^{\infty} h_i^2. \quad (13)$$

Man kann den Entwicklungssatz 8.5.21 verschärfen, wenn die darzustellende Funktion glatter ist. Es sei etwa $f := Mh$ *quellenmäßig dargestellt*. Dann ist nach Gl. (10)

$$\left(\sum_{i=m}^n |f_i u_i| \right)^2 = \left(\sum_{i=m}^n |h_i \frac{u_i}{\lambda_i}| \right)^2 \leq \left(\sum_{i=m}^n h_i^2 \right) \left(\sum_{i=m}^n \frac{u_i^2}{\lambda_i^2} \right) \leq k_3 \sum_{i=m}^n h_i^2 < \infty.$$

Es gilt also

Satz 8.5.22: Es sei $f = Mh$ mit $h \in \mathcal{L}^2(I)$. Dann ist

$$f(x) = \sum_{i=1}^{\infty} f_i u_i(x),$$

und die Reihe konvergiert absolut und gleichmäßig.

Es gibt auch Entwicklungssätze für die Greensche Funktion selbst. Naheliegen würde

$$G(x, y) = \sum_{i=1}^{\infty} \frac{u_i(x) u_i(y)}{\lambda_i}.$$

Um eine solche Darstellung herzuleiten, müssen wir aber etwas ausholen und beweisen zunächst ein allgemeineres Konvergenzresultat, nämlich den

Satz von Dini: Es seien $F_n \in C(J, \mathbb{R}_0^+)$ sowie

$$\forall x \in J \quad F_n(x) \leq F_{n+1}(x) \quad \text{und} \quad F_n(x) \rightarrow F(x) \in C(J).$$

Dann folgt

$$\|F_n - F\| \rightarrow 0.$$

Zum Beweis setzen wir

$$F_n(x) + R_n(x) = F(x)$$

und nehmen

$$\exists \varepsilon > 0 \quad \forall i \quad \exists n_i \quad \exists x_i \quad R_{n_i}(x_i) \geq \varepsilon$$

an. Dann ist

$$F_{n_i}(x_i) = F(x_i) - R_{n_i}(x_i) \leq F(x_i) - \varepsilon,$$

und insbesondere gilt für jedes feste N und alle $n_i \geq N$

$$F_N(x_i) \leq F_{n_i}(x_i) \leq F(x_i) - \varepsilon.$$

Wählen wir nun eine Teilfolge der (x_i) , die gegen $x_0 \in J$ konvergiert. Dann folgt

$$\forall N \in \mathbb{N} \quad F_N(x_0) \leq F(x_0) - \varepsilon.$$

Das steht aber im Widerspruch zur vorausgesetzten punktweisen Konvergenz der Folge $(F_N(x))$. Damit ist der Satz bewiesen; er ist nach U. DINI (1845–1918) benannt.

Um den Entwicklungssatz für G herzuleiten, glätten wir zunächst wieder und bilden die iterierte Greensche Funktion

$$G_2(x, y) := \int_I G(x, z)G(z, y)r(z) dz.$$

Offenbar wird $G_2(\cdot, y)$ quellenmäßig dargestellt, es gilt also bei festem y

$$G_2(x, y) = \sum_{i=1}^{\infty} \frac{u_i(x)u_i(y)}{\lambda_i^2}.$$

Diese Reihe konvergiert absolut und gleichmäßig bezüglich beider Variablen. Es gilt nämlich nach Gl. (10)

$$\left(\sum_{i=m}^n \frac{|u_i(x)u_i(y)|}{\lambda_i^2} \right)^2 \leq \left(\sum_{i=m}^n \frac{u_i^2(x)}{\lambda_i^2} \right) \left(\sum_{i=m}^n \frac{u_i^2(y)}{\lambda_i^2} \right) \leq k_3 \sum_{i=m}^n \frac{u_i^2(x)}{\lambda_i^2}.$$

Wegen

$$\sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i^2} \rightarrow G_2(x, x) \in C(J)$$

folgt dann die Konvergenz der Reihe in der Supremumsnorm, also

Satz 8.5.23: Es konvergiert absolut und gleichmäßig bezüglich beider Variablen

$$\sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i^2} \rightarrow G_2(x, y).$$

Korollar 8.5.24: Es konvergiert gleichmäßig bezüglich $x \in J$

$$\left\| G(x, \cdot) - \sum_{i=1}^n \frac{u_i(x)u_i(\cdot)}{\lambda_i} \right\|_{\mathcal{L}^2} \rightarrow 0.$$

Der Beweis folgt aus

$$\int_I \left(G(x, y) - \sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i} \right)^2 r(y) dy = G_2(x, x) - \sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i^2}.$$

Nach diesen Vorbereitungen können wir das Hauptresultat dieses Abschnitts zeigen, nämlich den

Satz von Mercer: *Es konvergiert*

$$\sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i} \rightarrow G(x, y),$$

und zwar absolut und gleichmäßig in beiden Variablen.

Der Satz geht auf J. MERCER (1909) zurück. Sein Beweis erfolgt in zwei Schritten:

1. Für alle $x \in J$ ist $G(x, x) \geq 0$. Denn aus $G(x_0, x_0) < 0$ würde $G(x, y) < 0$ in einer Umgebung von (x_0, x_0) folgen. Das steht aber im Widerspruch zu Lemma 8.5.16.

2. Der von

$$H(x, y) := G(x, y) - \sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i}$$

erzeugte Operator N hat nur positive Eigenwerte, nämlich

$$\mu_{n+1} > \mu_{n+2} > \dots > 0.$$

u_{n+1}, u_{n+2}, \dots sind die zugehörigen Eigenfunktionen. Das folgt aus

$$\forall f \in \mathcal{L}^2(I) \quad (Nf, f) = \sum_{i=n+1}^{\infty} \mu_i(f, u_i)^2.$$

Nach 1 ist dann auch $H(x, x) \geq 0$ und

$$\sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i} \leq G(x, x).$$

Die Summe links konvergiert deshalb für jedes feste x , und es folgt für festes y

$$\left(\sum_{i=m}^n \frac{u_i(x)u_i(y)}{\lambda_i} \right)^2 \leq \left(\sum_{i=m}^n \frac{u_i^2(x)}{\lambda_i} \right) \left(\sum_{i=m}^n \frac{u_i^2(y)}{\lambda_i} \right) \leq c \sum_{i=m}^n \frac{u_i^2(x)}{\lambda_i}.$$

Mithin konvergiert

$$\sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i}$$

bei festem y absolut und gleichmäßig bezüglich x , und zwar gegen $G(x, y)$ nach Korollar 8.5.24. Also gilt punktweise

$$\sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i} \rightarrow G(x, x).$$

Aus dem Satz von Dini folgt deshalb

$$\left\| \sum_{i=1}^n \frac{u_i^2(x)}{\lambda_i} - G(x, x) \right\| \rightarrow 0.$$

Mithin konvergiert

$$\sum_{i=1}^n \frac{u_i(x)u_i(y)}{\lambda_i}$$

absolut und gleichmäßig bezüglich beider Variablen. Das beweist den Satz von Mercer. Es folgt

Korollar 8.5.25:

$$\sum_{i=1}^n \frac{1}{\lambda_i} \rightarrow \int_I G(x, x) r(x) dx.$$

8.5.6 Anwendungen

Es sei nun wieder L der in §8.5.1 definierte Operator und $L_0 = L - q_0$. L_0 hat nur positive Eigenwerte $\lambda_n^0 = \lambda_n - q_0$. Es sei M_0 der entsprechend M zu L_0 gebildete Operator.

1. Als erstes wollen wir eine Entwicklung für den Operator L selbst nach Eigenfunktionen angeben. Aus §8.5.4,5 folgt für $h \in \mathcal{L}^2(I)$

$$f := M_0 h = \sum_{i=1}^{\infty} f_i u_i = \sum_{i=1}^{\infty} \frac{h_i}{\lambda_i^0} u_i, \quad (14)$$

und die Reihen konvergieren absolut und gleichmäßig in $C(J)$ und damit erst recht in $\mathcal{L}^2(I)$.

Es sei nun $f \in C(J)$ und $v := M_0 f \in \mathcal{D}$. Dann ist $L_0 v = f$, also

$$L_0 v = f = \sum_{i=1}^{\infty} f_i u_i = \sum_{i=1}^{\infty} \lambda_i^0 v_i u_i.$$

Diese Reihen konvergieren bezüglich der $\mathcal{L}^2(I)$ -Norm. Daraus folgt

$$Lv = \sum_{i=1}^{\infty} \lambda_i^0 v_i u_i + q_0 \sum_{i=1}^{\infty} v_i u_i = \sum_{i=1}^{\infty} \lambda_i v_i u_i.$$

Zur Abkürzung schreiben wir das als

$$L \cdot = \sum_{i=1}^{\infty} \lambda_i(\cdot, u_i) u_i \quad (15)$$

und haben damit eine *Spektraldarstellung* des Operators L erhalten. Es ist unmittelbar klar, daß die Voraussetzung $f \in C(J)$ zur Konvergenz der Reihen unnötig stark war. Es würde $f \in \mathcal{L}^2(I)$ genügen. Deshalb können wir L auf eine größere Funktionenmenge fortsetzen, ohne die wesentliche Eigenschaft (15) zu verändern. Wir bezeichnen die Fortsetzung mit \tilde{L} und wählen

$$\mathcal{D}(\tilde{L}) := \left\{ v \in \mathcal{L}^2(I) \mid \sum_{i=1}^{\infty} \lambda_i^2 |(v, u_i)|^2 < \infty \right\}.$$

und

$$\begin{aligned} \tilde{L} : \mathcal{D}(\tilde{L}) &\longrightarrow \mathcal{L}^2(I) \\ v &\longmapsto \sum_{i=1}^{\infty} \lambda_i(v, u_i) u_i. \end{aligned}$$

Dann ist $\mathcal{D} \subset \mathcal{D}(\tilde{L})$. Man kann sich überlegen, daß

$$\mathcal{D}(\tilde{L}) = \left\{ v \in C_1(J) \mid (pv')' \in \mathcal{L}^2(I), v(a) = v(b) = 0 \right\}$$

ist. Dabei soll $(pv')'$ nur noch im schwachen Sinne existieren, das heißt, es soll gelten

$$\exists g \in \mathcal{L}^2(I) \quad \forall \phi \in \mathring{C}_\infty(I) \quad (v, (p\phi')') = (g, \phi).$$

Diese Fortsetzung des Operators L , die Diskussion von L und M im Hilbertraum $\mathcal{L}^2(I)$ und ihre Spektraldarstellungen haben sich als äußerst praktikabel erwiesen. Beachten Sie, daß wir mit der Erweiterung von \mathcal{D} zu $\mathcal{D}(\tilde{L})$ den Lösungsbegriff verallgemeinert haben. Wir verlangen jetzt gar nicht mehr die klassische Differenzierbarkeit der Lösungen u , sondern nur noch $u \in \mathcal{D}(\tilde{L})$. Die in diese Definition eingehenden Eigenwerte und -funktionen haben wir aber über Integralgleichungen gewonnen. Genauere Regularitätsuntersuchungen müßten das alles präzisieren. Im folgenden lasse ich die Tilde über L wieder fort.

Mit Hilfe der Spektraldarstellung (15) lassen sich auch Potenzen und sogar stetige Funktionen von L definieren. Natürlich $L^0 := id : \mathcal{L}^2(I) \rightarrow \mathcal{L}^2(I)$

$$L^0 = \sum_{i=1}^{\infty} (\cdot, u_i) u_i,$$

aber auch

$$\mathcal{D}(L^2) := \left\{ v \in \mathcal{L}^2(I) \mid \sum_{i=1}^{\infty} \lambda_i^4 |(v, u_i)|^2 < \infty \right\}$$

und

$$\begin{aligned} L^2 &: \mathcal{D}(L^2) \longrightarrow \mathcal{L}^2(I) \\ v &\longmapsto \sum_{i=1}^{\infty} \lambda_i^2(v, u_i) u_i \end{aligned}$$

oder im Falle nicht negativer Eigenwerte

$$\mathcal{D}(\sqrt{L}) := \left\{ v \in \mathcal{L}^2(I) \mid \sum_{i=1}^{\infty} \lambda_i |v, u_i|^2 < \infty \right\}$$

$$\begin{aligned} \sqrt{L} &: \mathcal{D}(\sqrt{L}) \longrightarrow \mathcal{L}^2(I) \\ v &\longmapsto \sum_{i=1}^{\infty} \sqrt{\lambda_i} (v, u_i) u_i. \end{aligned}$$

2. Die Spektraldarstellung (15) liefert einen handlichen Kalkül. Wir wollen ihn benutzen, um endlich Gleichung (2) zu lösen. Es sei also $f \in \mathcal{L}^2(I)$ vorgegeben; wir suchen ein $v \in \mathcal{D}(L)$ mit

$$Lv = \lambda v + f. \quad (16)$$

Verwenden wir die Darstellung (15), dann folgt die Äquivalenz dieser Gleichung mit

$$(\lambda_i - \lambda)v_i = f_i. \quad (17)$$

Dabei sind v_i, f_i wieder die Fourierkoeffizienten (v, u_i) bzw. (f, u_i) von v und f .

Wenn λ kein Eigenwert von L ist, dann ist Gl. (17) eindeutig lösbar, und wir erhalten

$$v := (L - \lambda id)^{-1} f := \sum_{i=1}^{\infty} \frac{f_i}{\lambda_i - \lambda} u_i.$$

Offenbar konvergiert die Reihe in $\mathcal{L}^2(I)$, und es gilt $v \in \mathcal{D}(L)$.

Wenn $\lambda = \lambda_n$ ein Eigenwert ist, dann ist Gl. (17) nicht mehr für alle f lösbar. Vielmehr muß f die notwendige und hinreichende Lösungsbedingung $f_n = 0$ erfüllen. In diesem Falle ist mit $c_n \in \mathbb{R}$

$$v := c_n u_n + \sum_{\substack{i=1 \\ i \neq n}}^{\infty} \frac{f_i}{\lambda_i - \lambda} u_i$$

die allgemeine Lösung von Gl. (16).

3. Wir behandeln die schwingende Saite. Es sei $I = (0, \pi)$ und $Lu = -u''$. Dann ist

$$G(x, y) = \frac{1}{\pi} \begin{cases} x(\pi - y) & \text{für } x \leq y \\ y(\pi - x) & \text{für } x \geq y. \end{cases}$$

Aus $Lu = \lambda u$ folgt $u'' + \lambda u = 0$, also

$$u(x) = \alpha \cos \sqrt{\lambda} x + \beta \sin \sqrt{\lambda} x.$$

Aus $u(0) = 0$ folgt dann $\alpha = 0$ und aus $u(\pi) = 0$ schließlich

$$\lambda_n = n^2 \quad \text{und} \quad u_n(x) = \sqrt{\frac{2}{\pi}} \sin(nx).$$

Der Satz von Mercer besagt daher

$$G(x, y) = \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin(nx) \sin(ny)}{n^2}.$$

4. Es gibt eine Fülle weiterer Anwendungen. Zum Beispiel kann man die Legendreschen Polynome als Eigenfunktionen deuten und Entwicklungssätze danach beweisen. Darauf kann ich aber hier nicht mehr eingehen. Ich

möchte aber doch wenigstens skizzieren, wieso die klassische Fourierreihe in unsere Theorie paßt. Es seien $I := (-\pi, \pi)$ und

$$\mathcal{D} := \{u \in C_1(\bar{I}, \mathbb{C}) \mid u(-\pi) = u(\pi)\}$$

$$Lu := -iu'.$$

$Lu = \lambda u$ ist jetzt eine Eigenwertaufgabe für einen Operator erster Ordnung. Man überzeugt sich leicht von der Symmetrie von L ,

$$\forall u, v \in \mathcal{D} \quad (Lu, v) = (u, Lv).$$

Aus $Lu = \lambda u$ folgt

$$u(x) = e^{i\lambda x} \quad \text{mit } e^{-i\lambda\pi} = e^{i\lambda\pi},$$

oder $\sin \lambda\pi = 0$. Das bedeutet $\lambda = n \in \mathbb{Z}$ und

$$u_n(x) = \sqrt{\frac{1}{2\pi}} e^{inx}.$$

Weil $\lambda = 0$ Eigenwert ist, werden die weiteren Überlegungen etwas erschwert, und es existiert nur eine Greensche Funktion im erweiterten Sinne. Deshalb schränken wir L zunächst auf

$$\mathcal{L}_0^2(I) := \{f \in \mathcal{L}^2(I) \mid (f, 1) = 0\}$$

ein. Zu jedem $f \in \mathcal{L}^2(I)$ gibt es dann ein $h := f - (f, u_0) u_0 \in \mathcal{L}_0^2(I)$ mit

$$f = (f, u_0) u_0 + h.$$

Damit folgt aus dem Entwicklungssatz für h wieder die Spektraldarstellung für f und L . Mit

$$g(t) := \frac{i}{2\pi} \begin{cases} \pi - t & \text{für } 0 < t < \pi \\ -\pi - t & \text{für } -\pi < t < 0 \end{cases}$$

ist $g'(t) = -i/2\pi$, also konstant, und mit

$$(Mf)(x) := \int_{-\pi}^{\pi} f(y)g(x-y) dy = \int_{-\pi}^x \dots + \int_x^{\pi} \dots$$

folgt wegen $(f, 1) = 0$

$$(Mf)'(x) = f(x)\{g(x - (x-)) - g(x - (x+))\} = f(x)\{g(0+) - g(0-)\} = if(x).$$

g ist deshalb Greensche Funktion für den eingeschränkten Operator, und wir können unsere Theorie darauf anwenden. Insbesondere folgt aus dem Satz von Mercer

$$\sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{u_n(x) \overline{u_n(y)}}{n} = \frac{1}{2\pi} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{e^{in(x-y)}}{n} = \frac{i}{\pi} \sum_{n=1}^{\infty} \frac{\sin n(x-y)}{n} = g(x-y)$$

in Übereinstimmung mit

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n} = \frac{\pi - x}{2} \quad \text{für } 0 < x < 2\pi.$$

Es sei noch bemerkt, daß man im Falle $I = \mathbb{R}$ analog die Fouriertransformation erhält.

5. Schließlich sei noch bemerkt, daß es auch nichtlineare Randwertaufgaben gibt. Sie sind allgemein recht schwierig zu behandeln. Damit Sie eine Vorstellung von ihnen erhalten, betrachten wir

$$Lu = f(\cdot, u).$$

Dabei sei $f \in C(\bar{I} \times \mathbb{R}, \mathbb{R})$ und genüge einer globalen Lipschitzbedingung. $\lambda = 0$ sei kein Eigenwert von L . Dann erhalten wir die äquivalente Fixpunktgleichung

$$u = Mf(\cdot, u) = F(u).$$

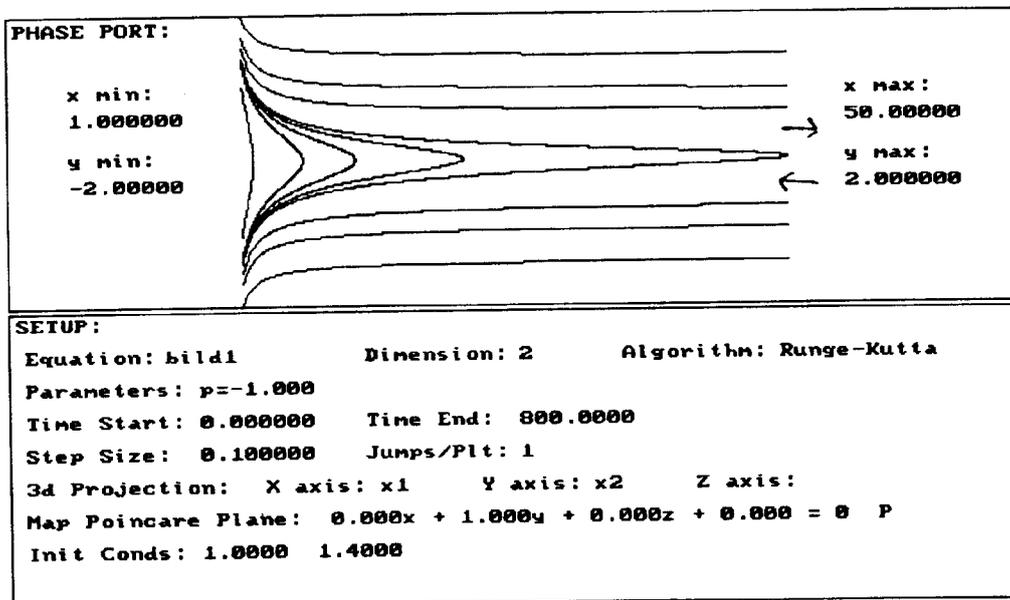
Diese läßt sich lösen, wenn die Lipschitzkonstante nicht zu groß ist. Das folgt aus

$$\|F(u_1) - F(u_2)\| = \|M(f(\cdot, u_1) - f(\cdot, u_2))\| \leq \|M\| L \|u_1 - u_2\|.$$

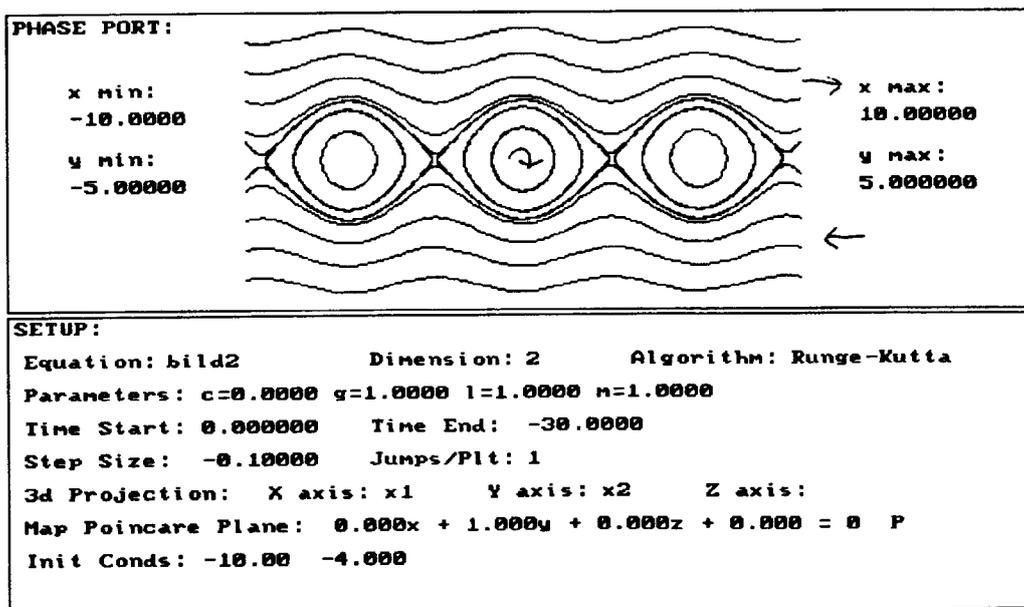
Der Banachsche Fixpunktsatz liefert dann die Lösung. Auch $f = f(\cdot, u, u')$ kann zugelassen werden.

8.6 Abbildungen

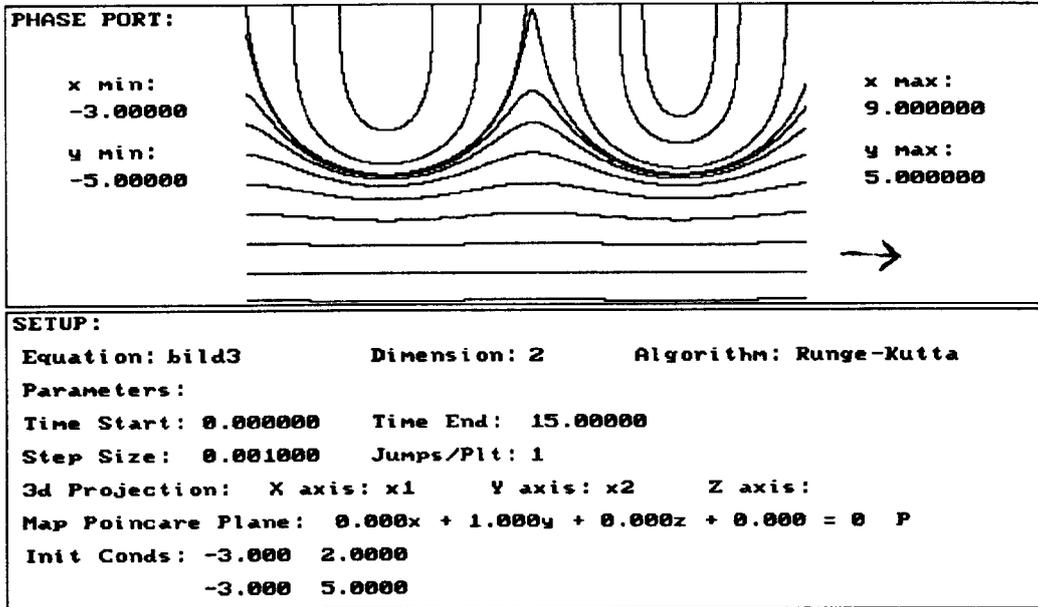
$$x' = \begin{pmatrix} x_2 \\ -1/(x_1)^2 \end{pmatrix}$$



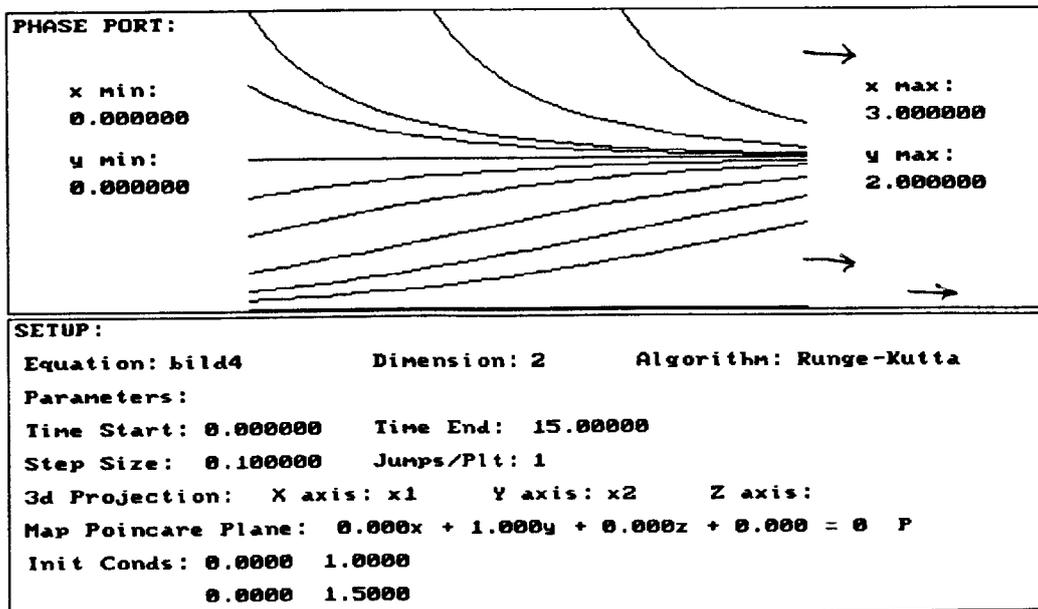
$$x' = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix}$$



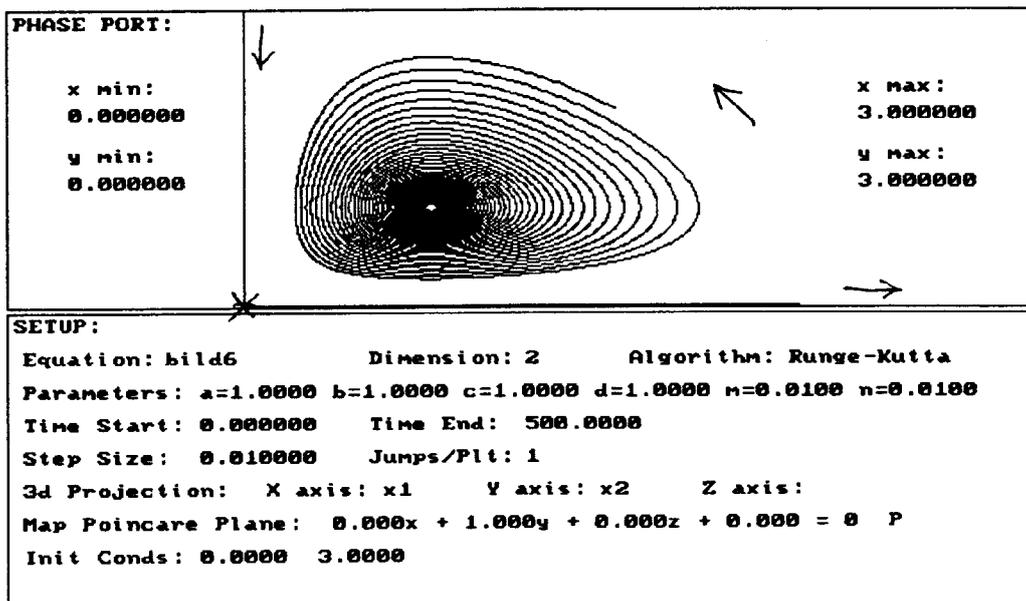
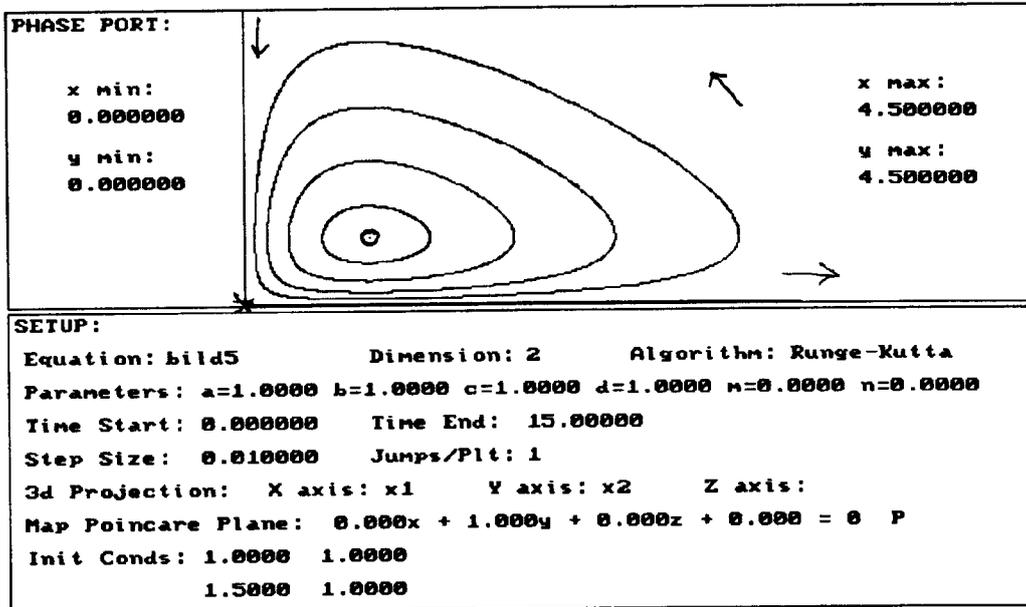
$$x' = \begin{pmatrix} 1 \\ e^{x_2} \sin x_1 \end{pmatrix}$$



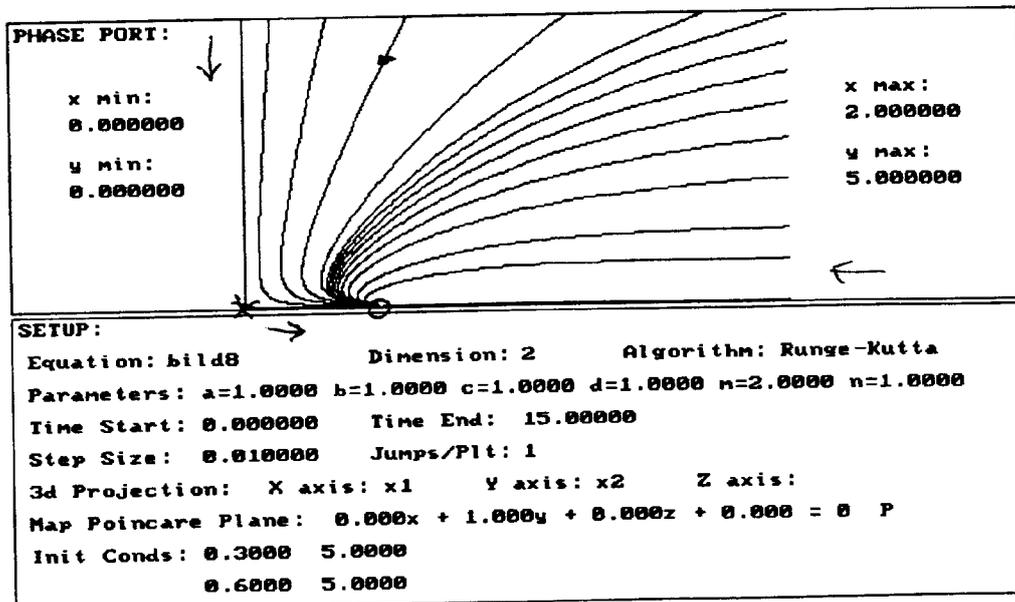
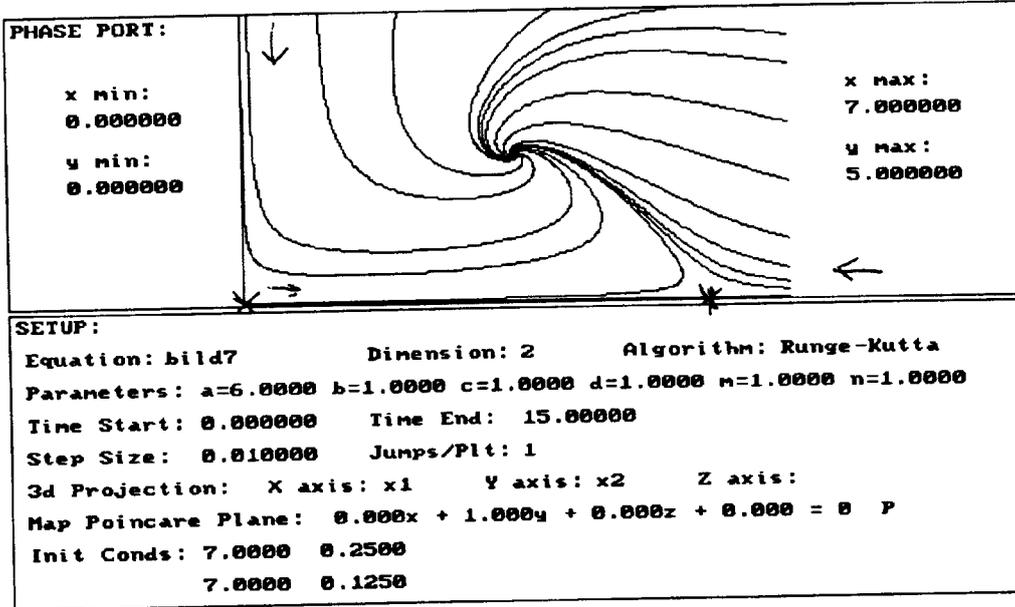
$$y' = y(1 - y)$$



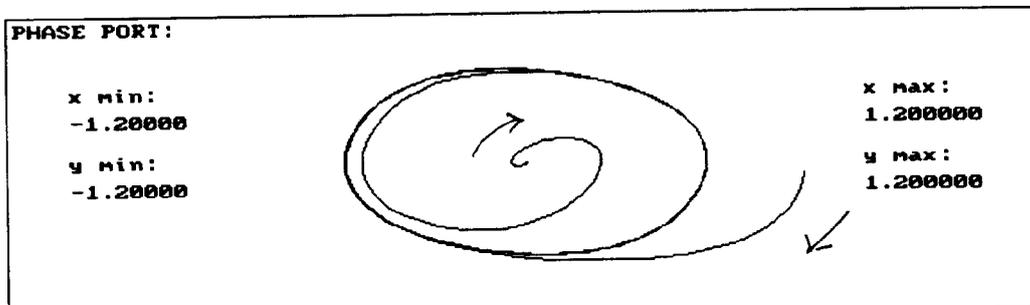
$$x' = \begin{pmatrix} (a - b x_2 - m x_1) x_1 \\ (-d + c x_1 - n x_2) x_2 \end{pmatrix}$$



$$x' = \begin{pmatrix} (a - b x_2 - m x_1) x_1 \\ (-d + c x_1 - n x_2) x_2 \end{pmatrix}$$

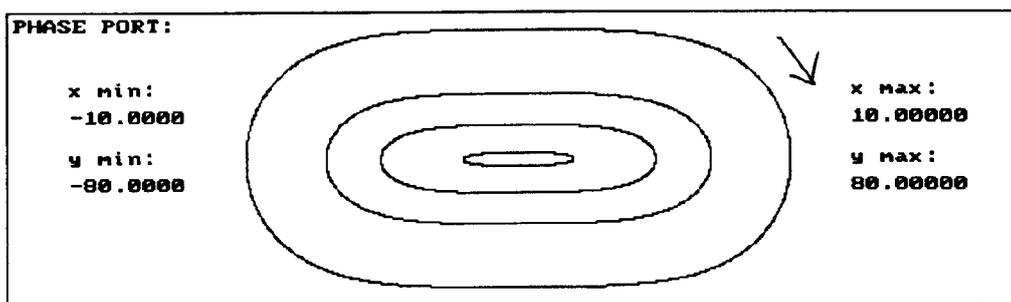


$$x' = \begin{pmatrix} x_2 \\ -x_1 + x_2(1 - x_1^2 - 2(x_2)^2) \end{pmatrix}$$



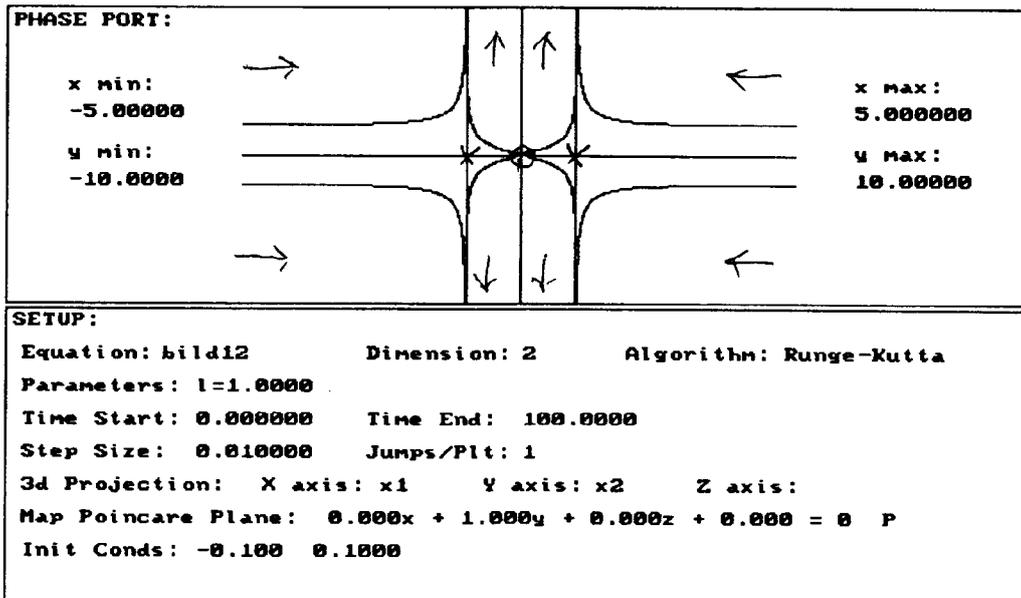
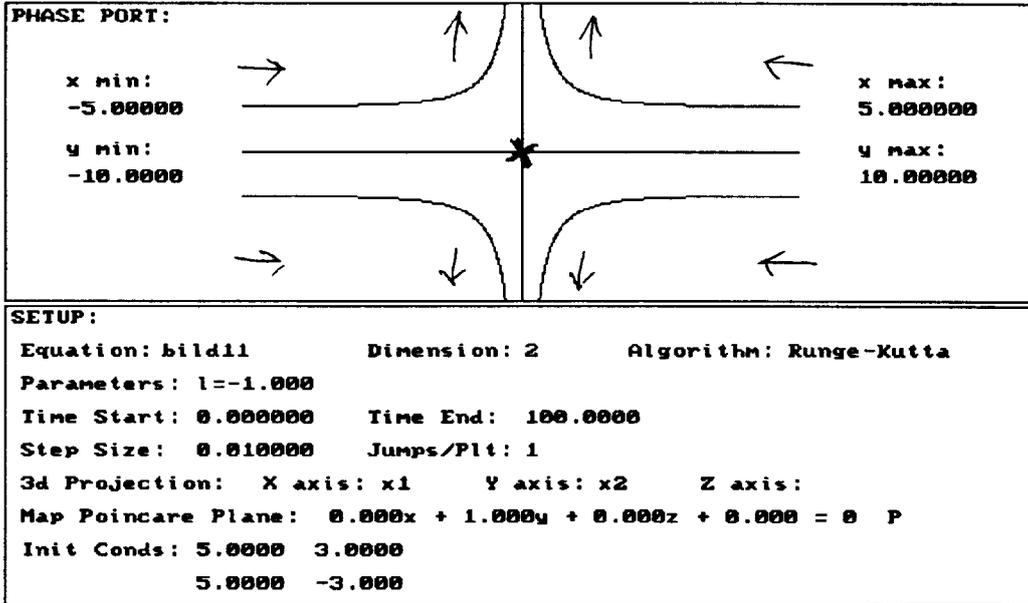
SETUP:
 Equation: bild9 Dimension: 2 Algorithm: Runge-Kutta
 Parameters:
 Time Start: 0.00000 Time End: 15.00000
 Step Size: 0.10000 Jumps/Plt: 1
 3d Projection: X axis: x1 Y axis: x2 Z axis:
 Map Poincare Plane: 0.000x + 1.000y + 0.000z + 0.000 = 0 P
 Init Conds: 0.0100 0.0000
 1.2000 0.0000

$$x' = \begin{pmatrix} x_2 \\ -x_1 - (x_1)^3 \end{pmatrix}$$



SETUP:
 Equation: bild10 Dimension: 2 Algorithm: Runge-Kutta
 Parameters:
 Time Start: 0.00000 Time End: 15.00000
 Step Size: 0.001000 Jumps/Plt: 1
 3d Projection: X axis: x1 Y axis: x2 Z axis:
 Map Poincare Plane: 0.000x + 1.000y + 0.000z + 0.000 = 0 P
 Init Conds: 5.0000 0.0000
 7.0000 0.0000

$$x' = \begin{pmatrix} lx^1 - (x_1)^3 \\ x_2 \end{pmatrix}$$



8.7 Ergänzungen

8.7.1 Bahn eines Satelliten im Feld Erde-Mond

Zweikörperproblem: Es seien im \mathbb{R}^2 x^1 und x^2 die Lagen zweier Körper (Erde, Mond) mit den Massen m_1 und m_2 . Dann ist nach NEWTON das Gravitationsfeld durch

$$g(x) = -G \sum_{k=1}^2 m_k \frac{x - x^k}{|x - x^k|^3} \quad (1)$$

gegeben. Dabei ist G die Gravitationskonstante. Es folgen die Bewegungsgleichungen

$$(x^j)'' = g(x^j), \quad (2)$$

also

$$(x^1)'' = -G m_2 \frac{x^1 - x^2}{|x^1 - x^2|^3} \quad (3)$$

$$(x^2)'' = -G m_1 \frac{x^2 - x^1}{|x^2 - x^1|^3}. \quad (4)$$

Wählt man die relativen Koordinaten

$$x := x^2 - x^1 \quad \text{und} \quad \xi := \frac{m_1 x^1 + m_2 x^2}{m_1 + m_2},$$

dann erhält man mit $M := m_1 + m_2$

$$x'' = -G M \frac{x}{|x|^3} \quad (5)$$

$$\xi'' = 0. \quad (6)$$

Der Schwerpunkt ξ bewegt sich also mit konstanter Geschwindigkeit. Wir wählen im folgenden $\xi = 0$. Dann ist mit $\alpha := m_1/M$ und $\beta := m_2/M$

$$x^1 = -\beta x \quad \text{und} \quad x^2 = \alpha x. \quad (7)$$

Aus Gl. (5) folgt

$$x'' x' = -GM \frac{x x'}{|x|^3}$$

oder

$$\frac{(x')^2}{2} = \frac{GM}{|x|} + \text{const.}$$

Damit haben wir ein erstes Integral (die Energie)

$$I_1(x) := \frac{(x')^2}{2} - \frac{GM}{|x|} = c$$

der Bewegung gefunden. Eine zweite Invariante ist das *Impulsmoment*, nämlich

$$I_2(x) := x \times x' = d.$$

Das folgt aus

$$I_2'(x) = x' \times x' + x \times x'' = 0.$$

Mit Hilfe dieser Invarianten läßt sich die Bewegung vollständig beschreiben. Wir verwenden Polarkoordinaten

$$x = r \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}.$$

Dann ist $r^2 \varphi' = d$,

$$x'' = -\frac{GM}{r^2} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

und

$$\frac{d}{d\varphi} x' = \frac{1}{\varphi'} x'' = -\frac{GM}{d} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix},$$

also

$$x' = \frac{GM}{d} \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} + a.$$

Das heißt, es ist

$$(x'_1 - a_1)^2 + (x'_2 - a_2)^2 = \left(\frac{GM}{d}\right)^2.$$

Weiterhin folgt aus

$$x' = r' \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + r\varphi' \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} = \frac{GM}{d} \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} + a$$

durch Elimination von r'

$$r\varphi' = \frac{GM}{d} - a_1 \sin \varphi + a_2 \cos \varphi$$

oder

$$\frac{1}{r} = \frac{GM}{d^2} - \frac{a_1}{d} \sin \varphi + \frac{a_2}{d} \cos \varphi.$$

Das ist die Gleichung eines Kegelschnitts in Polarkoordinaten.

Näherung: Zur Vereinfachung nehmen wir im folgenden $r' = 0$ an, die Bewegung soll also kreisförmig verlaufen. Dann ist $\varphi' = d r^{-2}$ konstant, also

$$\varphi(t) = \frac{d}{r^2} t,$$

und aus $I_1(x) = c$ folgt

$$\frac{d^2}{2r^2} - \frac{GM}{r} = c.$$

Ferner seien die Zeit- und Radiusskalen so gewählt, daß

$$\varphi(t) = t \quad \text{und} \quad r = 1$$

gilt. Dann ist $d = 1$, $GM = 1$, $c = -1/2$, und wir erhalten die Bewegung

$$x(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}.$$

Mitrotierendes Koordinatensystem: Wir wählen nun ein mitrotierendes Koordinatensystem, in dem die beiden Körper ruhend erscheinen, nämlich

$$X := D(t)x := \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} x,$$

also speziell

$$D(t) \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Mit

$$A(t) := D'(t) = \begin{pmatrix} -\sin t & \cos t \\ -\cos t & -\sin t \end{pmatrix}$$

folgt nach leichter Rechnung

$$\begin{aligned} X' &= D x' + AD^{-1}X \\ X'' &= D x'' + 2AD^{-1}X' - (AD^{-1})^2X \\ &= Dx'' + 2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} X' + X. \end{aligned}$$

Bahn des Satelliten: Für einen Satelliten gilt nun analog Gl. (2) die Bewegungsgleichung

$$s'' = -Gm_1 \frac{s - x_1}{|s - x_1|^3} - Gm_2 \frac{s - x_2}{|s - x_2|^3}.$$

Wir nehmen an, daß der Satellit die Bewegung der Erde und des Mondes nicht beeinflußt, und benutzen die angegebene Näherung. Dann folgt mit $S = Ds$ wegen der Invarianz des Absolutbetrages unter Rotationen

$$S'' = 2 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S' + S - \alpha \frac{S - X^1}{|S - X^1|^3} - \beta \frac{S - X^2}{|S - X^2|^3}. \quad (8)$$

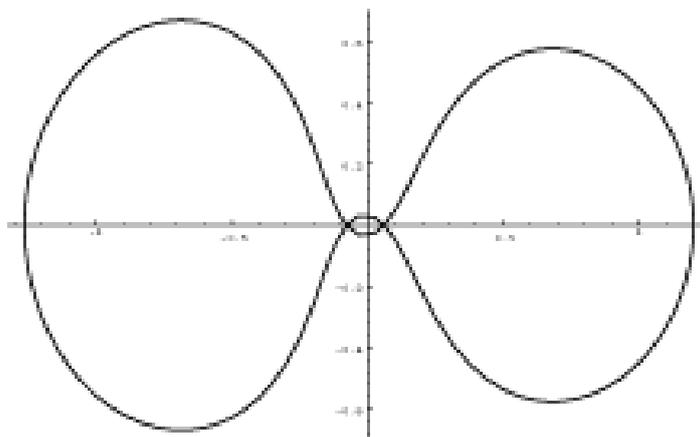
Das ist die *Satellitengleichung*. Darin ist $\beta = 0,01213$ und $\alpha = 1 - \beta$, sowie nach Gl. (7)

$$X^1 = D(t)x^1(t) = -\beta \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad X^2 = D(t)x^2(t) = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Wählen wir die Anfangsbedingungen

$$S_1(0) = 1,2 \quad S'_1(0) = 0 \quad S_2(0) = 0 \quad \text{und} \quad S'_2(0) = -1,04936,$$

dann erhalten wir in der (S_1, S_2) -Ebene folgenden Bahnverlauf:



8.7.2 Rechnung zur Seite 10 oben

$$K(k) := \int_0^{\pi/2} \frac{dt}{\sqrt{1 - k^2 \sin^2 t}} = \int_0^1 \frac{dx}{\sqrt{(1 - x^2)(1 - k^2 x^2)}}.$$

Es sei

$$y := (1 + k') \frac{x \sqrt{1 - x^2}}{\sqrt{1 - k^2 x^2}}.$$

Dann ist

$$\frac{dy}{dx} = (1 + k') \frac{1 - 2x^2 + k^2 x^4}{\sqrt{(1 - x^2)(1 - k^2 x^2)}(1 - k^2 x^2)}$$

und

$$\left(\frac{1 - 2x^2 + k^2 x^4}{1 - k^2 x^2} \right)^2 = (1 - y^2)(1 - k^2 y^2).$$

Es sei $x_0 = \sqrt{1/(1 + k')}$ die Nullstelle von

$$1 - 2x^2 + k^2 x^4 = 0$$

in $(0, 1)$. Dann ist $y_0 = 1$ und

$$K(k) = \int_0^1 \dots = \int_0^{x_0} \dots + \int_{x_0}^1 \dots = \frac{2}{1+k'} \int_0^{y_0} \frac{dy}{\sqrt{(1-y^2)(1-\bar{k}^2 y^2)}},$$

also

$$K(k) = \frac{2}{1+k'} K(\bar{k}).$$

Nun ist mit $\bar{k}' := \sqrt{1-\bar{k}^2} = 2\sqrt{k'}/(1+k')$

$$\begin{aligned} K(\bar{k}) &= \int_0^{\pi/2} \frac{dt}{\sqrt{1-\bar{k}^2 \sin^2 t}} = \int_0^{\pi/2} \frac{dt}{\sqrt{\cos^2 t + (\bar{k}')^2 \sin^2 t}} \\ &= \frac{1+k'}{2} \int_0^{\pi/2} \frac{dt}{\sqrt{(\frac{1+k'}{2})^2 \cos^2 t + k' \sin^2 t}}. \end{aligned}$$

Es folgt also

$$K(k) = \int_0^{\pi/2} \frac{dt}{\sqrt{\cos^2 t + k'^2 \sin^2 t}} = \int_0^{\pi/2} \frac{dt}{\sqrt{(\frac{1+k'}{2})^2 \cos^2 t + k' \sin^2 t}}.$$

8.7.3 Zur Nagumobedingung

Es sei $y(t)$ die Differenz zweier Lösungen, es sei f stetig, und es erfülle f die Nagumobedingung. Dann verschwindet y in einer Umgebung des Nullpunktes; das soll gezeigt werden. Dazu bilden wir

$$g(t) := \left| \frac{y(t)}{t} \right|.$$

Wegen $g(t) = |y'(\theta t)|$, $0 < \theta < 1$, ist g in J stetig mit $g(0) = 0$ und nimmt dort das Maximum M an. Es sei nun $M > 0$ und

$$t_0 = \inf \{t \in J \mid g(t) = M\}.$$

Dann ist $t_0 > 0$, $g(t_0) = M$, und aus der Nagumobedingung und dem Mittelwertsatz der Integralrechnung folgt der Widerspruch

$$M = g(t_0) = \frac{1}{t_0} \left| \int_0^{t_0} y'(s) ds \right| \leq \frac{1}{t_0} \int_0^{t_0} g(s) ds = g(\theta t_0) < M.$$

8.7.4 Der RLC-Kreis

Die Gleichung für einen elektrischen Oszillator lautet

$$y'' + \frac{R}{L} y' + \frac{1}{LC} y = f$$

mit

$$f(t) = \frac{k}{L} e^{i\omega t}.$$

Als erstes wollen wir eine Lösung φ der homogenen Gleichung mit $\varphi(0) = 0$ und $\varphi'(0) = 1$ bestimmen. Dazu machen wir den Ansatz $\varphi(t) = Ae^{\lambda t}$ und erhalten für λ die charakteristische Gleichung

$$0 = \lambda^2 + \frac{R}{L} \lambda + \frac{1}{LC} = (\lambda - \lambda_1) \cdot (\lambda - \lambda_2).$$

Mit den Abkürzungen

$$\alpha := \frac{R}{2L} \quad \text{und} \quad \beta := \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}$$

ist

$$\lambda_1 = -\alpha + \beta \quad \text{und} \quad \lambda_2 = -\alpha - \beta.$$

Daraus folgt

$$\varphi(t) = \frac{1}{2\beta} e^{-\alpha t} \{e^{\beta t} - e^{-\beta t}\}.$$

Eine Lösung y der inhomogenen Gleichung mit $y(0) = y'(0) = 0$ erhält man dann gemäß

$$y(t) = \int_0^t \varphi(t-\tau) f(\tau) d\tau.$$

Mit $\gamma := -\alpha \pm \beta$ sind also Integrale der Form

$$I(t) = \int_0^t e^{\gamma(t-\tau)} e^{i\omega\tau} d\tau = e^{\gamma t} \int_0^t e^{(i\omega-\gamma)\tau} d\tau$$

zu berechnen. Hier sind zwei Fälle möglich:

1. Es sei $i\omega \neq \gamma$. Dann ist

$$I(t) = \frac{1}{i\omega - \gamma} \{e^{i\omega t} - e^{\gamma t}\}.$$

2. Es sei $i\omega = \gamma$. Dann ist

$$I(t) = te^{\gamma t}.$$

Wir wollen nun die Lösungen der inhomogenen Gleichung diskutieren. Als erstes sei $R \neq 0$. Dann tritt Fall 1 auf, und wir erhalten

$$y_\omega(t) = \frac{k}{2\beta L} \left\{ \frac{e^{i\omega t} - e^{(-\alpha+\beta)t}}{i\omega + \alpha - \beta} - \frac{e^{i\omega t} - e^{(-\alpha-\beta)t}}{i\omega + \alpha + \beta} \right\}.$$

Zweitens sei $R = 0$. Dann ist $\alpha = 0$ und $\beta =: i\omega_0$. Es sei zunächst $\omega \neq \omega_0$. Dann ist

$$y_\omega(t) = \frac{k}{2L\omega_0} \left\{ -\frac{e^{i\omega t} - e^{i\omega_0 t}}{\omega - \omega_0} + \frac{e^{i\omega t} - e^{-i\omega_0 t}}{\omega + \omega_0} \right\}.$$

Es sei schließlich $R = 0$ und $\omega = \omega_0$, also ω gleich der Resonanzfrequenz ω_0 . Dann tritt Fall 2 auf, und wir erhalten

$$y_0(t) = \frac{ik}{2\omega_0 L} \left\{ -te^{i\omega_0 t} + \frac{1}{\omega_0} \sin \omega_0 t \right\}.$$

Man sieht sofort, daß bei festem t

$$\lim_{\omega \rightarrow \omega_0} y_\omega(t) = y_0(t)$$

ist. Physikalisch können wir das als Einschwingvorgang deuten; $y_0(t)$ divergiert für große t .

9 Funktionen von mehreren Veränderlichen

In diesem Kapitel wollen wir Abbildungen

$$f : D(f) \subset X \longrightarrow Y$$

behandeln. Dabei sind $D(f)$ der Definitionsbereich von f und X, Y metrische Räume. Insbesondere sollen uns hier die Fälle

$$X = \mathbb{R}^n \quad \text{und} \quad Y = \mathbb{R}^m$$

interessieren.

9.1 Stetige Funktionen

Am Ende von §5.1 haben wir bereits die Stetigkeit von Abbildungen in beliebigen metrischen Räumen definiert. Ich wiederhole:

Definition 9.1.1: f heißt stetig in $x_0 \in X$: \iff

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X, d_X(x, x_0) < \delta \quad d_Y(f(x), f(x_0)) < \varepsilon$$

und

Definition 9.1.2:

$$f \in C(X, Y) \quad \iff \quad \forall x \in X \text{ ist } f \text{ in } x \text{ stetig.}$$

Den folgenden Satz haben wir bewiesen (Satz 5.1.20):

Satz 9.1.3: $f : X \rightarrow Y$ ist stetig \iff

$$\forall V \subset Y, V \text{ offen, ist } f^{-1}(V) \subset X \text{ offen.}$$

Satz 9.1.3 läßt sich auch in seiner dualen Form aussprechen, nämlich

Satz 9.1.4: $f : X \rightarrow Y$ ist stetig \iff

$$\forall A \subset Y, A \text{ abgeschlossen, ist } f^{-1}(A) \subset X \text{ abgeschlossen.}$$

Zum Beweis von Satz 9.1.4 nutzt man $\complement f^{-1}(A) = f^{-1}(\complement A)$ aus, also $X = f^{-1}(A) \cup f^{-1}(\complement A)$.

Ich erinnere auch an die Definition des Trägers von f :

Definition 9.1.5: Der Träger (Support) von f ist

$$\text{supp } f := \overline{\{x \in D(f) \mid f(x) \neq 0\}}.$$

Ist $D(f) \neq X$, dann wird $C(D(f), Y)$ analog $C(X, Y)$ definiert. Will man Satz 9.1.3 verwenden, dann muß man aber berücksichtigen, daß der zugrundeliegende metrische Raum nicht mehr X , sondern $D(f)$ ist. Man muß also die Begriffe offen oder abgeschlossen auf $D(f)$ beziehen. Dafür sagt man, $f^{-1}(V)$ sei *relativ $D(f)$ offen*, genauer

Definition 9.1.6: Es sei $A \subset B \subset X$. Dann sagt man:

A ist abgeschlossen relativ B : \iff $A = \bar{A} \cap B$,

A ist offen relativ B : \iff $B \setminus A$ ist abgeschlossen relativ B .

Beispiele: Jedes B ist relativ B offen und abgeschlossen. Es ist

1. $(0, 1]$ abgeschlossen relativ \mathbb{R}^+ .
2. $[0, 1)$ offen relativ \mathbb{R}_0^+ .

Auch jetzt gilt wie im Falle $B = X$

Lemma 9.1.7: A ist relativ B abgeschlossen \iff $B \setminus A$ ist relativ B offen.

Beweis: Es sei $C := B \setminus A$. Dann sind A , C disjunkt und $B = A \cup C$, also $B \setminus C = A$. Damit folgt:

$\Rightarrow A$ abgeschlossen $\rightsquigarrow A = B \setminus C$ abgeschlossen $\rightsquigarrow C$ offen.

$\Leftarrow C$ offen $\rightsquigarrow B \setminus C = A$ abgeschlossen.

Es sei hervorgehoben, daß in der Definition der Stetigkeit das „ $\forall x \in X, d_X(x, x_0) < \delta$ “ wichtig ist, und zwar mit der Betonung des „für alle“. Das gilt natürlich schon im \mathbb{R}^1 . Denken Sie an Sprungfunktionen, die an der Sprungstelle unstetig sind; die einseitige Stetigkeit läßt sich aber noch definieren. Dieses Phänomen tritt natürlich in höher-dimensionalen Räumen deutlicher hervor. Man spricht im \mathbb{R}^n von *partieller Stetigkeit*, wenn die Funktion nur bezüglich einer Variablen stetig ist, wobei die anderen festgehalten werden. Ein Beispiel möge das erläutern.

Beispiel 9.1.8: Es seien $X = \mathbb{R}^2$, $Y = \mathbb{R}$ und

$$f(x, y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{für } x^2 + y^2 > 0 \\ 0 & \text{für } x = y = 0. \end{cases}$$

Wir zeigen

1. In $(x_0, y_0) \neq (0, 0)$ ist f stetig: Es sei $r := \sqrt{x^2 + y^2}$ und $r_0 := \sqrt{x_0^2 + y_0^2}$ mit

$$\frac{r_0}{2} \leq r \leq 2r_0.$$

Dann ist

$$\begin{aligned} |f(x, y) - f(x_0, y_0)| &= \left| \frac{(x-x_0)y}{r^2} + \frac{(y-y_0)x_0}{r_0^2} + x_0y \left(\frac{1}{r^2} - \frac{1}{r_0^2} \right) \right| \\ &\leq \frac{1}{r_0} (2|x-x_0| + |y-y_0|) + \frac{|r_0^2 - r^2|}{rr_0} \\ &\leq \frac{1}{r_0} (2|x-x_0| + |y-y_0|) + \frac{2}{r_0^2} (|x-x_0||x+x_0| + |y-y_0||y+y_0|) \\ &\leq \frac{8}{r_0} (|x-x_0| + |y-y_0|) \leq \frac{8\sqrt{2}}{r_0} \sqrt{|x-x_0|^2 + |y-y_0|^2}. \end{aligned}$$

Das heißt, daß f in (x_0, y_0) stetig ist.

2. Der Nullpunkt ist kritisch. Halten wir zunächst y fest und betrachten f als Funktion nur von x . Für $y \neq 0$ ist f dann stetig. Für $y = 0$ ist $f(x, 0) = 0$, also ebenfalls stetig. Analoges gilt, wenn man x festhält. Mithin ist f partiell stetig. f ist jedoch im Nullpunkt nicht stetig, denn es gibt Folgen $(x_n, y_n) \rightarrow (0, 0)$, für die $f(x_n, y_n)$ nicht gegen Null konvergiert. Es sei etwa $y = ax$ mit $a \neq 0$. Dann ist

$$f(x, ax) = \begin{cases} \frac{a}{1+a^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Mithin gilt

$$\exists \varepsilon := \frac{a}{1+a^2} \quad \forall x \neq 0, x \rightarrow 0 \quad |f(x, ax)| \geq \varepsilon,$$

und f ist in $(0, 0)$ unstetig.

Wir wollen nun einige Eigenschaften stetiger Funktionen herausstellen und beginnen mit folgendem Lemma (dabei denken wir an $B := D(f)$):

Lemma 9.1.9: Es sei $B \subset X$ kompakt. Dann ist B auch relativ zu sich selbst kompakt.

Beweis: Es sei wieder $A \subset B \subset X$.

1. A sei relativ B abgeschlossen. Dann gibt es ein abgeschlossenes C mit $A = C \cap B$. Man wähle nur $C := \bar{A}$.

2. A sei relativ B offen. Dann gibt es ein offenes D mit $A = D \cap B$. Es gilt nämlich

$$\forall a \in A \quad \exists r_a > 0 : \quad \forall b \in B \text{ mit } d(b, a) < r_a \text{ ist } b \in A.$$

Es seien $V_a := \{x \in X \mid d(x, a) < r_a\}$ und

$$D := \bigcup_{a \in A} V_a.$$

Dann leistet D das Gewünschte.

3. Es sei nun $\{U_\lambda \mid \lambda \in \Lambda\}$ eine relativ B offene Überdeckung von B . Dann gibt es zu jedem U_λ ein offenes U_λ^* mit $U_\lambda = U_\lambda^* \cap B$, und die $\{U_\lambda^*\}$ überdecken B . Mithin genügen endlich viele dieser U_λ^* zur Überdeckung, etwa $U_{\lambda_1}^*, \dots, U_{\lambda_n}^*$. Dann überdecken aber auch bereits $U_{\lambda_1}, \dots, U_{\lambda_n}$ die Menge B . Das war zu zeigen.

Es sei nun $f \in C(D, Y)$ mit $D = D(f) \subset X$. Dann gilt analog Satz 5.1.11

Satz 9.1.10: *Es sei D kompakt. Dann ist auch $R(f)$ kompakt.*

Beweis: Es sei \mathcal{V} eine offene Überdeckung von $R(f)$. Dann ist für $V \in \mathcal{V}$

$$f^{-1}(V) \subset D \text{ relativ offen,}$$

und weil D auch relativ kompakt ist, genügen endlich viele der $\{f^{-1}(V) \mid V \in \mathcal{V}\}$ zur Überdeckung von D . Wegen $f(f^{-1}(V)) = V$ überdecken dann aber auch endlich viele der $\{V \in \mathcal{V}\}$ die Menge $R(f)$, das heißt, diese ist kompakt.

Folgerung 9.1.11: *Es seien D kompakt und $Y = \mathbb{R}^m$. Dann existieren das Maximum und das Minimum von $|f|$.*

Eine kompakte Menge ist nämlich beschränkt und abgeschlossen (Satz 4.5.2). Mithin werden

$$\sup_{x \in D} |f(x)| \quad \text{und} \quad \inf_{x \in D} |f(x)|$$

angenommen.

Es seien nun $D \subset \mathbb{R}^n$ und $Y = \mathbb{R}^m$. Dann ist $C(D, \mathbb{R}^m)$ wieder ein Vektorraum. Auch das Produkt und die Verknüpfung stetiger Funktionen sind stetig, wenn diese Operationen möglich sind. Wie im \mathbb{R}^1 geben wir die folgende

Definition 9.1.12: $f : D \rightarrow \mathbb{R}^m$ heißt *gleichmäßig stetig* : \iff

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, x_0 \in D, |x - x_0| < \delta \quad |f(x) - f(x_0)| < \varepsilon.$$

Es gilt wieder

Satz 9.1.13: *Es seien $K \subseteq \mathbb{R}^n$ und $f \in C(K, \mathbb{R}^m)$. Dann ist f gleichmäßig stetig.*

Der Beweis ist wörtlich derselbe wie der zu Satz 5.1.17. Eine gleichmäßig stetige Abbildung $f : D \rightarrow \mathbb{R}^m$ kann auch stetig auf \bar{D} fortgesetzt werden (Satz 5.1.15).

Wir möchten ein Analogon zum Zwischenwertsatz zeigen. Dazu benötigen wir den Begriff des *Zusammenhangs*.

Definition 9.1.14: *Ein metrischer Raum X heißt nichtzusammenhängend* : \iff

$$\exists A \subsetneq X \text{ mit } A \text{ ist offen und abgeschlossen.}$$

Dazu äquivalent ist

$$X = A \cup B$$

mit $A, B \neq \emptyset, A \cap B = \emptyset$ und A, B sind abgeschlossen (oder offen).

X heißt nun *zusammenhängend*, wenn X nicht nichtzusammenhängend ist. Damit folgt der

Zwischenwertsatz: *Es seien X ein zusammenhängender metrischer Raum, Y ein metrischer Raum und $f \in C(X, Y)$. Dann ist $R(f)$ zusammenhängend.*

Wir beweisen indirekt: Es seien $A \neq \emptyset, A \subsetneq R(f)$ und A sowohl offen als auch abgeschlossen. Dann ist

$$R(f) = A \cup B,$$

und auch $B := R(f) \setminus A$ ist relativ $R(f)$ offen und abgeschlossen. Folglich sind $f^{-1}(A)$ und $f^{-1}(B)$ offen und abgeschlossen in X . Es gilt $f^{-1}(A) \neq \emptyset, f^{-1}(B) \neq \emptyset, f^{-1}(A) \cap f^{-1}(B) = \emptyset$ (denn f ist eindeutig) und $X = f^{-1}(A) \cup f^{-1}(B)$. Also wäre X nichtzusammenhängend.

Bemerkung: Oft definiert man zusammenhängende Mengen im \mathbb{R}^n als *wegzusammenhängend*. Ich möchte das erläutern:

Definition 9.1.15: Eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt *Weg*. Man sagt, die Punkte $\gamma(a)$ und $\gamma(b)$ seien durch den Weg γ stetig verbunden.

Definition 9.1.16: Eine Teilmenge $M \subset \mathbb{R}^n$ heißt *wegzusammenhängend*, wenn man je zwei ihrer Punkte durch einen in M verlaufenden Weg stetig miteinander verbinden kann.

Ein wegzusammenhängender Raum ist zusammenhängend. Die Umkehrung gilt aber nur unter zusätzlichen Annahmen. Das Beispiel

$$\{(x, y) \mid y = \sin(1/x), 0 < x \leq 1\} \cup \{(0, y) \mid -1 < y < 1\}$$

zeigt dies.

Wir wollen nun noch den Fall

$$f : D \rightarrow \mathbb{R}^m \quad \text{injektiv und stetig}$$

betrachten. Es existiert dann

$$f^{-1} : R(f) \rightarrow \mathbb{R}^n.$$

Im \mathbb{R}^1 haben wir für offene D bewiesen, daß f^{-1} stetig ist (Satz 5.1.19). Die entsprechende Aussage gilt auch für beliebige $m = n$. Zum Beweis muß man für eine Kugel $B \subset D \subset \mathbb{R}^n$ zeigen, daß dann auch $f(B) \subset \mathbb{R}^n$ offen ist. Dazu fehlen uns aber die topologischen Hilfsmittel. Wir zeigen hier nur

Satz 9.1.17: Es seien $K \in \mathbb{R}^n$ und $f \in C(K, \mathbb{R}^m)$ injektiv. Dann ist $f^{-1} : R(f) \rightarrow K$ stetig.

Der Beweis unterscheidet sich nicht vom ersten Teil des Beweises zu Satz 5.1.19. Es seien $x_0 \in K$, $y_0 := f(x_0)$ und f^{-1} in y_0 unstetig. Dann gilt

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists y, |y - y_0| < \delta \quad |f^{-1}(y) - f^{-1}(y_0)| \geq \varepsilon.$$

Zu jedem $\delta = 1/n$ gibt es daher ein $y_n \in R(f)$ mit $|y_n - y_0| < 1/n$ und $|f^{-1}(y_n) - f^{-1}(y_0)| \geq \varepsilon$. Es seien $x_n := f^{-1}(y_n) \in K$. Dann existieren eine Teilfolge von (x_n) — es sei (x_n) bereits diese Teilfolge — und ein $x \in K$ mit $x_n \rightarrow x$ (Satz 4.5.6). Wegen $|x_n - x_0| \geq \varepsilon$ ist $x \neq x_0$. Aus der Stetigkeit von f folgt aber

$$f(x) = \lim_{x_n \rightarrow x} f(x_n) = \lim_{n \rightarrow \infty} y_n = y_0 = f(x_0).$$

Damit ist der Satz bewiesen.

9.2 Lineare Abbildungen

In diesem Abschnitt behandeln wir lineare Abbildungen. Es seien X, Y normierte \mathbb{K} -Vektorräume und A eine lineare Abbildung, also

$$A : X \rightarrow Y$$

mit

$$\forall x_1, x_2 \in X \quad \forall \alpha_1, \alpha_2 \in \mathbb{K} \quad A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2.$$

Dafür schreiben wir $A \in \mathcal{L}(X, Y)$. Ist A linear, dann gilt auch $A^{-1} \in \mathcal{L}(R(A), X)$, sofern diese Abbildung existiert. Speziell heißt $A \in \mathcal{L}(X, \mathbb{K})$ *lineares Funktional*.

Eine wichtige Teilklasse von $\mathcal{L}(X, Y)$ bilden nun die beschränkt linearen Abbildungen, nämlich

Definition 9.2.1:

$$A \in \mathcal{L}_b(X, Y) \quad :\iff \quad \sup_{\|x\|_X=1} \|Ax\|_Y \text{ existiert.}$$

Mit

$$\|A\|_{\mathcal{L}_b} := \sup_{\|x\|_X=1} \|Ax\|_Y$$

wird $\mathcal{L}_b(X, Y)$ selbst zu einem normierten Vektorraum.

Natürlich gibt es unbeschränkt lineare Abbildungen. Gerade die Differentialoperatoren sind unbeschränkt. Denken Sie zum Beispiel an $X = C_1([0, 1], \mathbb{R})$, versehen mit der Supremumsnorm. Es seien $f_n(x) := x^n$ und $A \in \mathcal{L}(X, C([0, 1], \mathbb{R}))$, $Ax := x'$. Dann ist

$$\|f_n\| = 1 \quad \text{und} \quad \|A f_n\| = n.$$

A ist also unbeschränkt.

Offenbar sind die beschränkt linearen Abbildungen stetig. Wegen der Linearität gilt sogar

Satz 9.2.2: Es sei $A \in \mathcal{L}(X, Y)$. Dann gilt:

1. A ist entweder in allen $x \in X$ stetig oder nirgends.
2. $C(X, Y) \cap \mathcal{L}(X, Y) = \mathcal{L}_b(X, Y)$.

Beweis:

1. A sei in $x_0 \in X$ stetig, das heißt

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, \|x - x_0\| < \delta \quad \|Ax - Ax_0\| < \varepsilon.$$

Es sei $x_1 \in X$ beliebig und $x \in B(x_1, \delta)$. Dann gilt

$$\|Ax - Ax_1\| = \|A(x - x_1 + x_0) - Ax_0\| < \varepsilon$$

wegen $\|(x - x_1 + x_0) - x_0\| = \|x - x_1\| < \delta$. A ist also auch in x_1 stetig.

2. Aus $A \in \mathcal{L}_b(X, Y)$ folgt $\|Ax\| \leq c \|x\|$. Mithin ist A im Nullpunkt stetig, also $A \in C(X, Y)$.

Umgekehrt, es sei A im Nullpunkt stetig. Dann gibt es ein $\delta = \delta(1)$ mit

$$\forall x, \|x\| < \delta \quad \|Ax\| < 1.$$

Es sei $x \neq 0$ und

$$x_0 := \frac{\delta}{2\|x\|}x.$$

Dann ist $\|x_0\| < \delta$, $\|Ax_0\| < 1$, und es folgt

$$\|Ax\| = \frac{2\|x\|}{\delta} \|Ax_0\| < \frac{2}{\delta} \|x\|.$$

Das heißt, es ist $\|A\| < 2/\delta$.

Eine wichtige Klasse nicht notwendig stetiger linearer Abbildungen ist die Klasse der *abgeschlossenen* linearen Abbildungen. Um sie zu definieren, erinnere ich an den Graphen

$$G(A) := \{(x, Ax) \mid x \in D\} \subset X \times Y$$

einer Abbildung $A : D \subset X \rightarrow Y$.

Definition 9.2.3: $A \in \mathcal{L}(D, Y)$ heißt *abgeschlossen*, wenn $G(A)$ in $X \times Y$ abgeschlossen ist.

A ist also abgeschlossen, wenn aus $x_n \rightarrow x$ und $Ax_n \rightarrow y$ auch $x \in D$ und $Ax = y$ folgt. A heißt *abschließbar*, wenn ein linearer Operator B mit $G(B) = \overline{G(A)}$ existiert. $\bar{A} := B$ heißt dann *Abschluß* von A . In den Anwendungen ist es ein wichtiges Problem zu prüfen, ob ein gegebener Operator abschließbar ist. Gerade die Differentialoperatoren sind oft abgeschlossen bzw. abschließbar. Ich erwähne nur, daß in vollständigen Räumen X, Y ein abgeschlossener Operator mit $D(A) = X$ auch beschränkt ist (Satz vom abgeschlossenen Graphen). Wir zeigen

Satz 9.2.4: Es sei $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Dann gilt:

1. A ist abgeschlossen.
2. $A \in \mathcal{L}_b(\mathbb{R}^n, \mathbb{R}^m)$.

Zum Beweis wählen wir eine Folge (x_k) mit $x_k \rightarrow x$ und $y_k := Ax_k \rightarrow y$. Es sei $z_k := x_k - x$. Dann ist (z_k) eine Nullfolge, und es gilt $Az_k = y_k - Ax \rightarrow y - Ax$. Es sei

$$z_k := \sum_{i=1}^n \zeta_{i,k} e_i.$$

Dann ist

$$|\zeta_{i,k}| = |(z_k, e_i)| \leq |z_k|$$

und

$$|Az_k| \leq \sum_{i=1}^n |\zeta_{i,k}| \cdot |Ae_i| \leq c |z_k|$$

mit

$$c := n \cdot \max_{1 \leq i \leq n} |Ae_i|.$$

Mithin gilt auch $Az_k \rightarrow 0$, und es ist $y = Ax$. Die zweite Behauptung folgt analog aus

$$\|A\| \leq n \cdot \max_{1 \leq i \leq n} |Ae_i|.$$

Damit haben wir gleichzeitig für diesen endlich-dimensionalen Fall den Satz vom abgeschlossenen Graphen bewiesen.

Im Beweis des letzten Satzes haben wir das Skalarprodukt im \mathbb{R}^n benutzt. Es ist

$$|x| = \sqrt{(x, x)}$$

und

$$(x, y) = \sum_{i=1}^n \xi_i \eta_i \quad \text{mit } x = \sum_{i=1}^n \xi_i e_i \quad \text{und } y = \sum_{i=1}^n \eta_i e_i.$$

Dieses Skalarprodukt werden wir im folgenden häufig verwenden. Ich möchte daher an dieser Stelle drei leicht beweisbare und allgemein gültige Resultate aus der Theorie der Hilberträume bereitstellen. Ein Hilbertraum ist ein vollständiger Vektorraum mit dem Skalarprodukt (\cdot, \cdot) und der davon abgeleiteten Norm $\|\cdot\|$ (vgl. §3.4).

Wir beginnen mit dem

Approximationssatz: *Es seien M ein abgeschlossener Teilraum des Hilbertraumes H , $h \in H$ und*

$$d := \inf_{m \in M} \|m - h\|$$

der Abstand von h zu M . Dann gilt:

$$\exists m \in M \quad \|m - h\| = d.$$

Es gibt also genau ein Element bester Approximation $m \in M$. Zum Beweis des Satzes wählen wir eine *Minimalfolge* (m_n) mit

$$\lim_{n \rightarrow \infty} \|m_n - h\| = d.$$

Das Problem ist zu zeigen, daß (m_n) — oder eine Teilfolge davon — konvergiert. Nun gilt die *Parallelogrammgleichung*

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Aus ihr folgt mit $u := h - m_n, v := h - m_l$

$$4 \underbrace{\left\| h - \frac{m_n + m_l}{2} \right\|^2}_{=d^2 + p_{nl}, p_{nl} \geq 0} + \|m_n - m_l\|^2 = 2(\|h - m_n\|^2 + \|h - m_l\|^2) \rightarrow 4d^2.$$

Für $l, n \rightarrow \infty$ gilt also

$$\|m_n - m_l\| \rightarrow 0.$$

Das heißt, (m_n) ist eine Cauchyfolge. H ist vollständig, und M ist abgeschlossen. Mithin gibt es ein $m \in M$ mit $m_n \rightarrow m$.

m ist eindeutig bestimmt. Es sei nämlich (z_n) eine andere Minimalfolge. Dann folgt aus der Parallelogrammgleichung mit $u := h - m_n$ und $v := h - z_n$

$$\|m_n - z_n\| \rightarrow 0.$$

Mithin konvergiert (z_n) ebenfalls gegen m .

Bemerkung 9.2.5: *Es sei bemerkt, daß die Voraussetzung, M sei ein Teilraum von H , unnötig ist. Es genügt, daß M eine abgeschlossene konvexe Teilmenge ist. Für eine konvexe Teilmenge gilt*

$$\forall s \in [0, 1] \quad \forall m_1, m_2 \in M \quad s m_1 + (1 - s) m_2 \in M.$$

Mit $m_n, m_l \in M$ ist dann

$$\frac{m_n + m_l}{2} \in M,$$

und mehr wurde nicht benötigt.

Aus dem Approximationssatz folgt nun der

Projektionssatz: Es seien M ein abgeschlossener Teilraum von H und

$$M^\perp := \{h \in H \mid \forall m \in M \quad (h, m) = 0\}.$$

Dann gilt

$$\forall h \in H \quad \exists m \in M \quad \exists m^\perp \in M^\perp \quad h = m + m^\perp.$$

Zum Beweis wählen wir m aus dem Approximationssatz. Dann gilt für alle $\lambda \in \mathbb{C}$ und $x \in M$ mit $x \neq 0$

$$\|h - m\|^2 \leq \|h - m - \lambda x\|^2 = \|h - m\|^2 - 2 \operatorname{Re} \bar{\lambda}(h - m, x) + |\lambda|^2 \|x\|^2.$$

Wählt man

$$\lambda := \frac{(h - m, x)}{\|x\|^2},$$

dann folgt daraus

$$0 \leq -|\lambda|^2 \cdot \|x\|^2$$

oder $\lambda = 0$, das heißt

$$h = m + m^\perp \quad \text{mit} \quad m^\perp := h - m \in M^\perp.$$

Diese Zerlegung ist eindeutig. Aus

$$h = m' + m'^\perp$$

folgt nämlich

$$0 = (m - m') + (m^\perp - m'^\perp),$$

also

$$0 = (m - m', m - m') = \|m - m'\|^2.$$

M^\perp ist ebenfalls ein abgeschlossener Teilraum von H .

Vermöge des Projektionssatzes können wir eine lineare Abbildung, die *orthogonale Projektion*

$$\begin{aligned} P : H &\longrightarrow M \\ h &\longmapsto m, \end{aligned}$$

definieren. Aus $(Ph, g) = (Ph, Pg) = (h, Pg)$ folgt ihre Symmetrie und Beschränktheit ($\|P\| = 1$ falls $M \neq \{0\}$). Es ist $P^2 = P$. H selbst wird in eine direkte Summe

$$H = M \oplus M^\perp = R(P) \oplus N(P)$$

zerlegt. Es ist nämlich

$$R(P)^\perp = \{x \mid \forall h \quad (Ph, x) = 0\} = \{x \mid \forall h \quad (h, Px) = 0\} = \{x \mid Px = 0\} = N(P).$$

$Q := id - P$ ist die orthogonale Projektion auf M^\perp .

Als letztes allgemeines Resultat folgern wir noch einen Darstellungssatz, der auf F. RIESZ (1880–1956) zurückgeht.

Rieszscher Darstellungssatz: Es sei $F \in \mathcal{L}_b(H, \mathbb{K})$ ein beschränkt lineares Funktional. Dann gilt:

$$\exists f \in H \quad \forall h \in H \quad Fh = (h, f)$$

und $\|f\| = \|F\|$.

Es sei nämlich $N := \{h \in H \mid Fh = 0\}$. N ist ein abgeschlossener Teilraum von H . Aus $x_n \in N$, $x_n \rightarrow x$ folgt nämlich

$$|Fx| = |Fx_n + F(x - x_n)| \leq \|F\| \cdot |x - x_n| \rightarrow 0.$$

Nun sind zwei Fälle möglich:

1. Es sei $N = H$. Dann wähle man $f := 0$, und dann ist $\|f\| = \|F\|$.
2. Es sei $N \neq H$. Dann ist

$$H = N \oplus N^\perp,$$

und es gibt ein $w \in N^\perp$ mit $w \neq 0$ und $Fw \neq 0$. Es folgt

$$h - w \frac{Fh}{Fw} \in N$$

oder

$$(h, w) = \frac{\|w\|^2}{Fw} Fh.$$

Man wähle deshalb

$$f := \frac{\overline{Fw}}{\|w\|^2} w.$$

Dann ist $Fh = (h, f)$. Es sei auch $(h, f') = Fh$. Dann folgt

$$\forall h \in H \quad (h, f - f') = 0,$$

also $\|f - f'\|^2 = 0$, das heißt $f = f'$. Es bleibt $\|f\| = \|F\|$ zu zeigen. Das folgt aus

$$\|f\| = \sup_{\|w\|=1} \frac{|Fw|}{\|w\|} \leq \|F\|$$

und

$$\|F\| = \sup_{\|h\|=1} |Fh| = \sup_{\|h\|=1} |(h, f)| \leq \|f\|.$$

Damit ist der Darstellungssatz bewiesen.

Wir fahren nun mit der Diskussion linearer Abbildungen

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

fort und wollen Gleichungen der Form

$$Ax = f \tag{*}$$

lösen. $N(A)$ ist abgeschlossen. Das gilt allgemein für abgeschlossene Abbildungen. Wichtig ist nun

Lemma 9.2.6: *Es sei $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Dann ist $R(A)$ abgeschlossen.*

Beweis: Es sei $Ax_n = y_n \rightarrow y$, $N := N(A)$ und

$$\lfloor x \rfloor := \inf_{n \in N} |x + n|$$

der Abstand von x zu N . Dann ist $\lfloor x \rfloor \leq |x|$. Wir wählen $z_j := x_j + n_j$, $n_j \in N$, mit

$$\lfloor x_j \rfloor - \frac{1}{j} \leq |z_j| \leq \lfloor x_j \rfloor + \frac{1}{j}.$$

Dann folgt auch $Az_j = y_j \rightarrow y$. Die Folge $(\lfloor x_j \rfloor)$ ist beschränkt. Es sei nämlich $\lfloor x_j \rfloor \rightarrow \infty$ und

$$v_j := \frac{z_j}{\lfloor x_j \rfloor}.$$

Dann ist (v_j) beschränkt, und es gilt $Av_j = y_j / \lfloor x_j \rfloor \rightarrow 0$. (v_j) enthält eine konvergente Teilfolge; es sei (v_j) bereits diese Teilfolge und $v_j \rightarrow v$. Dann ist $v \in N$, weil A abgeschlossen ist. Dem widerspricht aber

$$1 \leq \frac{1}{\lfloor x_j \rfloor} \left| x_j + \underbrace{(n_j - \lfloor x_j \rfloor v)}_{\in N} \right| = |v_j - v| \rightarrow 0.$$

Also ist $(\lfloor x_j \rfloor)$ beschränkt und damit auch $(|z_j|)$. Mithin gibt es eine Teilfolge von (z_j) – o.B.d.A. wieder (z_j) selbst – mit $z_j \rightarrow z$. Dann ist $Az = y$ und damit $y \in R(A)$. Das war zu zeigen.

Aus diesem Lemma und dem Projektionssatz folgt die Lösungstheorie zur Gl. (*), nämlich

Satz 9.2.7: *Es sei $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. Dann ist*

$$\mathbb{R}^n = R(A) \oplus N(A^t)$$

eine orthogonale Zerlegung.

Dabei ist A^t die zu A transponierte Matrix. Der Beweis folgt aus

$$\mathbb{R}^n = R(A) \oplus R(A)^\perp$$

und

$$R(A)^\perp = \{x \mid \forall y (Ay, x) = 0\} = \{x \mid \forall y (y, A^t x) = 0\} = \{x \mid A^t x = 0\} = N(A^t).$$

Es ist auch $\dim N(A) = \dim N(A^t)$. Damit gilt die *Fredholmsche Alternative* für Gleichungen von der Form (*). Aus Satz 9.2.4 folgt die Stetigkeit von $A^{-1} : R(A) \rightarrow \mathbb{R}^n$.

Als letztes zeigen wir noch folgendes Resultat:

Satz 9.2.8: *Es sei Ω die Menge der bijektiven linearen Abbildungen des \mathbb{R}^n in sich. Dann gilt:*

1. $A \in \Omega, \alpha := \|A^{-1}\|^{-1}, B \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n), \|B - A\| < \alpha \implies B \in \Omega$.
2. $\Omega \subset \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ ist offen bezüglich $\|A\| := \sup_{|x|=1} |Ax|$.
3. $F : \Omega \rightarrow \Omega, A \mapsto A^{-1}$ ist stetig.

Beweis:

1. Es sei $\beta := \|B - A\|$. Dann folgt aus $|x| = |A^{-1}Ax| \leq \frac{1}{\alpha} |Ax|$

$$\begin{aligned} (\alpha - \beta)|x| &\leq |Ax| - \beta|x| = |Ax| - \|B - A\| \cdot |x| \\ &\leq |Ax| - |(B - A)x| \leq |(A + (B - A))x| = |Bx|. \end{aligned}$$

Mithin ist B für $\beta < \alpha$ invertierbar.

2. Das folgt aus 1.
3. Aus 1 folgt

$$\|B^{-1}\| \leq \frac{1}{\alpha - \beta}.$$

Deshalb ist

$$\begin{aligned} \|F(A) - F(B)\| &= \|A^{-1} - B^{-1}\| = \|A^{-1}(B - A)B^{-1}\| \\ &\leq \|A^{-1}\| \cdot \|B - A\| \cdot \|B^{-1}\| \leq \frac{\beta}{\alpha(\alpha - \beta)} \rightarrow 0 \end{aligned}$$

für $\beta = \|A - B\| \rightarrow 0$.

Der Satz gilt analog in Banachräumen für stetige bijektive Abbildungen A mit stetigem A^{-1} .

9.3 Gleichmäßige Konvergenz

Es sei jetzt wieder

$$f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

und $\mathcal{F}(D, \mathbb{R}^m)$ die Menge dieser Abbildungen. \mathcal{F} ist ein Vektorraum. Es sei $\mathcal{B}(D, \mathbb{R}^m)$ der Teilraum der beschränkten Abbildungen, nämlich

$$\mathcal{B}(D, \mathbb{R}^m) := \{f \in \mathcal{F}(D, \mathbb{R}^m) \mid \|f\| < \infty\}.$$

Dabei sei

$$\|f\| := \sup_{x \in D} |f(x)|$$

wieder die Supremumsnorm. $\mathcal{B}(D, \mathbb{R}^m)$ ist ein normierter Vektorraum.

Wie im \mathbb{R}^1 geben wir die

Definition 9.3.1: *Es seien $D \subset \mathbb{R}^n, f_j \in \mathcal{B}(D, \mathbb{R}^m)$ und $f \in \mathcal{B}(D, \mathbb{R}^m)$. Dann heißt die Folge (f_j) gleichmäßig konvergent gegen $f : \iff$*

$$\lim_{j \rightarrow \infty} \|f_j - f\| = 0.$$

Die Cauchy-Konvergenz definiert man analog.

Analog Satz 5.2.4 folgt

Satz 9.3.2: Es seien $f_n \in C(D, \mathbb{R}^m)$, $f \in \mathcal{F}(D, \mathbb{R}^m)$ und

$$\|f_n - f\| \rightarrow 0.$$

Dann gilt auch $f \in C(D, \mathbb{R}^m)$.

Speziell ergibt sich die Vollständigkeit der folgenden Räume bezüglich der Supremumsnorm:

$$\mathcal{B}(D, \mathbb{R}^m), \mathcal{BC}(D, \mathbb{R}^m), C^*(D, \mathbb{R}^m), C(K, \mathbb{R}^m).$$

Dabei ist $K \subseteq \mathbb{R}^n$ und C^* die Menge der gleichmäßig stetigen Funktionen.

Es sei im folgenden $K \subseteq \mathbb{R}^n$, und es seien $f, g \in C(K, \mathbb{R})$. Dann ist auch

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|) \in C(K, \mathbb{R})$$

$$\min(f, g) = \frac{1}{2}(f + g - |f - g|) \in C(K, \mathbb{R}).$$

Analoges gilt für das Maximum oder Minimum endlich vieler Funktionen.

Wir wollen nun in $C(K, \mathbb{R})$ Approximationssätze beweisen. Es sei $\mathcal{V} \subset C(K, \mathbb{R})$, und $f \in C(K, \mathbb{R})$ soll durch $f_j \in \mathcal{V}$ bezüglich der Supremumsnorm approximiert werden. Dann stellt sich die Frage, wie „groß“ \mathcal{V} sein muß. Als erstes zeigen wir den Satz von M. STONE (1903–), nämlich

Satz 9.3.3: Es sei \mathcal{V} eine Teilmenge von $C(K, \mathbb{R})$ mit folgenden Eigenschaften:

1. Mit f, g gehören auch $\max(f, g)$ und $\min(f, g)$ zu \mathcal{V} .
2. $\forall a, b \in \mathbb{R} \quad \forall y, z \in K, y \neq z \quad \exists v \in \mathcal{V} \quad v(y) = a \wedge v(z) = b.$

Dann gilt:

$$\forall f \in C(K, \mathbb{R}) \quad \forall \varepsilon > 0 \quad \exists v \in \mathcal{V} \quad \|v - f\| < \varepsilon.$$

Beweis: Es seien $f \in C(K, \mathbb{R})$ und $\varepsilon > 0$ vorgegeben. Zu $y, z \in K$ und $a := f(y)$, $b := f(z)$ wähle man v_{yz} nach 2. Dann gilt

1. Es sei y fest. Dann folgt aus $v_{yz}(z) = f(z)$

$$\forall z \quad \exists U(z) \quad \forall x \in U(z) \cap K \quad v_{yz}(x) - f(x) > -\varepsilon.$$

Die $U(z)$ überdecken K . Mithin gibt es $z_1, \dots, z_r \in K$, so daß $U(z_1), \dots, U(z_r)$ die kompakte Menge K überdecken. Es sei

$$v_y := \max(v_{yz_1}, \dots, v_{yz_r}) \in \mathcal{V}.$$

Dann ist

$$\forall x \in K \quad v_y(x) - f(x) > -\varepsilon. \quad (*)$$

2. Wir wiederholen diesen Prozeß bezüglich y . Wegen $v_{yz_i}(y) = f(y)$ ist auch $v_y(y) = f(y)$. Daraus folgt

$$\forall y \in K \quad \exists V(y) \quad \forall x \in V(y) \cap K \quad v_y(x) - f(x) < \varepsilon.$$

Wiederum gibt es $y_1, \dots, y_s \in K$, so daß $U(y_1), \dots, U(y_s)$ K überdecken. Es sei

$$g := \min(v_{y_1}, \dots, v_{y_s}) \in \mathcal{V}.$$

Dann ist

$$\forall x \in K \quad g(x) - f(x) < \varepsilon, \quad (**)$$

nach (*) gilt auch

$$\forall x \in K \quad g(x) - f(x) > -\varepsilon.$$

Damit haben wir

$$\forall x \in K \quad |g(x) - f(x)| < \varepsilon$$

bewiesen.

In der Praxis ist die Menge \mathcal{V} oft eine Unteralgebra $\mathcal{A} \subset C(K, \mathbb{R})$ (\mathcal{A} ist Vektorraum und Ring mit Eins). Dafür gilt der

Satz von Stone-Weierstraß: Es sei $\mathcal{A} \subset C(K, \mathbb{R})$ mit

1. $\forall f, g \in \mathcal{A}$ und $\forall c \in \mathbb{R}$ gilt $f + g \in \mathcal{A}$ und $cf \in \mathcal{A}$.
2. $\forall f, g \in \mathcal{A}$ gilt $f \cdot g \in \mathcal{A}$.
3. $1 : x \mapsto 1$ gehört zu \mathcal{A} .
4. $\forall y, z \in K$ mit $y \neq z$ gibt es ein $g \in \mathcal{A}$ mit $g(y) \neq g(z)$.

Dann gilt

$$\forall f \in C(K, \mathbb{R}) \quad \forall \varepsilon > 0 \quad \exists g \in \mathcal{A} \quad \|f - g\| < \varepsilon.$$

Eine Teilmenge von $C(K, \mathbb{R})$ mit den Eigenschaften 1–4 nennt man eine *punktetrennende Unteralgebra* von $C(K, \mathbb{R})$.

Beweis: Es sei $\bar{\mathcal{A}}$ die abgeschlossene Hülle von \mathcal{A} . Wir wollen $\bar{\mathcal{A}} = C(K, \mathbb{R})$ zeigen. Der Beweis erfolgt in vier Schritten:

1. Auch $\bar{\mathcal{A}}$ ist eine punktgetrennende Unteralgebra.
2. Es sei $f \in \bar{\mathcal{A}}$. Dann folgt $|f| \in \bar{\mathcal{A}}$. Das sieht man so: Die Abbildung $t \mapsto |t|$ läßt sich in jedem symmetrisch zum Nullpunkt liegenden kompakten Intervall gleichmäßig durch Polynome approximieren (Satz von Weierstraß). Das heißt

$$\forall c > 0 \quad \forall \varepsilon > 0 \quad \exists P \quad \forall t \in [-c, c] \quad |P(t) - |t|| < \varepsilon.$$

Es sei nun $\|f\| \leq c$. Dann folgt

$$\forall x \in K \quad |P(f(x)) - |f(x)|| < \varepsilon.$$

Ist $f \in \mathcal{A}$, dann ist es auch $P(f)$, also

$$f \in \mathcal{A} \implies |f| \in \bar{\mathcal{A}}.$$

Daraus folgt

$$f \in \bar{\mathcal{A}} \implies |f| \in \bar{\mathcal{A}},$$

denn zu $\varepsilon > 0$ wähle man $g \in \mathcal{A}$ mit $\|f - g\| < \varepsilon$. Dann ist $|g| \in \bar{\mathcal{A}}$ und

$$\||f| - |g|\| \leq \|f - g\| < \varepsilon.$$

3. $\bar{\mathcal{A}}$ besitzt die Eigenschaften 1 und 2 im Satz von Stone: Die erste ist zum Beispiel wegen

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|)$$

erfüllt. Es sei $g \in \bar{\mathcal{A}}$ mit Eigenschaft 4 der Voraussetzungen. Man wähle mit $\lambda, \mu \in \mathbb{R}$

$$v(x) = \lambda + \mu g(x),$$

so daß

$$v(y) = \lambda + \mu g(y) = a$$

$$v(z) = \lambda + \mu g(z) = b$$

ist. Das ist möglich, denn die Determinante des Gleichungssystems ist gerade $g(z) - g(y)$. Also hat $\bar{\mathcal{A}}$ auch die zweite Eigenschaft.

4. Aus dem Satz von Stone folgt dann

$$\forall \varepsilon > 0 \quad \forall f \in C(K, \mathbb{R}) \quad \exists g \in \bar{\mathcal{A}} \quad \|g - f\| < \varepsilon/2.$$

Zusammen mit

$$\forall \varepsilon > 0 \quad \forall g \in \bar{\mathcal{A}} \quad \exists h \in \mathcal{A} \quad \|g - h\| < \varepsilon/2$$

bedeutet das

$$\forall \varepsilon > 0 \quad \forall f \in C(K, \mathbb{R}) \quad \exists h \in \mathcal{A} \quad \|f - h\| < \varepsilon.$$

Das war zu zeigen.

Folgerung 9.3.4: Zu jedem $f \in C(K, \mathbb{R})$ und $\varepsilon > 0$ gibt es ein Polynom P mit

$$\|f - P\| < \varepsilon.$$

Beweis: Es sei \mathcal{A} die Menge der Polynome (der ganzrationalen Funktionen) eingeschränkt auf K . Dann erfüllt \mathcal{A} die Voraussetzungen des Satzes von Stone-Weierstraß. Für die Voraussetzungen 1–3 ist das klar. Wir zeigen 4: Es seien $x = (\xi_1, \dots, \xi_n)$, y, z analog und etwa $\eta_i \neq \xi_i$. Es sei

$$\begin{aligned} g_i &: K \longrightarrow \mathbb{R} \\ x &\longmapsto \xi_i. \end{aligned}$$

Dann ist $g_i \in \mathcal{A}$ und $g_i(y) \neq g_i(z)$.

9.4 Differenzierbare Abbildungen

In §5.4 habe ich schon auf die Schwierigkeiten bei der Übertragung der klassischen Definition der Differenzierbarkeit im \mathbb{R}^1 auf den \mathbb{R}^n oder Banachräume hingewiesen. Die Taylorsche Formel — die weiterhin gelten sollte

$$f(x+h) = f(x) + f'(x)h + r(x, h)|h|$$

mit

$$\lim_{h \rightarrow 0} r(x, h) = 0$$

zeigt aber, wie man vorzugehen hat: Es sei $U \subset \mathbb{R}^n$ offen und

$$f : U \longrightarrow \mathbb{R}^m.$$

Es seien $x \in U$ fest, $V := \{h \in \mathbb{R}^n \mid x+h \in U\}$, $f'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ und

$$r(x, \cdot) : V \longrightarrow \mathbb{R}^m \quad \text{mit} \quad \lim_{h \rightarrow 0} r(x, h) = 0.$$

Dann ist $f'(x)$ die Ableitung von f an der Stelle x . Im \mathbb{R}^1 ist also

$$r(x, h) = \frac{h}{|h|} \left(\frac{f(x+h) - f(x)}{h} - f'(x) \right).$$

Die Ableitung von f ist also eine lineare Abbildung des \mathbb{R}^n in den \mathbb{R}^m ; wir erklären

Definition 9.4.1: Es seien $U \subset \mathbb{R}^n$ offen und $V := \{h \in \mathbb{R}^n \mid x+h \in U\}$. Dann heißt $f : U \rightarrow \mathbb{R}^m$ an der Stelle $x \in U$ differenzierbar $:\Leftrightarrow$

1. Es gibt ein $A(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.
2. Es gibt eine im Nullpunkt stetige Abbildung

$$r(x, \cdot) : V \longrightarrow \mathbb{R}^m \quad \text{mit} \quad r(x, 0) = 0$$

und

$$f(x+h) = f(x) + A(x) \cdot h + r(x, h) \cdot |h|.$$

Definition 9.4.2: f heißt in U differenzierbar, wenn f für alle $x \in U$ differenzierbar ist. Man nennt $A(x)$ auch die — erste — Ableitung von f und schreibt

$$f'(x) := A(x).$$

Als erstes müssen wir zeigen, daß die Ableitung eindeutig definiert wurde. Es seien $A_1(x)$ und $A_2(x)$ Ableitungen. Dann folgt

$$(A_1(x) - A_2(x))h + (r_1(x, h) - r_2(x, h))|h| = 0$$

oder für festes $y \neq 0$ und alle $t \in \mathbb{R}$ mit $t \rightarrow 0$

$$\frac{|(A_1(x) - A_2(x))y|}{|y|} = \frac{|(A_1(x) - A_2(x))ty|}{|ty|} = |r_1(x, ty) - r_2(x, ty)| \xrightarrow{t \rightarrow 0} 0.$$

Es gilt also

$$\forall y \in \mathbb{R}^n \quad A_1(x)y = A_2(x)y.$$

Das heißt, es ist $A_1(x) = A_2(x)$. Das beweist die Eindeutigkeit.

Aus der Definition der Ableitung folgt sofort

Satz 9.4.3: f sei in x differenzierbar. Dann ist f dort stetig.

Im Falle $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$ folgt aus dem Rieszischen Darstellungssatz die Existenz von $f_1(x), \dots, f_n(x)$ mit

$$f'(x) \cdot h = \sum_{i=1}^n f_i(x) h_i,$$

und durch Spezialisieren erhält man

$$f_i(x) = \frac{\partial f}{\partial x_i}(x).$$

Die $\partial f / \partial x_i$ sind die *partiellen Ableitungen* von f (die Ableitungen in Richtung der Koordinatenachsen). Man differenziert also nach x_i und hält dabei die anderen Variablen fest. Zur Abkürzung verwendet man den *Nabla-Operator*

$$\nabla = \begin{pmatrix} \partial_1 \\ \vdots \\ \partial_n \end{pmatrix} \quad \text{mit } \partial_i := \frac{\partial}{\partial x_i}.$$

Das Wort Nabla stammt aus dem Griechischen (lateinisch nablium) und bezeichnet ein antikes Saiteninstrument. Das Symbol ∇ (ein umgekehrtes Delta) soll an eine Harfe erinnern. Es gilt also

$$f'(x) \cdot h = (\nabla f(x), h) = (h, \nabla) f(x) = (h^t \nabla) f(x).$$

∇f nennt man auch den *Gradienten* von f und schreibt $\text{grad } f := \nabla f$ dafür.

Im Falle $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ geht man analog vor. Aus dem Rieszischen Darstellungssatz folgt nun für $k = 1, 2, \dots, m$

$$(f'(x)h) \cdot e_k = \sum_{i=1}^n f_{ki}(x) h_i$$

oder

$$f'(x) \cdot h = \begin{pmatrix} f_{11}(x) & \cdots & f_{1n}(x) \\ \vdots & & \vdots \\ f_{m1}(x) & \cdots & f_{mn}(x) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Durch Spezialisieren erhält man

$$f_{ki}(x) = \frac{\partial}{\partial x_i} f_k(x) \quad \text{mit } f = (f_1, \dots, f_m)^t.$$

Die Matrix

$$J_f(x) := \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots \\ \partial_1 f_m & \cdots & \partial_n f_m \end{pmatrix} (x) = (\partial_1 f, \dots, \partial_n f)$$

bezeichnet man als *Jacobi-Matrix* von f , benannt nach C.G. JACOBI (1804–1851). Es ist also

$$f'(x) \cdot h = J_f(x)h = (h, \nabla) f(x) = (h^t \nabla) f(x).$$

Beispiel 9.4.4: Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f(x) = \begin{pmatrix} (x_1)^2 - (x_2)^2 \\ x_1 x_2 \end{pmatrix}.$$

Dann ist

$$f'(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{pmatrix}.$$

Beispiel 9.4.5: Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} \frac{x_1(x_2)^2}{|x|^2} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Dann ist für $x \neq 0$

$$f'(x) = \frac{1}{|x|^4} \begin{pmatrix} (x_2)^2(- (x_1)^2 + (x_2)^2) \\ 2(x_1)^3 x_2 \end{pmatrix}.$$

Kritisch ist der Punkt $x = 0$. Es sei ausdrücklich darauf hingewiesen, daß – wie bei der Stetigkeit – der Grenzübergang $h \rightarrow 0$ für alle Nullfolgen h durchgeführt werden muß. Das Ergebnis darf nicht von der speziellen Wahl von h abhängen, und insbesondere genügt es nicht, die partiellen Ableitungen auszurechnen.

In unserem Beispiel ist an der Stelle $x = 0$

$$1. f(te_1) = 0 \rightsquigarrow \partial_1 f(0) = 0$$

$$2. f(te_2) = 0 \rightsquigarrow \partial_2 f(0) = 0.$$

3. Wenn $f'(0)$ existieren würde, wäre also $f'(0) = 0$. Das steht aber im Widerspruch zu

$$\frac{1}{2} = \frac{f(t(e_1 + e_2))}{t} \xrightarrow{t \rightarrow 0} f'(0)(e_1 + e_2) = 0.$$

Mithin ist f im Nullpunkt nicht differenzierbar.

Wir geben noch eine leicht abgewandelte Formulierung der Differenzierbarkeit. Es sei wieder

$$f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

und $x \in U$. Wenn f in x differenzierbar ist, dann existiert $f'(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Ist f in U differenzierbar, dann existiert also

$$\begin{aligned} f' : U &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ x &\longmapsto f'(x). \end{aligned} \quad (*)$$

Wir schreiben nun auch

$$f'(x) \cdot h =: f'(x, h),$$

also

$$\begin{aligned} f' : U \times \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x, h) &\longmapsto f'(x) \cdot h. \end{aligned} \quad (**)$$

Dabei ist $f'(\cdot, \cdot)$ linear in der zweiten Variablen.

Die differenzierbaren Abbildungen bilden einen Vektorraum. Für die Menge der stetig differenzierbaren Abbildungen von $U \subset \mathbb{R}^n$ in den \mathbb{R}^m schreibt man kurz $C_1(U, \mathbb{R}^m)$. Für differenzierbare Abbildungen $f, g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$ gilt wieder die *Produktregel*

$$(f \cdot g)' = gf' + fg'$$

und für $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m, g : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ mit $f(U) \subset V$ die *Kettenregel*

$$(g \circ f)'(x) = (g' \circ f)(x) \cdot f'(x).$$

Mit $H := f(x+h) - f(x)$ ist nämlich

$$(g \circ f)(x+h) - (g \circ f)(x) = (g' \circ f)(x) \cdot H + r_g(f(x), H) \cdot |H|.$$

Nun ist

$$H = f'(x)h + r_f(x, h) \cdot |h|,$$

also

$$(g' \circ f)(x) \cdot H = (g' \circ f)(x) \cdot f'(x)h + (g' \circ f)(x) \cdot r_f(x, h)|h|.$$

Daraus erhält man die Behauptung.

Folgerung 9.4.6: Es seien $x : U \subset \mathbb{R} \rightarrow \mathbb{R}^n$ mit $R(x) \subset V$ und $g : V \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar. Dann ist

$$(g(x))' = \nabla g(x) \cdot x' = \sum_{i=1}^n \frac{\partial g}{\partial x_i}(x) \cdot x'_i.$$

Anschauliche Deutung des Gradienten: Es seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und $x = a + h$ mit $h = t h_0$, $t \in \mathbb{R}$ und $|h_0| = 1$, $h_0 \in \mathbb{R}^n$ fest. Es sei ferner

$$F(t) = f(a + t h_0).$$

Dann ist

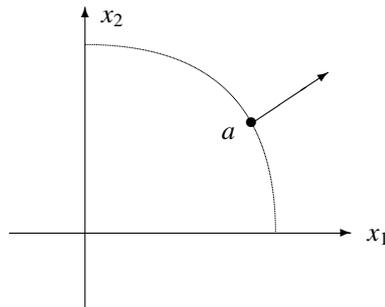
$$F'(0) = \nabla f(a) \cdot h_0$$

die *Richtungsableitung* von f im Punkte a in Richtung h_0 . Der Gradient von f an der Stelle a ist daher ein Vektor in Richtung der stärksten Änderung der Funktion f , und sein Betrag hat den Wert der Ableitung in dieser Richtung.

Beispiel 9.4.7: Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x) = x_1^2 + x_2^2.$$

Dann ist $\nabla f(a) = 2a = 2|a|a_0$. Die stärkste Änderung erfolgt in radialer Richtung, für $h_0 = a_0$, und es ist $|\nabla f(a)| = 2|a|$.



Die Linien $f = \text{const}$ nennt man *Höhenlinien*. Hier im Beispiel sind es Kreise. Die Kurven in Richtung der Gradienten nennt man *Falllinien*.

Es sei $\eta = \eta(\xi)$ eine solche Kurve mit $a_2 = \eta(a_1)$. Dann ist im Beispiel wegen $\nabla f = 2x$

$$\eta' = \frac{\eta}{\xi},$$

also $\eta = k\xi$ oder

$$\eta(\xi) = \frac{a_2}{a_1} \xi \quad \text{für } a_1 \neq 0.$$

Die Falllinien sind hier also Geraden durch den Nullpunkt.

Aus $f = \text{const}$ folgt allgemein für den Anstieg der Höhenlinien

$$\partial_1 f \cdot x'_1 + \partial_2 f \cdot x'_2 = 0,$$

also

$$-\frac{\partial_1 f}{\partial_2 f} = \frac{x'_2}{x'_1} = \frac{dx_2}{dx_1}.$$

Der Anstieg der Falllinien war

$$\eta' = \frac{\partial_2 f}{\partial_1 f}.$$

Mithin stehen Höhenlinien und Falllinien senkrecht aufeinander.

Weitere Beispiele:

1. Es sei $f(x) = (x_1)^2 + 2(x_2)^2$. Dann sind die Höhenlinien Ellipsen, und für die Falllinien gilt

$$\eta' = \frac{2\eta}{\xi},$$

also $\eta = k\xi^2$. Das sind Parabeln.

2. Es sei $f(x) = (x_1)^2 - (x_2)^2$. Dann sind die Höhenlinien Hyperbeln, und für die Falllinien folgt

$$\eta' = -\frac{\eta}{\xi},$$

also $\eta = k/\xi$. Das sind ebenfalls Hyperbeln.

3. Es sei jetzt $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Man nennt $\operatorname{div} f := \nabla^t f = \sum_{i=1}^3 \partial_i f_i$ die Divergenz von f ,

$$\operatorname{rot} f := \nabla \times f := \begin{pmatrix} \partial_2 f_3 - \partial_3 f_2 \\ \partial_3 f_1 - \partial_1 f_3 \\ \partial_1 f_2 - \partial_2 f_1 \end{pmatrix} \quad \text{die Rotation von } f$$

und

$$\Delta := \nabla^t \nabla = \operatorname{div} \operatorname{grad} = \sum_{i=1}^3 (\partial / \partial x_i)^2$$

den Δ - oder Laplace-Operator. $\mathcal{L}^2(\mathbb{R}^3)$ läßt sich orthogonal zerlegen in divergenz- und rotationsfreie (Gradienten-) Felder. Zum Beispiel sei

$$f(x) = \frac{x}{|x|^3}.$$

Dann ist f offenbar ein Gradientenfeld, nämlich

$$f = \nabla \varphi \quad \text{mit } \varphi(x) = -\frac{1}{|x|}.$$

Es ist für $x \neq 0$ $\operatorname{div} f = 0$ und $\operatorname{rot} f = 0$.

Differenzierbarkeit der Umkehrabbildung: Es seien nun wieder $U \subset \mathbb{R}^n$ offen sowie $f : U \rightarrow \mathbb{R}^n$ injektiv und an der Stelle $a \in U$ differenzierbar mit $\det f'(a) \neq 0$. Ferner seien

$$g := f^{-1} : f(U) \longrightarrow \mathbb{R}^n,$$

sowie $b := f(a)$ innerer Punkt von $f(U)$ und g an der Stelle b stetig. Beides ist beweisbar, man vergleiche den Satz von der lokalen Umkehrbarkeit 9.7.1 und Satz 9.1.17. Dann gilt

Satz 9.4.8: *Unter den angegebenen Voraussetzungen ist g an der Stelle b differenzierbar mit*

$$g'(b) = (f'(a))^{-1}.$$

Aus $g \circ f = \operatorname{id}_{\mathbb{R}^n}$ folgt also

$$g'(b) \cdot f'(a) = \operatorname{id}_{\mathbb{R}^n}.$$

Beweis: Es ist

$$\begin{aligned} f(a+x) &= f(a) + f'(a) \cdot x + r(a, x) \cdot |x| \\ &= f(a) + F(a, x) \cdot x \end{aligned}$$

mit

$$F(a, x) := \begin{cases} f'(a) + r(a, x) \cdot x_0 & \text{für } x \neq 0 \\ f'(a) & \text{für } x = 0. \end{cases}$$

Bei festen a, x ist $F(a, x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$. $F(a, x)$ ist an der Stelle $x = 0$ stetig als Funktion von x , und es gilt

$$F(a, 0) = f'(a) \quad \text{mit } \det f'(a) \neq 0.$$

Dann folgt aus der Matrizenrechnung die Existenz einer Umgebung $U(0)$, in der $(F(a, x))^{-1}$ existiert, und in der

$$\lim_{x \rightarrow 0} (F(a, x))^{-1} = (f'(a))^{-1}$$

ist.

Es seien nun

$$f(a+x) =: b+y, \quad a+x = g(b+y)$$

und

$$G(y) := (F(a, x))^{-1}.$$

Dann folgt aus $G(y) \cdot y = G(y) \cdot F(a, x) \cdot x = x$

$$g(b+y) = g(b) + G(y) \cdot y,$$

und wegen der Stetigkeit von g ist

$$\lim_{y \rightarrow 0} G(y) = \lim_{y \rightarrow 0} (F(a, g(b+y) - a))^{-1} = (F(a, 0))^{-1} = (f'(a))^{-1}.$$

Das heißt, g ist an der Stelle b differenzierbar, und es gilt

$$g'(b) = (f'(a))^{-1}.$$

Das war zu beweisen.

Beispiel 9.4.9: Es seien $r \in \mathbb{R}^+$, $\varphi \in (0, 2\pi)$ und

$$f(r, \varphi) = r \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix},$$

also

$$g(x, y) = \begin{pmatrix} \sqrt{x^2 + y^2} \\ \arctan \frac{y}{x} \end{pmatrix}.$$

Dann ist

$$f'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},$$

und

$$g'(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\frac{\sin \varphi}{r} & \frac{\cos \varphi}{r} \end{pmatrix}.$$

Ist f' stetig, dann heißt f stetig differenzierbar, und man schreibt $f \in C_1(U, \mathbb{R}^n)$.

Definition 9.4.10: Es seien $U \in \mathbb{R}^n$ offen, $f \in C_1(U, \mathbb{R}^n)$ invertierbar, $f(U)$ offen und g die Umkehrabbildung $g: f(U) \rightarrow \mathbb{R}^n$ sei ebenfalls stetig differenzierbar. Dann nennt man f (und auch g) einen Diffeomorphismus.

Die Voraussetzung, daß $f(U)$ offen ist, ist entbehrlich. Das zeigt wieder der Satz von der lokalen Umkehrbarkeit 9.7.1. Die Polarkoordinatenabbildung im letzten Beispiel ist ein Diffeomorphismus, etwa

$$f: \{(r, \varphi) \mid 0 < r, 0 < \varphi < 2\pi\} \longrightarrow \mathbb{R}^2 \setminus (\mathbb{R}_0^+ \times \{0\}),$$

also die Abbildung des Streifens $\{(r, \varphi) \mid 0 < r, 0 < \varphi < 2\pi\}$ auf die längs der positiven x -Achse aufgeschnittene Ebene. Hingegen ist

$$f: \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto x^3$$

kein Diffeomorphismus in $U(0)$, denn dort ist $f'(0) = 0$, und $g'(0)$ existiert nicht.

9.5 Der Mittelwertsatz

Die Übertragung des Mittelwertsatzes auf reellwertige Funktionen im \mathbb{R}^n ist möglich, wenn die Verbindungsstrecke Γ der betrachteten Punkte a und b

$$\gamma(t) := a + t(b - a) \quad 0 \leq t \leq 1$$

in der Definitionsmenge liegt.

Satz 9.5.1: Es seien $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar; $a, b \in U$ und $\Gamma \subset U$. Dann gibt es ein $c \in \Gamma$ mit

$$f(b) = f(a) + f'(c) \cdot (b - a).$$

Beweis: Es sei

$$F: [0, 1] \longrightarrow \mathbb{R} \\ t \longmapsto f(\gamma(t)).$$

Dann folgt aus dem Mittelwertsatz im \mathbb{R} die Existenz eines $\theta \in (0, 1)$ mit

$$F(1) = F(0) + F'(\theta).$$

Nun ist

$$F'(t) = f'(\gamma(t)) \cdot \gamma'(t) = f'(\gamma(t)) \cdot (b - a),$$

und die Behauptung folgt mit $c := \gamma(\theta)$.

Bemerkung 9.5.2:

1. Die Verbindungsstrecke Γ von a nach b kann man durch einen beliebigen ganz in U verlaufenden differenzierbaren Weg ersetzen. Dann gilt $f(b) = f(a) + f'(\gamma(\theta), \gamma'(\theta))$.
2. Eine zweite Formulierung des Satzes ist wieder

$$f(x+h) = f(x) + f'(x+\theta h) \cdot h \quad \text{mit } 0 < \theta < 1.$$

3. Wenn f vektorwertig ist, $f: U \rightarrow \mathbb{R}^m$, dann läßt sich der Satz auf jede Komponente von f anwenden; also (f_j sei die j -te Komponente von f)

$$f_j(x+h) = f_j(x) + \nabla f_j(x+\theta_j h) \cdot h.$$

Man beachte, daß θ jetzt von j abhängt.

Es gilt der folgende

Satz 9.5.3: Es seien $U \subset \mathbb{R}^n$ offen, wegzusammenhängend und $f: U \rightarrow \mathbb{R}$ differenzierbar. Dann gilt

$$f' = 0 \iff f = \text{const.}$$

Auch dieses Resultat gilt für vektorwertige Abbildungen. Aus $f = \text{const}$ folgt $f' = 0$. Es seien also $f' = 0$ und $a, b \in U$. Weil U offen ist, kann man a, b durch einen Streckenzug $a = a_0, a_1, \dots, a_k = b$ mit $a_i \in U$ verbinden. Nun wende man den Mittelwertsatz zwischen a_i und a_{i-1} an. Dann folgt $f(a_i) = f(a_{i-1})$ und damit die Behauptung.

Aus dem Mittelwertsatz folgt auch

Satz 9.5.4: Es sei wieder $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Die partiellen Ableitungen von f mögen an der Stelle $a \in U$ existieren und dort stetig sein. Dann ist f in a stetig differenzierbar.

Wir beweisen den Satz für $m = 1$. Zu a, h seien

$$a =: a_0, a_1, \dots, a_n := a + h$$

Zwischenpunkte mit $a_i \in U$ und

$$\begin{aligned} a_1 &:= a_0 + h_1 e_1 \\ a_2 &:= a_1 + h_2 e_2 \\ &\vdots \\ a_n &= a_{n-1} + h_n e_n. \end{aligned}$$

Dann ist

$$f(a+h) - f(a) = (f(a+h) - f(a_{n-1})) + (f(a_{n-1}) - f(a_{n-2})) + \dots + (f(a_1) - f(a)).$$

Wendet man jeweils den Mittelwertsatz in \mathbb{R}^1 an, dann folgt mit geeigneten $c_j \in B(a, |h|)$

$$f(a+h) - f(a) = \sum_{j=1}^n (\partial_j f)(c_j) h_j$$

und

$$\left| f(a+h) - f(a) - \sum_{j=1}^n (\partial_j f)(a) h_j \right| = \left| \sum_{j=1}^n ((\partial_j f)(c_j) - (\partial_j f)(a)) h_j \right| \leq |h| \sum_{j=1}^n |(\partial_j f)(c_j) - (\partial_j f)(a)|,$$

und wegen der Stetigkeit der partiellen Ableitungen konvergiert die Summe gegen Null für $h \rightarrow 0$. Aus

$$f'(a) \cdot h = \sum_{j=1}^n (\partial_j f)(a) h_j$$

folgt dann auch die Stetigkeit von f' .

9.6 Höhere Ableitungen

Die höheren Ableitungen werden analog zu den ersten definiert. Es sei also wieder $U \in \mathbb{R}^n$, U offen, und

$$f : U \longrightarrow \mathbb{R}^m.$$

Dann hatten wir

$$f' : U \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \quad (*)$$

oder

$$f' : U \times \mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad (**)$$

und zwar linear in der zweiten Variablen.

1. Bezogen auf Standardbasen handelt es sich bei $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ um $(n \times m)$ -Matrizen. $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ ist ein $(n \cdot m)$ -dimensionaler Vektorraum. Es sei f' differenzierbar. Dann ist nach (*)

$$f'' := (f')'$$

mit

$$f'' : U \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$$

und analog

$$f''' : U \longrightarrow \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)))$$

und so weiter.

2. Schreiben wir f' in der Form (**), dann ist

$$f'(x, h) := f'(x) \cdot h,$$

und wir erhalten die zweite Ableitung in der Form

$$\begin{aligned} f'' : U \times \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x, h, k) &\longmapsto f''(x, h, k). \end{aligned}$$

f'' ist in den beiden letzten Variablen linear. Analog ist mit $h \in (\mathbb{R}^n)^p$, also $h = (h_1, h_2, \dots, h_p)$,

$$\begin{aligned} f^{(p)} : U \times (\mathbb{R}^n)^p &\longrightarrow \mathbb{R}^m \\ (x, h) &\longmapsto f^{(p)}(x, h). \end{aligned}$$

$f^{(p)}$ ist in den h_i linear.

Bleiben wir bei der zweiten Ableitung. Bezogen auf die Standardbasis im \mathbb{R}^n hatten wir

$$f'(x, h) = \sum_{i=1}^n (\partial_i f)(x) h_i = (h^t \nabla) f(x).$$

Analog ist

$$f''(x, h, k) = \sum_{j=1}^n (\partial_j f')(x, h) k_j = \sum_{i,j=1}^n (\partial_j \partial_i f)(x) h_i k_j = (k^t \nabla)(h^t \nabla) f(x)$$

und

$$f^{(p)}(x, h) = (h_1^t \nabla) \cdots (h_p^t \nabla) f(x).$$

Mit f'' existieren also auch die zweiten partiellen Ableitungen.

Wir zeigen als erstes die Symmetrie der bilinearen Abbildung

$$(h, k) \longrightarrow f''(x, h, k).$$

Es gilt der

Satz 9.6.1: *Es sei $f : U \rightarrow \mathbb{R}^m$ zweimal differenzierbar. Dann ist*

$$\forall h, k \in \mathbb{R}^n \quad f''(x, h, k) = f''(x, k, h).$$

Beweis: Es ist

$$f'(x+k, h) - f'(x, h) = f''(x, h, k) + R(x, h, k)|k|.$$

R ist linear in h , und es gilt

$$\lim_{k \rightarrow 0} R(x, h, k) = 0.$$

Wir zeigen

$$\lim_{s \downarrow 0} \frac{1}{s^2} \{f(x+sh+sk) - f(x+sh) - f(x+sk) + f(x)\} = f''(x, h, k).$$

Daraus folgt dann die Behauptung, weil die linke Seite in h und k symmetrisch ist.

Es genügt, diese Gleichung für eine Komponente von f zu beweisen, es sei also o.B.d.A. $m = 1$. Für $0 \leq t \leq 1$ sei ferner

$$F(t) := f(x+th+k) - f(x+th).$$

Dann ist

$$F'(t) = f'(x+th+k, h) - f'(x+th, h),$$

und mit

$$A(x, h, k) := F(1) - F(0) = f(x+h+k) - f(x+h) - f(x+k) + f(x)$$

folgt aus dem Mittelwertsatz

$$A(x, h, k) = F'(\theta) \quad \text{mit } 0 < \theta < 1$$

oder

$$\begin{aligned} A(x, h, k) &= \{f'(x+\theta h+k, h) - f'(x, h)\} - \{f'(x+\theta h, h) - f'(x, h)\} \\ &= f''(x, h, \theta h+k) + R(x, h, \theta h+k) \cdot |\theta h+k| - f''(x, h, \theta h) - R(x, h, \theta h) \cdot |\theta h| \\ &= f''(x, h, k) + R(x, h, \theta h+k) \cdot |\theta h+k| - R(x, h, \theta h) \cdot |\theta h|. \end{aligned}$$

Es sei nun $s > 0$. Man ersetze h, k durch sh bzw. sk . Dann folgt aus der Linearität

$$A(x, sh, sk) = s^2 \{f''(x, h, k) + R(x, h, s(\theta h+k)) \cdot |\theta h+k| - R(x, h, s\theta h) \cdot |\theta h|\}$$

und daraus die gewünschte Identität

$$\lim_{s \downarrow 0} \frac{A(x, sh, sk)}{s^2} = f''(x, h, k).$$

Der Beweis zeigt, daß f'' nur an der Stelle x existieren muß, f' natürlich in einer Umgebung $U(x)$.

Folgerung 9.6.2: f sei zweimal differenzierbar. Dann ist

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Wenn alle zweiten partiellen Ableitungen von f existieren und stetig sind, dann folgt analog Satz 9.5.4 auch die Existenz und Stetigkeit von f'' .

Wir wollen nun für reellwertige Funktionen die Taylorsche Formel übertragen. Es gilt

Taylorsche Formel: Es seien $U \subset \mathbb{R}^n$ offen, $f \in C_p(U, \mathbb{R})$, $f^{(p)}$ differenzierbar und die Verbindungsstrecke von x nach $x+h$ gehöre zu U . Dann gilt

$$f(x+h) = f(x) + f'(x, h) + \cdots + \frac{1}{p!} f^{(p)}(x, \underbrace{h, \dots, h}_{p\text{-mal}}) + R_p(x, h)$$

mit

$$R_p(x, h) = \frac{1}{(p+1)!} f^{(p+1)}(x + \theta h, \underbrace{h, \dots, h}_{(p+1)\text{-mal}}), \quad 0 < \theta < 1.$$

Beweis: Analog zum Beweis des Mittelwertsatzes sei

$$\gamma(t) = x + th \quad \text{mit } 0 \leq t \leq 1$$

eine Darstellung der Verbindungsstrecke Γ . Es sei ferner

$$F = f \circ \gamma : [0, 1] \longrightarrow \mathbb{R}.$$

Dann gilt

$$F(1) = F(0) + F'(0) + \dots + \frac{1}{p!} F^{(p)}(0) + \frac{1}{(p+1)!} F^{(p+1)}(\theta)$$

mit

$$\begin{aligned} F'(t) &= f'(\gamma(t), \gamma'(t)) = f'(\gamma(t), h) \\ F''(t) &= f''(\gamma(t), h, \gamma'(t)) = f''(\gamma(t), h, h) \\ &\vdots \\ F^{(p)}(t) &= f^{(p)}(\gamma(t), \underbrace{h, \dots, h}_{p\text{-mal}}) \end{aligned}$$

sowie

$$F(0) = f(x) \quad \text{und} \quad F(1) = f(x + h).$$

Daraus folgt die Behauptung.

Ist f $(p + 1)$ -mal stetig differenzierbar, dann erhält man auch wieder eine Integraldarstellung für das Restglied, zunächst (vgl. Gl. (**)) in §7.3)

$$R_p(x, h) = \frac{1}{p!} \int_0^1 (1-t)^p F^{(p+1)}(t) dt$$

und daraus

$$R_p(x, h) = \frac{1}{p!} \int_0^1 (1-t)^p f^{(p+1)}(x + th, \underbrace{h, \dots, h}_{(p+1)\text{-mal}}) dt.$$

Bemerkung 9.6.3: Die Übertragung der Taylorschen Formel auf vektorwertige Funktionen ist möglich. Man beachte nur, daß dabei in jeder Komponente ein anderes θ auftritt.

Aus der Taylorschen Formel erhält man für beliebig oft differenzierbare Funktionen analog Definition 7.3.4 die Taylorreihe. Alles dort Gesagte gilt entsprechend.

Wir wollen die in der Taylorschen Formel auftretenden Ableitungen noch etwas anders schreiben. Es ist

$$f^{(k)}(x, h, \dots, h) = \sum_{i_1, \dots, i_k=1}^n \partial_{i_1} \dots \partial_{i_k} f(x) h_{i_1} \dots h_{i_k}.$$

Wegen der Vertauschbarkeit der partiellen Ableitungen können von den n^k Summanden jeweils mehrere zusammengefaßt werden. Eine bequeme Schreibweise erreicht man durch Verwenden von *Multi-Indizes*. Für $\alpha_i \in \mathbb{N}_0$ seien

$$\begin{aligned} \alpha &:= (\alpha_1, \alpha_2, \dots, \alpha_n) \\ \partial^\alpha &:= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} \\ h^\alpha &:= h_1^{\alpha_1} h_2^{\alpha_2} \dots h_n^{\alpha_n} \\ \alpha! &:= (\alpha_1!) (\alpha_2!) \dots (\alpha_n!) \\ |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n. \end{aligned}$$

Bei vorgegebenem $k = |\alpha|$ gibt es insgesamt

$$\frac{k!}{(\alpha_1!) (\alpha_2!) \dots (\alpha_n!)} = \frac{k!}{\alpha!}$$

Summanden mit demselben Beitrag

$$\partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n} f(x) h_1^{\alpha_1} h_2^{\alpha_2} \dots h_n^{\alpha_n} = \partial^\alpha f(x) h^\alpha,$$

also

$$\frac{1}{k!} f^{(k)}(x, \underbrace{h, \dots, h}_{k\text{-mal}}) = \sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha.$$

Damit lautet die Taylorsche Formel

$$f(x+h) = \sum_{k=0}^p \left(\sum_{|\alpha|=k} \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha \right) + R_p(x, h).$$

Dabei ist

$$R_p(x, h) = \sum_{|\alpha|=p+1} \frac{1}{\alpha!} (\partial^\alpha f)(x + \theta h) h^\alpha.$$

Folgerungen: Als Folgerung aus der Taylorschen Formel lassen sich wieder Kriterien für die Existenz lokaler Extrema angeben.

Definition 9.6.4: Es sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Dann heißt $a \in U$ kritische Stelle von $f : \iff$

$$\forall h \in \mathbb{R}^n \quad f'(a, h) = 0.$$

Der Gradient von f verschwindet dann also in a . Als erstes zeigen wir

Satz 9.6.5: Es sei a eine lokale Extremalstelle der differenzierbaren Funktion f . Dann ist a kritische Stelle von f .

Denn es sei a etwa ein Maximum. Dann gibt es eine Umgebung V mit

$$\forall x \in V \quad f(x) \leq f(a).$$

Man wähle h so klein, daß auch $a + th$ mit $t \in [-1, 1]$ in V liegt. Es sei

$$F(t) := f(a + th) \quad \text{mit } t \in [-1, 1].$$

Wegen $F(t) \leq F(0)$ ist dann nach Satz 5.5.1 $F'(0) = f'(a, h) = 0$. Das gilt zunächst nur für kleine h , wegen der Linearität von f' dann aber für alle h .

Beispiel 9.6.6: Es sei für $x_i \geq 0$

$$f(x) := \frac{(x_1 \cdot x_2 \cdots x_n)^{1/(n+1)}}{1 + x_1 + x_2 + \cdots + x_n}.$$

f ist stetig in \bar{U} und differenzierbar in U mit

$$U := (\mathbb{R}^+)^n.$$

Mit $f := Z/N$ ist

$$\partial_i f(x) = \frac{1}{N^2(x)} \left\{ -Z(x) + \frac{N(x)Z(x)}{(n+1)x_i} \right\}.$$

Aus $\partial_i f(a) = 0$ folgt also für alle i

$$(n+1)a_i = 1 + a_1 + a_2 + \cdots + a_n.$$

Die rechten Seiten hängen nicht von i ab. Mithin sind alle a_i gleich, etwa $a_i = \alpha$; es folgt

$$(n+1)\alpha = 1 + n\alpha,$$

also $\alpha = 1$. Mithin ist

$$a = (1, 1, \dots, 1)$$

die einzige kritische Stelle von f . Dort ist

$$f(a) = \frac{1}{n+1}.$$

Wir wollen direkt nachweisen, daß f an dieser Stelle ein absolutes Maximum annimmt. Dazu betrachte man f auf dem Kompaktum

$$K := \{x \in \bar{U} \mid 0 \leq x_1 + x_2 + \cdots + x_n \leq c\}$$

mit genügend großem $c > 0$. An den Randpunkten ∂K gilt

$$\begin{aligned} \text{für ein } x_i = 0 : & \quad f(x) = 0 \\ \text{für } x_1 + \cdots + x_n = c : & \quad f(x) \leq \frac{c^{n/(n+1)}}{1+c} < c^{-1/(n+1)}. \end{aligned}$$

Wir können c deshalb so groß wählen (etwa $c := (n+1)^{n+1}$), daß

$$\forall x \in \partial K \quad f(x) < \frac{1}{n+1}$$

ist. Diese Abschätzung gilt dann erst recht für größere c . Die stetige Funktion f besitzt in K ein Maximum. Wegen $f|_{\partial K} < f(a)$ liegt es im Inneren. Somit hat f in a ein absolutes Maximum. Das war zu zeigen. Mithin gilt

$$\forall x \in U \quad f(x) \leq \frac{1}{n+1}.$$

Wählen wir $y_{n+1} > 0$ und $y_i := y_{n+1}x_i$, $i = 1, \dots, n$, dann folgt

$$\sqrt[n]{y_1 \cdots y_n} \leq (y_{n+1})^{\frac{n}{n+1}} \frac{y_1 + \cdots + y_{n+1}}{(n+1)y_{n+1}}$$

oder

$$\sqrt[n]{y_1 \cdot y_2 \cdots y_{n+1}} \leq \frac{y_1 + y_2 + \cdots + y_{n+1}}{n+1}.$$

Das ist die allgemeine Abschätzung für das arithmetisch-geometrische Mittel.

In diesem Beispiel wurde die Entscheidung, ob ein Extremum vorliegt, direkt getroffen – ohne die zweiten Ableitungen heranzuziehen. Jetzt wollen wir aber auch diese benutzen. Die Rolle der zweiten Ableitungen spielen nun *quadratische Formen*

$$Q(h, h) = \sum_{i=1}^n \sum_{j=1}^n q_{ij} h_i h_j,$$

und es geht darum festzustellen, ob diese *positiv* oder *negativ* definit sind. $Q(h, h)$ heißt

positiv definit : $\iff \forall h \neq 0 \quad Q(h, h) > 0$

negativ definit : $\iff \forall h \neq 0 \quad Q(h, h) < 0$

indefinit : $\iff Q$ nimmt sowohl positive als auch negative Werte an.

Es sind auch *positiv semidefinit* und *negativ semidefinit* möglich. Kriterien dafür, ob eine Form definit ist, werden in der linearen Algebra bewiesen. Hier haben wir es mit einem besonders einfachen Fall zu tun. Weil die partiellen Ableitungen vertauschbar sind, können wir von symmetrischen (q_{ij}) ausgehen. Dann ist diese Matrix diagonal-ähnlich. Man hat also nur die Eigenwerte der (q_{ij}) auszurechnen, und Q ist zum Beispiel positiv definit, wenn alle Eigenwerte positiv sind.

Kehren wir zu unserer Aufgabe zurück. Wir zeigen

Satz 9.6.7: $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sei zweimal stetig differenzierbar, und $a \in U$ sei kritische Stelle von f . Dann gilt

1. $f''(a, h, h)$ positiv definit $\implies a$ ist Minimalstelle von f ,
2. $f''(a, h, h)$ negativ definit $\implies a$ ist Maximalstelle von f ,
3. $f''(a, h, h)$ indefinit $\implies a$ ist keine Extremalstelle.

Der Beweis folgt wieder aus der Taylorschen Formel

$$f(a+h) = f(a) + f'(a, h) + \frac{1}{2} f''(a+\theta h, h, h).$$

Wegen der Stetigkeit der zweiten Ableitungen ist

$$f''(a+\theta h, h, h) = f''(a, h, h) + g(a, \theta h)|h|^2$$

mit

$$\lim_{h \rightarrow 0} |g(a, \theta h)| = 0.$$

Daraus folgt die Behauptung. Die Matrix $f''(a, \cdot, \cdot)$ heißt *Hesse-Matrix* von f an der Stelle a , benannt nach LUDWIG HESSE (1811–1874).

Beispiel 9.6.8: Es seien die Winkel α, β, γ gemessen worden und $\delta := \pi - (\alpha + \beta + \gamma)$. Die „wahren Werte“ seien $\alpha + x$, $\beta + y$ und $\gamma + (\delta - x - y)$. Man bestimme x und y so, daß

$$f(x, y) = x^2 + y^2 + (\delta - x - y)^2$$

minimal wird.

Es geht also um die *Methode der kleinsten Fehlerquadrate*. Es ist

$$\nabla f(x, y) = \begin{pmatrix} 4x + 2y - 2\delta \\ 4y + 2x - 2\delta \end{pmatrix}.$$

Mithin ist

$$a = \frac{\delta}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

kritische Stelle. Für die zweiten Ableitungen erhalten wir

$$(\partial_i \partial_j f)(x, y) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

Diese Matrix hat die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 6$. Sie ist also positiv definit. f hat daher in a ein Minimum.

Beispiel 9.6.9: Es seien in der (x, y) -Ebene $n \geq 2$ Punkte (x_i, y_i) gegeben. Die x_i seien nicht alle gleich. Es sei $g(x) = \alpha x + \beta$ eine Gerade und $r_i := y_i - g(x_i)$. Man bestimme g so, daß

$$|r|^2 = \sum_{i=1}^n (y_i - g(x_i))^2$$

minimal wird.

Es geht also wieder um die Methode der kleinsten Fehlerquadrate. Es seien

$$r := \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}, \quad y := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad z = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}.$$

Dann folgt aus $r_i = y_i - \alpha x_i - \beta$

$$r = -Az + y$$

und

$$|r|^2 = (-Az + y)^t (-Az + y) = z^t A^t A z - 2z^t A^t y + y^t y =: f(z).$$

Für die ersten Ableitungen erhalten wir

$$(\nabla f)(z) = 2A^t A z - 2A^t y,$$

und die Matrix der zweiten Ableitungen ist $A^t A$. Es sei $(\nabla f)(a) = 0$. Dann ist

$$A^t A a = A^t y.$$

Diese Beziehung wird *Normalgleichung* genannt. Sie ist also nach a aufzulösen. Es ist

$$A^t A = \begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} = \begin{pmatrix} |x|^2 & \sigma \\ \sigma & n \end{pmatrix}$$

mit $\sigma := \sum_{i=1}^n x_i$. Diese Matrix ist positiv definit. Aus der Cauchy-Schwarzschen Ungleichung folgt nämlich

$$\sigma^2 = \left(\sum_{i=1}^n 1 \cdot x_i \right)^2 \leq n \sum_{i=1}^n x_i^2 = n|x|^2,$$

und das Gleichheitszeichen gilt nur, wenn alle x_i gleich sind. Das ist aber ausgeschlossen. Also ist $\sigma^2 < n|x|^2$ und damit $A^t A$ positiv definit. Mithin ist die Normalgleichung nach a auflösbar, und a ist Minimalstelle von f .

Konvexe Funktionen: In diesem Zusammenhang wollen wir uns noch kurz mit konvexen Funktionen beschäftigen.

Definition 9.6.10: Eine Teilmenge M des \mathbb{R}^n heißt konvex, wenn mit x_1 und x_2 auch die Verbindungsstrecke

$$x_1 + t(x_2 - x_1) \quad \text{mit } t \in [0, 1]$$

zu M gehört.

Offenbar ist der Durchschnitt beliebiger konvexer Mengen konvex.

Definition 9.6.11: Es sei $U \subset \mathbb{R}^n$ offen und konvex. Dann heißt $f : U \rightarrow \mathbb{R}$ konvex : \Leftrightarrow

$$\forall x_1, x_2 \in U, x_1 \neq x_2 \quad \forall t \in [0, 1] \quad f(x_1 + t(x_2 - x_1)) \leq f(x_1) + t(f(x_2) - f(x_1)).$$

f heißt streng konvex, wenn sogar

$$f(x_1 + t(x_2 - x_1)) < f(x_1) + t(f(x_2) - f(x_1))$$

gilt.

Es sei nun f eine in $I = (a, b)$ definierte konvexe Funktion. Dann folgt für $a < u < v < w < b$

$$\frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(u)}{w - u} \leq \frac{f(w) - f(v)}{w - v}.$$

Es sei $v := u + t(w - u)$ und $y := f(u) + t(f(w) - f(u))$. Dann ist $f(v) \leq y$ und

$$y = f(u) + \frac{v - u}{w - u}(f(w) - f(u)) = f(w) - \frac{w - v}{w - u}(f(w) - f(u)),$$

also

$$\frac{f(v) - f(u)}{v - u} \leq \frac{y - f(u)}{v - u} = \frac{f(w) - f(u)}{w - u} = \frac{f(w) - y}{w - v} \leq \frac{f(w) - f(v)}{w - v}.$$

Insbesondere folgt daraus mit $a < \alpha < u$

$$\frac{f(v) - f(\alpha)}{v - \alpha} \leq \frac{f(v) - f(u)}{v - u} \leq \frac{f(w) - f(v)}{w - v}.$$

Mithin ist

$$\lim_{u \rightarrow v} \frac{f(v) - f(u)}{v - u} = \frac{f(w) - f(v)}{w - v}.$$

f ist also stetig. Es existieren auch die einseitigen Ableitungen von f . Das folgt aus der Monotonie und Beschränktheit der Differenzenquotienten. Das Beispiel $f(x) := |x|$ zeigt aber, daß die Differenzierbarkeit selbst nicht zu erwarten ist. f'_+ und f'_- wachsen jedoch monoton. Damit existiert f' mit Ausnahme höchstens abzählbar unendlich vieler Punkte. Wichtig ist nun der

Satz 9.6.12: Es seien $U \subset \mathbb{R}^n$ offen und konvex und $f : U \rightarrow \mathbb{R}$ differenzierbar. Dann gilt: f ist streng konvex \Leftrightarrow

$$\forall x \in U \quad \forall h \in \mathbb{R}^n \setminus \{0\}, x + h \in U \quad f(x + h) > f(x) + f'(x, h).$$

Beweis:

1. \Rightarrow

$$f(x + th) = f(x + t((x + h) - x)) < f(x) + t(f(x + h) - f(x))$$

oder

$$\frac{f(x + th) - f(x)}{t} < f(x + h) - f(x).$$

Daraus folgt

$$f'(x, h) \leq f(x + h) - f(x).$$

Um das Gleichheitszeichen in der letzten Abschätzung auszuschließen, ersetzen wir darin h durch th . Aus der Linearität von f' folgt dann

$$f'(x, h) \leq \frac{f(x + th) - f(x)}{t} < f(x + h) - f(x).$$

2. \Leftarrow

Nun ist $f(x) + f'(x, h) < f(x + h)$. Es seien $x_1, x_2 \in U$ und $t \in (0, 1)$. Dann folgt aus

$$f(x) + f'(x, x_1 - x) < f(x_1)$$

$$f(x) + f'(x, x_2 - x) < f(x_2)$$

durch Addition nach vorheriger Multiplikation mit $(1 - t)$ bzw. t

$$f(x) + f'(x, x_1 + t(x_2 - x_1) - x) < f(x_1) + t(f(x_2) - f(x_1))$$

oder mit $x = x_1 + t(x_2 - x_1)$

$$f(x_1 + t(x_2 - x_1)) + \underbrace{f'(x, 0)}_{=0} < f(x_1) + t(f(x_2) - f(x_1)).$$

Das war zu zeigen.

Konvexe Funktionen lassen sich entsprechend durch

$$f(x + h) \geq f(x) + f'(x, h)$$

kennzeichnen.

Es sei f nun in U zweimal differenzierbar. Dann ist mit $\theta \in (0, 1)$

$$f(x + h) = f(x) + f'(x, h) + \frac{1}{2} f''(x + \theta h, h, h).$$

f ist deshalb streng konvex, wenn für alle $h \neq 0$

$$f''(x + \theta h, h, h) > 0$$

ist. Hieraus erhält man eine hinreichende Bedingung für die strenge Konvexität, nämlich

$$\forall x \in U \quad \forall h \neq 0 \quad f''(x, h, h) > 0,$$

also den

Satz 9.6.13: *Es sei $U \subset \mathbb{R}^n$ konvex und $f : U \rightarrow \mathbb{R}$ zweimal differenzierbar. Es sei f'' überall positiv definit. Dann ist f streng konvex.*

9.7 Lokale Umkehrbarkeit

In diesem Abschnitt wollen wir die Frage nach der Existenz der Umkehrabbildung anschneiden und zeigen, daß es sie jedenfalls lokal gibt. Im \mathbb{R}^1 lieferte uns die Monotonie die Injektivität; im \mathbb{R}^n ist die Situation aber nicht mehr so einfach. Wir nehmen hier $n = m$ an, eine allgemeinere Situation wird in §9.8 behandelt. Denkt man an den linearen Fall, dann liegt es nahe, $\det f'(x) \neq 0$ vorauszusetzen. Das ist aber schon im \mathbb{R}^1 nicht genug. Man betrachte etwa

$$f(x) = \begin{cases} x + x^2 \cos \frac{1}{x} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Dann ist f im Nullpunkt differenzierbar mit $f'(0) = 1$. f ist aber in $U(0)$ nicht injektiv, denn $x_n := ((3/2 + 2n\pi)^{-1})$ sind Maximalstellen. Um solche Fälle auszuschließen, fordern wir zusätzlich noch die Stetigkeit der ersten Ableitung. Wir zeigen den

Satz 9.7.1: *Es seien $U \subset \mathbb{R}^n$, $f \in C_1(U, \mathbb{R}^n)$, $a \in U$ und $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ invertierbar. Dann gibt es ein offenes $V \subset U$ mit $a \in V$ und*

1. $f|V$ ist injektiv.
2. $f(V)$ ist offen.
3. Die Umkehrabbildung g von $f|V$ ist stetig differenzierbar.

f bildet V also diffeomorph auf $f(V)$ ab.

Zum Beweis können wir uns auf den Fall $a = 0$ und $f(a) = 0$ beschränken (wir betrachten $x \mapsto f(x + a) - f(a)$). Außerdem normieren wir die Ableitung im Nullpunkt und nehmen $f'(0) = id$ an. Das erreicht man durch Multiplikation mit $(f'(0))^{-1}$; wir betrachten also $x \mapsto (f'(0))^{-1}f(x)$.

Es sei also im folgenden o.B.d.A. $a = 0$ sowie

$$f(0) = 0 \quad \text{und} \quad f'(0) = id.$$

Dann führen wir den Beweis in drei Schritten.

1. Wir legen V fest:

Es sei $\Phi_y(x) := x - f(x) + y$ mit $y \in \mathbb{R}^n$. Jeder Fixpunkt von Φ_y ist dann Urbild von y . Wir bestimmen im zweiten Schritt ein $r \in \mathbb{R}^+$ mit

$$\forall y \in B(0, r) \quad \Phi_y : \overline{B(0, 2r)} \rightarrow \overline{B(0, 2r)},$$

und zwar kontrahierend. Wenn das geschehen ist, folgt aus dem Banachschen Fixpunktsatz

$$\forall y \in B(0, r) \quad \exists! x \in \overline{B(0, 2r)} \quad f(x) = y.$$

Das Urbild von $B(0, r)$ sei nun V . $V \subset \overline{B(0, 2r)}$ ist offen, weil f stetig ist, und

$$f : V \rightarrow B(0, r)$$

ist dann bijektiv.

2. Wahl von r und Nachweis der Kontraktion:

Wegen $f'(0) = id$ und der Stetigkeit von f' gilt

$$\exists r > 0 \quad \forall x, |x| \leq 2r \quad \|id - f'(x)\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)} \leq \frac{1}{2}.$$

Aus dem Mittelwertsatz und $\Phi'_0(x) = (\Phi'_{0,j})(x) = id - f'(x)$ folgt daher mit $0 < \theta_j < 1$

$$|x - f(x)| = |\Phi_0(x) - \Phi_0(0)| = |(\Phi'_{0,j}(\theta_j x, x))| \leq \frac{1}{2}|x|.$$

Für $|y| < r$ und $|x| \leq 2r$ erhalten wir deshalb

$$|\Phi_y(x)| \leq |x - f(x)| + |y| < 2r,$$

also

$$\forall y, |y| < r \quad \Phi_y : \overline{B(0, 2r)} \rightarrow \overline{B(0, 2r)}.$$

Es bleibt der Nachweis der Kontraktion. Es seien $x_1, x_2 \in \overline{B(0, 2r)}$. Dann gilt mit $0 < \theta_j < 1$

$$\Phi_{y,j}(x_2) - \Phi_{y,j}(x_1) = \Phi'_{y,j}(x_1 + \theta_j(x_2 - x_1), x_2 - x_1).$$

Es ist $\Phi'_y = \Phi'_0$ und

$$x_1 + \theta_j(x_2 - x_1) = (1 - \theta_j)x_1 + \theta_j x_2 \in \overline{B(0, 2r)},$$

also

$$|\Phi_y(x_2) - \Phi_y(x_1)| \leq \frac{1}{2}|x_2 - x_1|.$$

3. Wir zeigen die dritte Behauptung:

Es seien $x_1, x_2 \in \overline{B(0, 2r)}$. Aus $\Phi_0(x_i) = x_i - f(x_i)$ folgt dann

$$x_2 - x_1 = \Phi_0(x_2) - \Phi_0(x_1) + f(x_2) - f(x_1)$$

und

$$|x_2 - x_1| \leq \frac{1}{2}|x_2 - x_1| + |f(x_2) - f(x_1)|,$$

also mit $y_i := f(x_i)$

$$|g(y_2) - g(y_1)| \leq 2|y_2 - y_1|.$$

Das heißt, g ist stetig. Aus dem Satz von der Differenzierbarkeit der Umkehrabbildung (Satz 9.4.8) folgt dann

$$g'(y) = (f'(x))^{-1} = (f'(g(y)))^{-1}.$$

Damit ist dann auch g' stetig. Denn f' ist stetig, linear und bijektiv. Nach Satz 9.2.8 ist dann auch $(f')^{-1}$ stetig. Damit ist unser Satz bewiesen.

Die globale Umkehrbarkeit ist schwieriger zu zeigen. Ich gebe hier nur zwei hinreichende Bedingungen an:

1. Es sei U konvex und $f : U \rightarrow \mathbb{R}^n$ differenzierbar. Dann folgt aus dem Mittelwertsatz

$$f(x+h) - f(x) = \begin{pmatrix} \partial_1 f_1(c_1) & \cdots & \partial_n f_1(c_1) \\ \vdots & \searrow & \vdots \\ \partial_1 f_n(c_n) & \cdots & \partial_n f_n(c_n) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$$

mit $c_j = x + \theta_j h$, $0 < \theta_j < 1$. Es sei nun für alle $z_1, \dots, z_n \in U$

$$\det \begin{pmatrix} \partial_1 f_1(z_1) & \cdots & \partial_n f_1(z_1) \\ \vdots & \searrow & \vdots \\ \partial_1 f_n(z_n) & \cdots & \partial_n f_n(z_n) \end{pmatrix} \neq 0.$$

Dann ist f invertierbar, denn aus $f(x+h) = f(x)$ folgt $h = 0$.

Beispiel 9.7.2: Es sei im \mathbb{R}^2

$$f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

Dann ist

$$f'(x, y) = 2 \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

und $\det f'(x, y) = 4(x^2 + y^2)$ sowie mit $z_i := (x_i, y_i)$

$$\det \begin{pmatrix} x_1 & -y_1 \\ y_2 & x_2 \end{pmatrix} = x_1 x_2 + y_1 y_2 = (z_1, z_2).$$

In $\mathbb{R}^2 \setminus \{0\}$ ist f lokal invertierbar, aber nicht global, denn es ist ja $f(-x, -y) = f(x, y)$. f ist jedoch global invertierbar in konvexen Teilbereichen $U \subset \mathbb{R}^2$, für die

$$\forall z_1, z_2 \in U \quad (z_1, z_2) \neq 0$$

ist. Solche U lassen sich leicht angeben, zum Beispiel $U = \mathbb{R}^+ \times \mathbb{R}^+$.

2. Lokal umkehrbare lineare Abbildungen sind auch global umkehrbar. Wir erwarten deshalb, daß eine beliebige C_1 -Abbildung, deren Ableitung sich nur „wenig“ von einer festen linearen unterscheidet, auch global umkehrbar ist. Genauer gilt

Satz 9.7.3: Es seien U konvex und $f \in C_1(U, \mathbb{R}^n)$ mit $\det f'(x) \neq 0$. Es gebe eine umkehrbare lineare Abbildung $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ mit

$$\forall x \in U \quad \|id - f'(x) \cdot A^{-1}\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)} < 1.$$

Dann ist f global umkehrbar.

Beweis: f ist lokal umkehrbar. Wir müssen also nur die Injektivität von f zeigen. Es seien $x, x+h \in U$, $h \neq 0$ und $f(x) = f(x+h)$. Dann ist

$$0 = f(x+h) - f(x) = \int_0^1 \left(\frac{d}{dt} f(x+th) \right) dt = \int_0^1 f'(x+th) h dt,$$

oder

$$Ah = \int_0^1 Ah dt = \int_0^1 (A - f'(x+th))h dt = \int_0^1 (id - f'(x+th) \cdot A^{-1})Ah dt.$$

Daraus folgt der Widerspruch

$$|Ah| \leq \int_0^1 \|id - f'(x+th) \cdot A^{-1}\|_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)} |Ah| dt < |Ah|.$$

Beispiel 9.7.4: Es seien $U = B(0, 3) \subset \mathbb{R}^2$ und $f : U \rightarrow \mathbb{R}^2$ mit

$$f(x, y) = \begin{pmatrix} x \\ y \end{pmatrix} + \lambda \begin{pmatrix} \varphi(x, y) \\ \psi(x, y) \end{pmatrix}$$

und differenzierbaren φ, ψ .

In diesem Beispiel ist

$$f'(x, y) = id + \lambda \begin{pmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{pmatrix}(x, y),$$

also (hier ist $A = id$)

$$\|id - f'(x, y)\|_{\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)} = |\lambda| \left\| \begin{pmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{pmatrix} \right\|_{\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)}.$$

Beispielsweise seien $\varphi(x, y) = y^2$ und $\psi(x, y) = x^2$. Dann ist

$$\left\| \begin{pmatrix} \varphi_x & \varphi_y \\ \psi_x & \psi_y \end{pmatrix} \right\| = \left\| 2 \begin{pmatrix} 0 & y \\ x & 0 \end{pmatrix} \right\| = \sup_{|h|=1} 2 \sqrt{(h_1 x)^2 + (h_2 y)^2} \leq 2 \sqrt{x^2 + y^2} \leq 6.$$

Für $|\lambda| < 1/6$ ist f deshalb global invertierbar.

9.8 Implizite Funktionen

In den Anwendungen tritt oft folgende Aufgabe auf: Gegeben sei ein Gleichungssystem

$$\begin{aligned} f_1(x_1, \dots, x_n, u_1, \dots, u_m) &= 0 \\ &\vdots \\ f_n(x_1, \dots, x_n, u_1, \dots, u_m) &= 0. \end{aligned}$$

Mit $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ und $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ schreiben wir dafür kurz

$$f(x, u) = 0.$$

Gesucht ist eine Auflösung $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^n$

$$x = \varphi(u)$$

mit $f(\varphi(u), u) = 0$.

Im allgemeinen ist eine solche Auflösung nicht möglich. Zum Beispiel sei $n = m = 1$ und

$$f(x, u) = 1 + x^2 + u^2.$$

Dann gibt es keine reelle Auflösung. Es sind also Voraussetzungen an f erforderlich, zum Beispiel die, daß es eine Stelle (x_1, u_1) mit $f(x_1, u_1) = 0$ geben soll.

Es seien $h \in \mathbb{R}^n$ und $k \in \mathbb{R}^m$. Dann schreiben wir

$$f' = f'(x, u; h, k) =: f'_1(x, u; h) + f'_2(x, u; k)$$

vermöge

$$\begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 & \partial_{n+1} f_1 & \cdots & \partial_{n+m} f_1 \\ \vdots & \diagdown & \vdots & \vdots & & \vdots \\ \partial_1 f_n & \cdots & \partial_n f_n & \partial_{n+1} f_n & \cdots & \partial_{n+m} f_n \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ k_m \end{pmatrix} = \begin{pmatrix} \partial_1 f_1 & \cdots & \partial_n f_1 \\ \vdots & \diagdown & \vdots \\ \partial_1 f_n & \cdots & \partial_n f_n \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} + \begin{pmatrix} \partial_{n+1} f_1 & \cdots & \partial_{n+m} f_1 \\ \vdots & & \vdots \\ \partial_{n+1} f_n & \cdots & \partial_{n+m} f_n \end{pmatrix} \begin{pmatrix} k_1 \\ \vdots \\ k_m \end{pmatrix}$$

und zeigen

Satz 9.8.1: Es seien $D \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f \in C_1(D, \mathbb{R}^n)$. Es gebe ein $a \in \mathbb{R}^n$ und ein $b \in \mathbb{R}^m$ mit

$$f(a, b) = 0 \quad \text{und} \quad \det f'_1(a, b) \neq 0.$$

Dann existieren eine Umgebung $U(b) \subset \mathbb{R}^m$ und eine eindeutig bestimmte C_1 -Abbildung

$$\varphi : U \longrightarrow \mathbb{R}^n$$

mit

$$\forall u \in U \quad f(\varphi(u), u) = 0.$$

Beweis: Es liegt nahe, den Satz von der lokalen Umkehrbarkeit bei festem u auf

$$f(x, u) = y$$

anzuwenden und in die Umkehrabbildung

$$g_u(y) = x$$

dann $y = 0$ einzusetzen, also $\varphi(u) := g_u(0)$ zu wählen. Um auch für variables u Aussagen zu erhalten, erweitern wir den Wertebereich von f und betrachten

$$F : D \longrightarrow \mathbb{R}^{n+m} \\ (x, u) \longmapsto (f(x, u), u).$$

Dann ist F stetig differenzierbar, und es gilt

$$F'(x, u) = \begin{pmatrix} f'_1(x, u) & f'_2(x, u) \\ 0 & id_{\mathbb{R}^m} \end{pmatrix}.$$

Hieraus folgt

$$\det F'(x, u) = \det f'_1(x, u) \quad \text{mit} \quad \det F'(a, b) \neq 0.$$

Jetzt können wir den Satz von der lokalen Umkehrbarkeit auf F anwenden. Zu $(a, b) \in D$ gibt es eine Umgebung $W \subset D$, so daß $F|_W$ eine C_1 -Umkehrabbildung

$$G : F(W) \longrightarrow W$$

besitzt. $F(W)$ ist offen.

Aus

$$F(x, u) = (y, v)$$

erhält man für G

$$G(y, v) = (x, u)$$

mit

$$x =: g(y, v) \quad \text{und} \quad u = v.$$

Dabei ist g aus $C_1(\mathbb{R}^{n+m}, \mathbb{R}^n)$. Dieses g gibt die gesuchte Abbildung φ , nämlich

$$\varphi(v) := g(0, v).$$

Es war $D(G) = F(W)$ offen. Mithin existiert eine Umgebung $U(b) \subset \mathbb{R}^m$ mit

$$\forall v \in U(b) \quad (0, v) \in F(W).$$

Dort ist φ erklärt. Aus $F \circ G = id$ folgt

$$f(\underbrace{g(y, v)}_{=x}, \underbrace{v}_{=u}) = y \quad \text{für} \quad (y, v) \in F(W).$$

Mit $y = 0$ und $v = u$ erhält man also

$$\forall u \in U(b) \quad f(\varphi(u), u) = 0,$$

das heißt, $\varphi(u)$ löst die Gleichung.

Daß $\varphi(u)$ die einzige Lösung ist, folgt aus

$$g(f(x, u), u) = x.$$

Denn wenn $f(x, u) = 0$ ist, folgt daraus

$$g(0, u) = x,$$

also $x = \varphi(u)$.

Wir bestimmen noch die Ableitung von φ . Aus

$$f(\varphi(u), u) = 0$$

und

$$f'(x, u; h, k) = f'_1(x, u; h) + f'_2(x, u; k)$$

folgt

$$f'_1(\varphi(u), u; \varphi'(u, k)) + f'_2(\varphi(u), u; k) = 0.$$

Wegen $\det f'_1(\varphi(u), u) \neq 0$ für $u \in U(b)$ kann man diese Gleichung nach $\varphi'(u, k)$ auflösen,

$$\varphi'(u, k) = -\left(f'_1(\varphi(u), u)\right)^{-1} f'_2(\varphi(u), u; k)$$

oder

$$\varphi'(u) = -\left(f'_1(\varphi(u), u)\right)^{-1} f'_2(\varphi(u), u).$$

Beispiel 9.8.2:

1. Es seien $n = m = 1$ und

$$f(x, u) = x + x^2 - u^2.$$

Dann ist $f(0, 0) = 0$ und

$$f'_1(x, u) = 1 + 2x,$$

also $f'_1(0, 0) = 1$. Die Voraussetzungen des Satzes sind also für $a = b = 0$ erfüllt. Formal erhält man aus $f(x, u) = 0$

$$x = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + u^2}.$$

Wegen $\varphi(0) = 0$ kommt aber nur

$$\varphi(u) = -\frac{1}{2} + \sqrt{\frac{1}{4} + u^2}$$

in Betracht. Die zweite Wurzel gehört zu $f(-1, 0) = 0$.

2. Es seien $n = 2, m = 1$ und

$$f(x, y, u) = \begin{pmatrix} 1 + x + y - u \\ -1 + x + y^2 + u \end{pmatrix}.$$

Dann ist $f(0, 0, 1) = 0$ und

$$f'_1(x, y, u) = \begin{pmatrix} 1 & 1 \\ 1 & 2y \end{pmatrix},$$

also $\det f'_1(x, y, u) = 2y - 1$. Die Voraussetzungen des Satzes sind also mit $a = (0, 0)$ und $b = 1$ wieder erfüllt. Formales Rechnen ergibt aus $f(x, y, u) = 0$ nach Elimination von x

$$0 = y^2 - y + 2(u - 1),$$

also

$$y = \frac{1}{2}(1 \pm \sqrt{9 - 8u}).$$

Es soll $\varphi(1) = (0, 0)$ sein. Daraus folgt dann

$$\varphi(u) = \frac{1}{2} \begin{pmatrix} 2u - 3 + \sqrt{9 - 8u} \\ 1 - \sqrt{9 - 8u} \end{pmatrix}.$$

Die zweite Wurzel würde zu $\varphi(1) = (-1, 1)$ passen.

9.9 Kurven und Flächen

Bei der Behandlung des Zwischenwertsatzes in §9.1 haben wir schon einen Weg definiert. Ich wiederhole die

Definition 9.9.1: Eine stetige Abbildung $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt Weg von $\gamma(a)$ nach $\gamma(b)$.

Solche Wege können noch recht „wild“ verlaufen. So stammt von PEANO das Beispiel eines stetigen Weges, der eine Dreiecksfläche ganz ausfüllt. Deshalb definiert man weiter:

Definition 9.9.2: Ist γ an der Stelle $s_0 \in [a, b]$ differenzierbar, dann heißt der Weg dort differenzierbar.

Definition 9.9.3: Ist γ überall stetig differenzierbar, und ist überall $\gamma' \neq 0$, dann heißt der Weg γ glatt. γ heißt stückweise glatt, wenn er aus endlich vielen glatten zusammengesetzt ist.

Es sei γ differenzierbar und g die Gerade

$$g(s) := \gamma(s_0) + \gamma'(s_0)(s - s_0).$$

Dann ist

$$\gamma(s) = g(s) + |s - s_0| \cdot r(s_0, s - s_0)$$

mit

$$\lim_{s \rightarrow s_0} r(s_0, s - s_0) = 0.$$

Ist $\gamma'(s_0) \neq 0$, dann ist g tangential, und man erhält den Tangenteneinheitsvektor t durch

$$t(s_0) := \frac{\gamma'(s_0)}{|\gamma'(s_0)|}.$$

Für glatte bzw. stückweise glatte Wege läßt sich nun die euklidische Länge erklären, nämlich durch

Definition 9.9.4: Es sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ein stückweise glatter Weg. Dann heißt die Zahl

$$L(\gamma) := \int_a^b |\gamma'(s)| ds$$

Länge des Weges γ .

Beispiel 9.9.5:

1. Es sei $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ mit $\gamma(s) := a + s(b - a)$ die Strecke von $a \in \mathbb{R}^n$ nach $b \in \mathbb{R}^n$. Dann ist $\gamma'(s) = b - a$ und

$$L(\gamma) = \int_0^1 |b - a| ds = |b - a|.$$

2. Es sei $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit

$$\gamma(s) := \begin{pmatrix} \cos s \\ \sin s \end{pmatrix}$$

die Kreislinie. Dann ist

$$\gamma'(s) = \begin{pmatrix} -\sin s \\ \cos s \end{pmatrix}$$

und $|\gamma'(s)| = 1$, also $L(\gamma) = 2\pi$.

In beiden Fällen erhalten wir also die übliche Länge.

Wir wollen die Definition der Länge $L(\gamma)$ noch etwas motivieren und halten dazu fest:

1. Für Strecken ist $L(\gamma)$ die euklidische Länge.

2. L ist additiv. Das heißt: Es seien $a < b < c$, $\gamma : [a, c] \rightarrow \mathbb{R}^n$, $\gamma_1 : [a, b] \rightarrow \mathbb{R}^n$ und $\gamma_2 : [b, c] \rightarrow \mathbb{R}^n$ mit $\gamma|_{[a, b]} = \gamma_1$ und $\gamma|_{[b, c]} = \gamma_2$, also $\gamma = \gamma_1 + \gamma_2$. Dann ist

$$L(\gamma) = L(\gamma_1) + L(\gamma_2).$$

3. Es sei

$$\|\gamma\|_1 := \sup_{s \in [a, b]} |\gamma(s)| + \sup_{s \in [a, b]} |\gamma'(s)|$$

die $C_1([a, b], \mathbb{R}^n)$ -Supremumsnorm von γ . Dann hängt L bezüglich dieser Topologie stetig von γ ab. Es sei nämlich (γ_n) eine Folge von Wegen $\gamma_n := [a, b] \rightarrow \mathbb{R}^n$ mit $\|\gamma_n - \gamma\|_1 \rightarrow 0$. Dann folgt aus der Dreiecksungleichung

$$|L(\gamma_n) - L(\gamma)| = \left| \int_a^b (|\gamma'_n(s)| - |\gamma'(s)|) ds \right| \leq \int_a^b |\gamma'_n(s) - \gamma'(s)| ds \rightarrow 0.$$

L ist durch die Eigenschaften 1–3 eindeutig festgelegt. Denn γ kann durch Treppenvektoren approximiert werden, und man erhält für glatte γ gerade den in der Definition gegebenen Wert $L(\gamma)$. Dabei ist wichtig, daß die Approximation in der C_1 -Norm und nicht in der C -Norm erfolgt. Zum Beispiel darf man die Strecke $\gamma : [0, 1] \rightarrow \mathbb{R}^2$

$$\gamma(s) := \begin{pmatrix} s \\ 0 \end{pmatrix}$$

nicht durch Sägezähne $\gamma_n(s) = (s, a_n(s))$ mit

$$a_n(s) = \begin{cases} s & \text{für } 0 \leq s \leq (1/2)^n \\ 2(\frac{1}{2})^n - s & \text{für } (1/2)^n \leq s \leq 2(1/2)^n \end{cases} \text{ usw.}$$

approximieren. Es ist nämlich

$$\gamma'(s) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \gamma'_n(s) = \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix},$$

also

$$L(\gamma) = 1 \quad \text{und} \quad L(\gamma_n) = \sqrt{2}.$$

Es folgen weitere Eigenschaften von $L(\gamma)$:

4. L ändert sich bei Isometrien des \mathbb{R}^n nicht. Denn es sei $x \mapsto Ax + a$ mit $(Ax, Ax) = (x, x)$ oder $A^t A = id$. Dann ist $(A\gamma + a)' = A\gamma'$ und $|A\gamma'| = |\gamma'|$.
5. Geht man von γ durch Einführen eines neuen Parameters \tilde{s}

$$s = \varphi(\tilde{s}) \quad \text{mit } \varphi \in C_1([c, d], [a, b]) \text{ stückweise und } \varphi'(\tilde{s}) > 0$$

zum Weg

$$\tilde{\gamma}(\tilde{s}) := \gamma(\varphi(\tilde{s})) \quad \text{mit } \tilde{\gamma} : [c, d] \rightarrow \mathbb{R}^n$$

über, dann gilt

$$L(\tilde{\gamma}) = L(\gamma).$$

Es ist nämlich $|\tilde{\gamma}'(\tilde{s})| = |\gamma'(\varphi(\tilde{s}))| \cdot \varphi'(\tilde{s})$, also

$$L(\tilde{\gamma}) = \int_c^d |\tilde{\gamma}'(\tilde{s})| d\tilde{s} = \int_c^d |\gamma'(\varphi(\tilde{s}))| \cdot \varphi'(\tilde{s}) d\tilde{s} = \int_{\varphi(c)}^{\varphi(d)} |\gamma'(s)| ds = \int_a^b |\gamma'(s)| ds = L(\gamma).$$

Man kann deshalb Äquivalenzklassen stückweise glatter Wege einführen

Definition 9.9.6: Der Weg $\tilde{\gamma}$ heißt zum Weg γ äquivalent, wenn beide durch eine Parametertransformation nach 5 mit $\varphi'(\tilde{s}) > 0$ auseinander hervorgehen. Jede solche Äquivalenzklasse von stückweise glatten Wegen heißt orientierte stückweise glatte Kurve.

Die Orientierung kommt dabei durch die Forderung $\varphi'(\tilde{s}) > 0$ hinein.

Wir wollen nun eine ausgezeichnete Parameterdarstellung angeben. Es soll nämlich überall

$$|\gamma'(\sigma)| = 1$$

sein. Das kann man folgendermaßen erreichen: Es sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ gegeben. Dann wird ein

$$\varphi \in C_1([c, d], [a, b]) \quad \text{mit } \varphi' > 0$$

und

$$|\gamma'(\varphi(\sigma))| \cdot \varphi'(\sigma) = 1$$

gesucht. Es sei $\psi := \varphi^{-1}$ die Umkehrabbildung von φ , $s = \varphi(\sigma)$. Dann ist

$$\psi(s) = c + \int_c^\sigma dt = c + \int_c^\sigma |\gamma'(\varphi(t))| \cdot \varphi'(t) dt = c + \int_{\varphi(c)}^{\varphi(\sigma)} |\gamma'(u)| du = c + \int_a^s |\gamma'(u)| du.$$

Wegen $|\gamma'(u)| > 0$ wächst $\psi(s)$ streng monoton, das heißt $\varphi = \psi^{-1}$ existiert und hat die geforderten Eigenschaften. Meist wählt man $c = \psi(a) = 0$ und $L := d = \psi(b)$. Dann ist L wieder die Länge des Weges γ , also

$$L(\gamma) = \psi(b).$$

Man nennt σ deshalb die *Bogenlänge* der Kurve γ .

In dieser ausgezeichneten Darstellung lassen sich die folgenden Resultate leichter formulieren. Deshalb wollen wir diese Darstellung verwenden. Allerdings ist es oft nicht einfach, die Bogenlänge wirklich auszurechnen, weil eine Integration dabei eingeht.

Beispiel 9.9.7: Wir betrachten in $0 \leq s \leq \pi/2$ mit $a, b \in \mathbb{R}^+$

$$\gamma(s) := \begin{pmatrix} a \sin s \\ b \cos s \end{pmatrix}.$$

γ ist eine Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

und es gilt

$$\gamma'(s) = \begin{pmatrix} a \cos s \\ -b \sin s \end{pmatrix},$$

also $|\gamma'(s)| = \sqrt{a^2 \cos^2 s + b^2 \sin^2 s}$ und

$$\sigma = \int_0^s \sqrt{a^2 \cos^2 t + b^2 \sin^2 t} dt.$$

Das ist ein elliptisches Integral, welches im allgemeinen nicht elementar berechnet werden kann. Ein Näherungsverfahren haben wir im Beispiel 8.1.2.6 bereits kennengelernt. Im Falle $a = b = r$ erhalten wir für den Kreisbogen

$$\sigma = r s$$

oder $\varphi(\sigma) = \sigma/r$. Das ergibt die ausgezeichnete Darstellung

$$\tilde{\gamma}(\sigma) = \gamma\left(\frac{\sigma}{r}\right) = r \begin{pmatrix} \sin(\sigma/r) \\ \cos(\sigma/r) \end{pmatrix}.$$

Jetzt ist in der Tat $|\tilde{\gamma}'(\sigma)| = 1$.

Es sei also von nun an $\gamma(s)$ ein Weg mit $|\gamma'(s)| = 1$. Dann ist

$$t(s) = \gamma'(s)$$

der Tangentenvektor. Es sei γ zweimal stetig differenzierbar. Dann nennt man $\gamma''(s)$ den *Krümmungsvektor*. Aus $|\gamma'| = 1$ folgt unmittelbar

$$(\gamma', \gamma'') = 0.$$

Tangenten- und Krümmungsvektor stehen also senkrecht aufeinander. Ist auch $\gamma''(s) \neq 0$, dann sind $\gamma'(s)$ und $\gamma''(s)$ linear unabhängig, und man nennt die von ihnen aufgespannte Ebene die *Schmiegebene* im Punkte $\gamma(s)$. Man nennt

$$\kappa(s) := |\gamma''(s)|$$

die *Krümmung* von γ in s .

Beispiel 9.9.8: Es seien a, b zwei orthonormale Vektoren, $m \in \mathbb{R}^n$ der Mittelpunkt und

$$k(s) = m + r \left(a \cos \frac{s}{r} + b \sin \frac{s}{r} \right)$$

ein Kreis um m mit dem Radius r .

Dann ist

$$k'(s) = -a \sin \frac{s}{r} + b \cos \frac{s}{r}$$

und

$$k''(s) = -\frac{1}{r} \left(a \cos \frac{s}{r} + b \sin \frac{s}{r} \right),$$

also

$$\kappa(s) = |k''(s)| = 1/r.$$

Für beliebige Kurven ist $\kappa = \kappa(s)$ natürlich variabel, und man kann jedem s_0 einen *Krümmungskreis* $k(s_0, s)$ zuordnen, nämlich

$$k(s_0, s) = m(s_0) + r(s_0) \left(a(s_0) \cos \frac{s-s_0}{r(s_0)} + b(s_0) \sin \frac{s-s_0}{r(s_0)} \right)$$

mit

$$m(s_0) = \gamma(s_0) + \frac{\gamma''(s_0)}{\kappa^2(s_0)}, \quad r(s_0) = \frac{1}{\kappa(s_0)}, \quad a(s_0) = -\frac{\gamma''(s_0)}{\kappa(s_0)}, \quad b(s_0) = \gamma'(s_0).$$

Dieser Kreis berührt $\gamma(s)$ von zweiter Ordnung, es ist nämlich

$$\begin{aligned} k(s_0, s) &= k(s_0, s_0) + k'(s_0, s_0) \cdot (s - s_0) + k''(s_0, s_0) \cdot (s - s_0)^2 + \dots \\ &= \gamma(s_0) + \gamma'(s_0) \cdot (s - s_0) + \gamma''(s_0) \cdot (s - s_0)^2 + \dots \end{aligned}$$

9.9.1 Ebene Kurven

Wir wollen kurz ebene Kurven, also $\gamma : [a, b] \rightarrow \mathbb{R}^2$ behandeln. Es sei

$$\gamma(s) = \begin{pmatrix} x(s) \\ y(s) \end{pmatrix}.$$

Dann ist

$$t(s) = \begin{pmatrix} x'(s) \\ y'(s) \end{pmatrix}$$

die Tangente, und es sei

$$n(s) := \begin{pmatrix} -y'(s) \\ x'(s) \end{pmatrix}$$

die Normale. t und n sind linear unabhängig. Folglich gibt es bei festem s Zahlen $\alpha, \beta, \gamma, \delta$ mit

$$t' = \alpha t + \beta n$$

$$n' = \gamma t + \delta n.$$

Diese Zahlen lassen sich leicht berechnen. Es ist

$$\alpha = (t', t) = 0 \quad (\text{wegen } |t| = 1)$$

$$\delta = (n', n) = 0 \quad (\text{wegen } |n| = 1)$$

$$\beta = (t', n) = -(n', t) = -\gamma \quad (\text{wegen } (n, t) = 0).$$

β ist die durch die Wahl des Vorzeichens von n orientierte Krümmung. Aus $n = \pm \gamma'' / |\gamma''|$ folgt nämlich

$$\beta = (t', n) = \pm |\gamma''| = \pm \kappa.$$

Es gilt also mit $|\beta| = \kappa$

$$\begin{pmatrix} t \\ n \end{pmatrix}' = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix} \begin{pmatrix} t \\ n \end{pmatrix}.$$

Das sind die *Frenetschen Gleichungen*, benannt nach J.F. FRENET (1816–1900), und es ist

$$\beta = (t', n) = x'y'' - x''y'.$$

Man schreibt auch

$$t = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$$

und nennt $\alpha = \alpha(s)$ den *Kontingenzwinkel*. Dann ist

$$t' = \alpha' \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$$

also $\beta = \alpha'$. Durch Vorgabe der Krümmung ist die Kurve bis auf Verschiebungen und Drehungen (bis auf $\gamma(s_0)$ und $\alpha(s_0)$) eindeutig festgelegt. Das folgt aus den Frenetschen Gleichungen. Es ist

$$\begin{aligned}\alpha(s) &= \int_{s_0}^s \beta(t) dt + \alpha(s_0) \\ x(s) &= \int_{s_0}^s \cos \alpha(t) dt + x(s_0) \\ y(s) &= \int_{s_0}^s \sin \alpha(t) dt + y(s_0).\end{aligned}$$

$$K := \int_0^L \kappa(s) ds$$

heißt *Gesamtkrümmung* von $\gamma(s)$.

9.9.2 Kurven im Raum

Es sei jetzt analog zum ebenen Fall

$$\gamma(s) = (x(s), y(s), z(s))',$$

$t = \gamma'$ wieder die Tangente, für $\kappa = |\gamma''| \neq 0$

$$n := \frac{t'}{|t'|}$$

die Normale und

$$b := t \times n$$

die Binormale. Es sei $\tau := (n', b) = -(b', n)$ die *Windung* von γ . Dann lauten die Frenetschen Gleichungen im \mathbb{R}^3

$$\begin{pmatrix} t \\ n \\ b \end{pmatrix}' = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} t \\ n \\ b \end{pmatrix}.$$

Der Beweis verläuft analog zum zweidimensionalen Fall. Insbesondere ist

$$\tau = (b, n') = \left(t \times \frac{1}{\kappa} t', \frac{1}{\kappa} t'' + \left(\frac{1}{\kappa} \right)' t' \right) = \frac{1}{\kappa^2} (t \times t', t'') = \frac{1}{\kappa^2} \det(\gamma', \gamma'', \gamma''').$$

Beispiel 9.9.9: Wir betrachten die Spirale

$$\gamma(s) := \begin{pmatrix} r \cos ks \\ r \sin ks \\ cks \end{pmatrix} \quad \text{mit } k = \frac{1}{\sqrt{r^2 + c^2}}.$$

Dann ist

$$\gamma'(s) = k \begin{pmatrix} -r \sin ks \\ r \cos ks \\ c \end{pmatrix}, \quad \gamma''(s) = -rk^2 \begin{pmatrix} \cos ks \\ \sin ks \\ 0 \end{pmatrix}, \quad \gamma'''(s) = rk^3 \begin{pmatrix} \sin ks \\ -\cos ks \\ 0 \end{pmatrix},$$

also

$$\kappa(s) = |\gamma''(s)| = rk^2 = \frac{r}{r^2 + c^2}$$

und

$$\tau(s) = \frac{1}{\kappa^2(s)} \det(\gamma', \gamma'', \gamma''')(s) = \frac{c r^2 k^6}{\kappa^2} = ck^2 = \frac{c}{r^2 + c^2}.$$

9.9.3 m -dimensionale Flächen im \mathbb{R}^n

Als letztes verallgemeinern wir die erhaltenen Resultate auf m -dimensionale Flächen im \mathbb{R}^n .

Definition 9.9.10: Es sei $U \subset \mathbb{R}^m$ offen, $m < n$. Die Abbildung

$$\gamma : U \rightarrow \mathbb{R}^n$$

sei stetig differenzierbar, und die Ableitung

$$\gamma' : U \rightarrow \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$$

habe für alle $u \in U$ den Rang m . Dann heißt γ Parameterdarstellung einer m -dimensionalen Fläche im \mathbb{R}^n (m -Fläche).

Aufgrund der Rangbedingung ist für jedes feste $u \in U$

$$\tau(h) := \gamma(u) + \gamma'(u, h) \quad \text{mit } h \in \mathbb{R}^m$$

ebenfalls eine Parameterdarstellung einer m -Fläche. τ ist in h linear. τ ist die Parameterdarstellung des m -dimensionalen Tangentialraumes für γ im Punkte $\gamma(u)$.

Beispiel 9.9.11: Wir betrachten Rotationsflächen im \mathbb{R}^3 .

$$\gamma(u, v) := \begin{pmatrix} f(u) \cos v \\ f(u) \sin v \\ g(u) \end{pmatrix}$$

mit $f(u) > 0$ und $f'(u)^2 + g'(u)^2 > 0$.

Es ist

$$\gamma'(u, v) = \begin{pmatrix} f'(u) \cos v & -f(u) \sin v \\ f'(u) \sin v & f(u) \cos v \\ g'(u) & 0 \end{pmatrix}.$$

γ' hat den Rang zwei.

Definition 9.9.12: Zwei Parameterdarstellungen $\gamma : U \rightarrow \mathbb{R}^n$ und $\tilde{\gamma} : \tilde{U} \rightarrow \mathbb{R}^n$ von m -Flächen heißen äquivalent $:\Leftrightarrow$

$$\exists \Phi : \tilde{U} \rightarrow U \quad \text{Diffeomorphismus}$$

mit $\tilde{\gamma} = \gamma \circ \Phi$.

Das ist wieder eine Äquivalenzrelation. Die Flächen haben gleiche Orientierung, wenn $\det \Phi' > 0$ ist.

Eine besonders einfache Darstellung erhält man durch

$$\begin{aligned} \gamma_1(u) &= u_1 \\ &\vdots \\ \gamma_m(u) &= u_m \\ \gamma_{m+1}(u) &= f_1(u) \\ &\vdots \\ \gamma_n(u) &= f_{n-m}(u) \end{aligned}$$

mit $u \in U \subset \mathbb{R}^m$. Dabei sind γ_i die Komponenten von γ und f_i C_1 -Funktionen. Es ist

$$\gamma'(u) = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \diagdown & \vdots \\ 0 & \cdots & 1 \\ \partial_1 f_1(u) & \cdots & \partial_m f_1(u) \\ \vdots & & \vdots \\ \partial_1 f_{n-m}(u) & \cdots & \partial_m f_{n-m}(u) \end{pmatrix}$$

$\gamma'(u)$ hat überall den Rang m .

Wir wollen nun zeigen, daß jede m -Fläche lokal äquivalent ist zu einem γ dieses speziellen Typs.

Satz 9.9.13: Es sei $\gamma : U \rightarrow \mathbb{R}^n$ die Parameterdarstellung einer m -Fläche. Für ein $u_0 \in U$ gelte

$$\det \begin{pmatrix} \partial_1 \gamma_1 & \cdots & \partial_m \gamma_1 \\ \vdots & \searrow & \vdots \\ \partial_1 \gamma_m & \cdots & \partial_m \gamma_m \end{pmatrix} (u_0) \neq 0.$$

Dann gibt es eine Umgebung W von u_0 mit

$$\gamma|_W \sim \tilde{\gamma} : V \rightarrow \mathbb{R}^n$$

und

$$\begin{aligned} \tilde{\gamma}_1(v) &= v_1 & \tilde{\gamma}_{m+1}(v) &= f_1(v) \\ \vdots & & \vdots & \\ \tilde{\gamma}_m(v) &= v_m & \tilde{\gamma}_n(v) &= f_{n-m}(v). \end{aligned}$$

Beweis: Für $u \in U$ betrachten wir

$$\begin{aligned} v_1 &= \gamma_1(u) \\ &\vdots \\ v_m &= \gamma_m(u) \end{aligned} \tag{*}$$

und wenden den Satz von der lokalen Umkehrbarkeit an. Dann gibt es zu $u_0 \in U$ eine Umgebung W von u_0 , in der Gl. (*) ein Diffeomorphismus von W auf eine offene Menge V ist. Die Umkehrabbildung davon sei

$$\Phi : V \rightarrow W.$$

Dann ist $\gamma \circ \Phi$ äquivalent zu $\gamma|_W$ und hat die angegebene Gestalt.

Beispiel 9.9.14: Wir betrachten die stereographische Projektion

$$\gamma(u, v) = \frac{1}{1 + u^2 + v^2} \begin{pmatrix} 2u \\ 2v \\ 1 - u^2 - v^2 \end{pmatrix}.$$

Dies ist eine Abbildung der Ebene auf die *gelochte* Kugeloberfläche mit

$$\gamma(0, 0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad \lim_{\alpha \rightarrow \infty} \gamma(\alpha, \alpha) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix},$$

also

$$\gamma : \mathbb{R}^2 \longrightarrow S^2 \setminus \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Es gilt

$$\gamma'(u, v) = \frac{2}{(1 + u^2 + v^2)^2} \begin{pmatrix} 1 - u^2 + v^2 & -2uv \\ -2uv & 1 + u^2 - v^2 \\ -2u & -2v \end{pmatrix}.$$

γ' hat überall den Rang zwei.

Es seien nun

$$x = \frac{2u}{1 + w^2}, \quad y = \frac{2v}{1 + w^2}, \quad z = \frac{1 - w^2}{1 + w^2}$$

mit $w^2 := u^2 + v^2$. Dann erhält man

$$x^2 + y^2 = \frac{4w^2}{(1 + w^2)^2}$$

und

$$z = \pm \sqrt{1 - x^2 - y^2}.$$

Etwa für $(u_0, v_0) = (0, 0)$ ist

$$\det \begin{pmatrix} \partial_1 \gamma_1 & \partial_2 \gamma_1 \\ \partial_1 \gamma_2 & \partial_2 \gamma_2 \end{pmatrix} (0, 0) = 1,$$

die Voraussetzung des Satzes also erfüllt. Die äquivalente Darstellung (x, y, z) gilt nur lokal, und man hat die Wahl des Vorzeichens von z .

Lokal läßt sich eine m -Fläche also in der Form

$$\begin{aligned}x_{m+1} &= f_{m+1}(x_1, \dots, x_m) \\ &\vdots \\ x_n &= f_n(x_1, \dots, x_m)\end{aligned}$$

schreiben, oder implizit als Nullstellenmenge von

$$\begin{aligned}0 &= g_1(x) = x_{m+1} - f_{m+1}(x_1, \dots, x_m) \\ &\vdots \\ 0 &= g_{n-m}(x) = x_n - f_n(x_1, \dots, x_m).\end{aligned}$$

Damit beenden wir erst einmal die Behandlung der m -Flächen. Die Bestimmung des Flächeninhalts werden wir vornehmen, nachdem wir in §10–12 die Integration weitergetrieben haben.

9.10 Extrema unter Nebenbedingungen

In den Anwendungen tritt häufig folgendes Problem auf: Es soll das Extremum von

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

unter $m < n$ Nebenbedingungen

$$g_1 = g_2 = \dots = g_m = 0$$

bestimmt werden, mit

$$g_j : \mathbb{R}^n \longrightarrow \mathbb{R}.$$

Zur Lösung wird man zunächst daran denken, aus den Nebenbedingungen etwa die letzten der x_j zu eliminieren und diese dann in f einzutragen.

Beispiel 9.10.1: *Es seien*

$$\begin{aligned}f : \mathbb{R}^2 &\rightarrow \mathbb{R}, & f(x, y) &= x^2 + y^2, \\ g : \mathbb{R}^2 &\rightarrow \mathbb{R}, & g(x, y) &= x + y - 1.\end{aligned}$$

Dann ist

$$f'(x, y) = 2 \begin{pmatrix} x \\ y \end{pmatrix}.$$

Es ist also $f'(x, y) = 0$ für $x = y = 0$. Der Nullpunkt ist eine Minimalstelle von f . Diese Stelle kommt aber wegen $g(0, 0) = -1$ nicht in Betracht. Eliminieren wir y aus $g(x, y) = 0$, dann folgt $y = 1 - x$ und

$$h(x) := f(x, 1 - x) = x^2 + (1 - x)^2 = 2x^2 - 2x + 1.$$

Jetzt läßt sich das Extremum von h wie gewohnt berechnen. Es ist $h'(x) = 4x - 2$ und $h''(x) = 4$. Man erhält also ein Minimum an der Stelle

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_0 \\ 1 - x_0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Wir wollen noch etwas anders an diese Aufgabe herangehen: Aus $g(x, y) = 0$ folge die Auflösung

$$x = \varphi(t) \quad \text{und} \quad y = \psi(t) \quad \text{mit} \quad t \in [a, b] \quad \text{und} \quad (\varphi', \psi') \neq 0.$$

Es sei

$$h(t) = f(\varphi(t), \psi(t)).$$

Dann suchen wir ein t_0 mit

$$h'(t_0) = \left(\partial_1 f \cdot \varphi' + \partial_2 f \cdot \psi' \right) \Big|_{t=t_0} \stackrel{!}{=} 0.$$

Wegen $g(\varphi(t), \psi(t)) = 0$ ist auch für alle t

$$g'(t) = (\partial_1 g \cdot \varphi' + \partial_2 g \cdot \psi')(t) = 0,$$

also

$$\begin{pmatrix} \partial_1 f & \partial_2 f \\ \partial_1 g & \partial_2 g \end{pmatrix} \begin{pmatrix} \varphi' \\ \psi' \end{pmatrix} \Big|_{t=t_0} \stackrel{!}{=} 0.$$

An der gesuchten kritischen Stelle t_0 von h sind deshalb die Vektoren ∇f und ∇g linear abhängig. Es gibt also eine Zahl λ_0 mit

$$(\nabla f + \lambda_0 \nabla g)(t_0) = 0.$$

Außerdem gilt natürlich $g = 0$. Aus diesen Gleichungen lassen sich $x_0 = x(t_0)$ und $y_0 = y(t_0)$ direkt bestimmen, ohne vorher nach t aufgelöst zu haben.

Setzen wir allgemein (im Beispiel ist $n = 2$ und $m = 1$)

$$\begin{aligned} F : \mathbb{R}^{n+m} &\longrightarrow \mathbb{R} \\ (x, \lambda) &\longmapsto f(x) + \sum_{i=1}^m \lambda_i g_i(x), \end{aligned}$$

dann ist

$$\nabla F = \begin{pmatrix} \nabla f + \sum_{i=1}^m \lambda_i \nabla g_i \\ g_1 \\ \vdots \\ g_m \end{pmatrix},$$

und unsere Aufgabe ist äquivalent zu

$$\nabla F \stackrel{!}{=} 0.$$

Das ist die notwendige Bedingung für eine Extremalstelle. Der Vorteil hierbei ist, daß man sich zu ihrer Formulierung vorher nicht um geeignetes Eliminieren kümmern muß. Im Beispiel erhalten wir

$$F(x, y, \lambda) = x^2 + y^2 + \lambda(x + y - 1)$$

und

$$\nabla F(x, y, \lambda) = \begin{pmatrix} 2x + \lambda \\ 2y + \lambda \\ x + y - 1 \end{pmatrix}.$$

$\nabla F \stackrel{!}{=} 0$ liefert unmittelbar

$$x_0 = y_0 \rightsquigarrow x_0 = y_0 = 1/2 \text{ und } \lambda_0 = -1,$$

also

$$\begin{pmatrix} x_0 \\ y_0 \\ \lambda_0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Den Vektor $\lambda \in \mathbb{R}^m$ nennt man Lagrangeschen Multiplikator. Wir formulieren das Ergebnis als

Satz 9.10.2: *Es seien $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ stetig differenzierbar, desgleichen $g_i : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$. Die Teilmenge $M \subset U$ sei durch $g_1 = g_2 = \dots = g_m = 0$ definiert. Die Ableitung von $g := (g_1, \dots, g_m)$ habe überall den Rang m , das heißt M sei eine $(n - m)$ -Fläche im \mathbb{R}^n . Es sei $\lambda \in \mathbb{R}^m$. Hat dann $f|_M$ in $x_0 \in M$ ein lokales Extremum, dann gibt es ein λ_0 , so daß $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$,*

$$F(x, \lambda) := f(x) + (\lambda, g)$$

in (x_0, λ_0) eine kritische Stelle hat (das heißt, es ist $\nabla F(x_0, \lambda_0) = 0$).

Damit haben wir eine notwendige Bedingung für die Existenz von Extremalstellen unter Nebenbedingungen erhalten. Möchte man auch hinreichende Bedingungen finden, dann muß man feststellen, ob

$$F''(x_0, \lambda_0; h, h)$$

definit ist. Das muß man aber nicht für alle $h \in \mathbb{R}^{n+m}$ überprüfen, denn λ_0 ist ja fest und nur x_0 wird variiert. Es genügt also, $h = (h_1, \dots, h_n, 0, \dots, 0)$ zu nehmen. Selbst diese Menge der h ist noch zu groß, weil auch die Nebenbedingungen $g_i(x_0 + h) = 0$ erfüllt sein müssen. Es soll also

$$F''(x_0, \lambda_0; h, h)$$

definit sein für alle $h \in H$ mit

$$H = \{h \in \mathbb{R}^{n+m} \mid h = (h_1, \dots, h_n, 0, \dots, 0) \neq 0 \text{ und } g_1(x_0 + h) = \dots = g_m(x_0 + h) = 0\}.$$

In unserem Beispiel ist

$$F''(x, \lambda) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Für alle $h = (h_1, h_2, 0)$ ist F'' deshalb positiv definit, und wir erhalten ein Minimum.

9.10.1 Quadratische Formen

Betrachten wir als Anwendungsbeispiel quadratische Formen. Es sei $A = A^t$ eine symmetrische $n \times n$ -Matrix, $A = (a_{ij})$, und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(x) := (x, Ax).$$

Es sei $g_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$g_1(x) := 1 - |x|^2.$$

Wir suchen Extremalstellen von f auf der Sphäre $g_1 = 0$.

Dazu wählen wir mit $\lambda \in \mathbb{R}$,

$$F_1(x, \lambda) := (x, Ax) + \lambda(1 - |x|^2)$$

und erhalten

$$\nabla F_1(x, \lambda) = \begin{pmatrix} 2(A - \lambda id)x \\ 1 - |x|^2 \end{pmatrix}.$$

Wegen der Symmetrie von A ist nämlich

$$\partial_i a_{jk} x_j x_k = a_{jk} \delta_{ji} x_k + a_{jk} x_j \delta_{ki} = 2 a_{ij} x_j.$$

Damit haben wir unsere Aufgabe auf ein Eigenwertproblem für A zurückgeführt. Wegen der Symmetrie sind alle Eigenwerte von A reell, und es gibt genau n Eigenwerte $\kappa_1, \dots, \kappa_n$ mit dazugehörigen orthonormalen Eigenvektoren u_1, \dots, u_n , also $Au_i = \kappa_i u_i$. Wir denken uns die Eigenwerte in der Form

$$\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_n$$

angeordnet. Dann ist für $i = 1, 2, \dots, n$

$$\nabla F_1(u_i, \kappa_i) = 0$$

und

$$F_1(u_i, \kappa_i) = f(u_i) = (u_i, Au_i) = \kappa_i.$$

Demnach kommt u_1 als Maximalstelle mit $f(u_1) = \kappa_1$ in Betracht.

Wir wollen auch die hinreichende Bedingung diskutieren. Es ist

$$F''_1(x, \lambda) = \begin{pmatrix} 2(A - \lambda id) & -2x \\ -2x^t & 0 \end{pmatrix}.$$

Überprüfen wir $(h, F''(u_1, \kappa_1)h)$ für alle h mit $h = (\dots, 0)$ und $|u_1 + h| = 1$. Es sei $h = \sum_{i=1}^n a_i u_i \neq 0$. Dann folgt

$$(h, F''_1(u_1, \kappa_1)h) = 2(h, (A - \kappa_1)h) = 2 \sum_{i=2}^n (\kappa_i - \kappa_1) a_i^2,$$

und aus $|u_1 + h| = 1$ erhalten wir

$$2a_1 + \sum_1^n a_i^2 = 0,$$

also $a_1 < 0$ und damit $\sum_{i=2}^n a_i^2 = -2a_1 - a_1^2 > 0$ für kleine a_1 . Ist nun $\kappa_1 > \kappa_2$, dann ist für diese h

$$(h, F_1''(u_1, \kappa_1)h) < 0.$$

Die hinreichende Bedingung für ein Maximum ist dann also ebenfalls erfüllt.

Wir wollen die Bedeutung der anderen Eigenwerte noch etwas besser verstehen: Dazu wählen wir

$$g_2(x) = (x, u_1)$$

als zweite Nebenbedingung. Dann ist

$$F_2(x, \lambda) := (x, Ax) + \lambda_1(1 - |x|^2) + \lambda_2(x, u_1)$$

und

$$\nabla F_2(x, \lambda) = \begin{pmatrix} 2(A - \lambda_1 id)x + \lambda_2 u_1 \\ 1 - |x|^2 \\ (x, u_1) \end{pmatrix}.$$

Für die kritischen Stellen erhalten wir

$$\lambda_2 = -2((A - \lambda_1 id)x, u_1) = 0;$$

also wieder die Eigenwertaufgabe

$$(A - \lambda id)x = 0,$$

aber jetzt mit zwei Nebenbedingungen, nämlich

$$|x| = 1 \quad \text{und} \quad (x, u_1) = 0.$$

Analog folgt für ein beliebiges $j > 1$: Es ist

$$\max_{|x|=1} (x, Ax) = \kappa_j,$$

und zwar unter den Nebenbedingungen

$$(x, u_1) = \dots = (x, u_{j-1}) = 0.$$

Bezeichnungen

| | |
|--|---|
| \equiv | $f \equiv y: \forall x \ f(x) = y$ |
| \subseteq | $A \subseteq B: A$ ist eine kompakte Teilmenge von B |
| δ_{ij} | Kroneckersymbol: $\delta_{ij} = 1$ für $i = j$, sonst 0 |
| $d_X(a, b)$ | Abstand von $a, b \in X$ (Metrik) |
| \bar{M} | Abschluß der Menge M |
| M^0 | Inneres der Menge M |
| $D(f), R(f)$ | Definitions- und Wertebereich der Abbildung f |
| $A \rightarrow B$ | von A wird in B abgebildet |
| $a \mapsto b$ | das Element a wird auf b abgebildet |
| \mathbb{C} | Komplexe Zahlen |
| \mathbb{R} | Reelle Zahlen |
| \mathbb{K} | bei linearen Räumen der zugrundeliegende Zahlkörper |
| $\ x\ _p$ | $\{\sum_{i=1}^{\infty} x_i ^p\}^{1/p}$, im Spezialfall $p = 2$ spricht man von der euklidischen Norm $ x $ |
| $\ \cdot\ _{\infty}$ | Supremumsnorm von $\mathbb{R}^n: \ x\ _{\infty} = \max(x_1 , x_2 , \dots, x_n)$ |
| ℓ^p | Folgenraum: $\ell^p = \{(x_k)_{k \in \mathbb{N}} \mid \forall i \ x_i \in \mathbb{C}, \sum_{i=1}^{\infty} x_i ^p < \infty\}$ |
| ℓ^{∞} | Raum der beschränkten unendlichen Folgen |
| $\ \cdot\ _{\ell^p}$ | p -Norm von $\ell^p: \ x\ _{\ell^p} = \{\sum_{i=1}^{\infty} x_i ^p\}^{1/p}$ |
| $\ \cdot\ _{\ell^{\infty}}$ | Supremumsnorm von $\ell^{\infty}: \ x\ _{\ell^{\infty}} = \sup_n x_n $ |
| (\cdot, \cdot) | Skalarprodukt von $\ell^2: (x, y) = \sum_{i=1}^{\infty} x_i \bar{y}_i$ |
| $\mathcal{F}(D, R)$ | Finite Abbildungen von D in R |
| $\mathcal{B}(D, R)$ | Beschränkte Abbildungen von D in R |
| $\mathcal{L}(D, R)$ | Lineare Abbildungen von D in R |
| $\mathcal{L}(D, \mathbb{K})$ | Lineare Funktionale: lineare Abbildungen von D in \mathbb{K} |
| $\mathcal{L}_b(D, R)$ | Beschränkt lineare Abbildungen von D in R |
| $\ \cdot\ _{\mathcal{L}_b}$ | Operatornorm von $\mathcal{L}_b(D, R): \ A\ _{\mathcal{L}_b} = \sup_{ x =1} \ Ax\ $ |
| $\mathcal{T}(D, R)$ | Treppenfunktionen von D in R |
| $C(D, R)$ | Stetige Abbildungen von D in R |
| $C^*(D, R)$ | gleichmäßig stetige Abbildungen von D in R |
| $\mathcal{BC}(D, R)$ | $\mathcal{B}(D, R) \cap C(D, R)$ |
| $\ \cdot\ = \ \cdot\ _{\mathcal{L}^{\infty}}$ | Supremumsnorm: $\ f\ = \sup_{x \in D} f(x) $ |
| $C_n(D, R)$ | n -mal stetig differenzierbare Abbildungen von D in R |
| $C_{\infty}(D, R)$ | beliebig oft stetig differenzierbare Abbildungen von D in R |
| $\mathring{C}_{\infty}(G, R)$ | Testfunktionen: beliebig oft differenzierbare Abbildungen mit in G kompaktem Träger |
| $\mathcal{L}^p(D, R)$ | Vervollständigung von $\mathcal{T}(D, R)$ bzgl. der \mathcal{L}^p -Norm |
| $\ \cdot\ _{\mathcal{L}^p}$ | \mathcal{L}^p -Norm: $\ f\ _{\mathcal{L}^p} = \{\int_D f(x) ^p dx\}^{1/p}$ |
| $\ \cdot\ _{\mathcal{L}^{\infty}}$ | Supremumsnorm: $\ f\ _{\mathcal{L}^{\infty}} = \sup_{x \in D} f(x) $ |
| $(\cdot, \cdot)_{\mathcal{L}^2}$ | Skalarprodukt von $\mathcal{L}^2: (f, g)_{\mathcal{L}^2} = \int_D f(x) \overline{g(x)} dx$ |
| ∂_t | $\frac{\partial}{\partial t}$ |
| ∂_x | $\frac{\partial}{\partial x}$ |
| ∂_i | $\frac{\partial}{\partial x_i}$ |
| $\nabla = \text{grad}$ | $(\partial_1, \dots, \partial_n)^t$, der Nabla-Operator |
| $O(x)$ | Landau-Symbol. $f = O(g)$ für $t \rightarrow \infty: \exists c > 0 \forall t > t_0 \ f(t) \leq c g(t) $ |
| $o(x)$ | Landau-Symbol. $f = o(g)$ für $t \rightarrow \infty: \lim_{t \rightarrow \infty} f(t)/g(t) = 0$ |
| $A \oplus B$ | Direkte Summe der Teilräume $A, B. \forall x \in A \oplus B$ gilt $\exists \overset{!}{a} \in A \exists \overset{!}{b} \in B \ x = a + b$ |
| $V \otimes W$ | Äußere Summe zweier \mathbb{K} -Vektorräume $V, W. V \otimes W := \{(v, w) \mid v \in V, w \in W\}$ |

Index

Druck der Seitenzahlen: Es bedeuten für das Stichwort

- schräg* : Personennamen
- fett** : definierendes Vorkommen
- von–bis : Abschnittsüberschriften
- normal : angewandtes Vorkommen

α -Limesmenge, **29**

Abbildung

- abgeschlossene lineare, **68**
- abschließbare lineare, **68**
- beschränkt lineare, **67**
- bilinare, 82
- differenzierbare, **75, 77, 75–80**
- global umkehrbare, 91
- Gradient, **76**
- lineare, 67–72
- lokal umkehrbare, 91
- Umkehr-, 80, 89, 91, 93, 96, 101
- unbeschränkt lineare, **68**
- abgeschlossene lineare Abbildung, **68**
- abgeschlossener Graph, 68
- abgeschlossener Operator, 68
- abgeschlossener Teilraum, 69
- Ableitung, **75**
 - höhere, 82–89
 - linksseitige, **29**
 - partielle, **76, 81, 82**
 - rechtsseitige, **29**
 - Richtungs-, **76, 78**
- abschließbare lineare Abbildung, **68**
- abschließbarer Operator, 68
- Abschluß, **68**
- Abstand, **69**
- äquivalenter Weg, 96
- Äquivalenzklassen
 - stückweise glatter Wege, **96**
- allgemeine Lösung, 26
- Alternative
 - Fredholmsche, **72**
- Anfangsbedingungen, 3, 61
- Anfangswertaufgabe, **37**
- Anfangszustand, 3
- Approximation, 16
- Approximationssatz, **69, 70, 73**
- arithmetisch–geometrisches Mittel
 - Abschätzung, **86**
- Arzelà, 15
- Ascoli, 15
- asymptotisch stabiler Punkt, **30, 31, 32, 36**
- Attraktor
 - negativer, **21**
 - positiver, **21**
- Auswahlsatz, **15**

autonome Differentialgleichung, **6**
 autonomes System, **6, 24**

- Banachraum, 72, 75
- Banachscher Fixpunktsatz, 11, 52, 90
- Bernoullische Differentialgleichung, **25, 26**
- beschränkt lineare Abbildung, **67**
- beschränkt lineares Funktional, 70
- Besselfunktion, **14**
- Besselsche Gleichung, **14**
- Besselsche Ungleichung, **46**
- Bewegung, **5**
- Bewegungsgleichung, 59, **61**
 - Newtonsche, 7
- Bifurkationsdiagramm, 36
- Bifurkationspunkt, **36**
- bilinare Abbildung, 82
- Binormale, **99**
- Bogenlänge einer Kurve, **97**
- Brennpunkt, **21**
- $C_1([a, b], \mathbb{R}^n)$ –Supremumsnorm, **96**
- Cauchy–Schwarzsche Ungleichung, 87
- Cauchyfolge, 69
- Courant, 45
- Courantsches Mini–Max–Prinzip, **45**
- Darstellungssatz
 - Rieszscher, **70, 76**
- Determinante
 - Wronskische, **19, 22, 40**
- Diagonalfolge, 15
- Diffeomorphismus, **34, 80, 100**
- Differentialgleichung, **5**
 - autonome, **6**
 - Bernoullische, **25, 26**
 - Besselsche, **14**
 - des beschränkten Wachstums, **27**
 - gewöhnliche, 3–63
 - homogene, **25**
 - höherer Ordnung, **22**
 - lineare, 14, 18, 23, 26, 33
 - logistische, **27**
 - nichtautonome, **6, 10**
 - nichtlineare, 14
 - parameterabhängige, 35
 - partielle, 3
 - Riccatische, **14, 26**
 - separable, **24**
- Differentialoperator, 10, 37, 68
 - Normalform, 37
- Differenzenquotient, 88
- differenzierbare Abbildung, **75, 77, 75–80**
- differenzierbare Funktion, **75, 77**
- differenzierbarer Weg, 81, **95**
- Differenzierbarkeit, 11, **75, 77, 88**
 - der Umkehrabbildung, **79, 91**

- Dini, 47
 direkte Summe, 70
 Dirichletsche Form, 38
 Dirichletsche Randbedingung, 37
 Dirichletsche Randwertaufgabe, 37
 Divergenz, 79
 divergenzfreies Vektorfeld, 35
 Dreiecksungleichung, 96
- ebene Kurven, 98–99
 Eigenfrequenz, 37
 Eigenfunktion, 37, 38, 41
 Eigenvektor, 20, 21, 104
 Eigenwert, 30, 34, 37, 38, 40, 44, 48, 86, 104
 Eigenwertaufgabe, 37, 41–45, 51, 105
 Eindeutigkeitsnachweis, 13
 Eindeutigkeitspunkt
 - globaler, 29
 - lokaler, 29
- Ellipse, 97
 elliptisches Integral, 9, 97
 Energieintegral, 9
 Entwicklungssätze, 46–49
 Erdbeschleunigung, 3
 Erdmasse, 7
 Erdradius, 7
 Erhaltungsgröße, 8, 25
 erstes Integral, 8, 25, 32, 35, 59
 Euler–Verfahren
 - explizites, 17
 - verbessertes, 17
- exakte Lösung, 4
 exaktes Vektorfeld, 25
 Existenz- und Eindeigkeitssätze, 10–18
 explizites Euler–Verfahren, 17
 Extremalstelle einer Funktion, 86
 - hinreichende Bedingung, 103
 - notwendige Bedingung, 103
- Falllinien, 78
 Fehlerabschätzung, 11
 Fixpunkt, 5, 8, 20, 27, 30, 34
 Fixpunktgleichung, 11, 52
 Fixpunktsatz
 - Banachscher, 11, 52
- Flächen, 95–102
 Fluchtgeschwindigkeit, 8
 Fokus, 21
 Fourierkoeffizient, 46, 50
 Fourierreihe, 37, 38, 51
 Fouriertransformation, 52
 Fredholmsche Alternative, 72
 freier Fall, 3
 Frenet, 98
 Frenetsche Gleichungen, 98
 Fundamentalsatz der Algebra, 34
 Funktion
 - abgeschlossene lineare, 68
 - Ableitung, 75
 - abschließbare lineare, 68
 - beschränkt lineare, 67
 - Bessel–, 14
 - differenzierbare, 75, 77
 - Eigen–, 37, 38, 41
 - Extremalstelle, 86
 - hinreichende Bedingung, 103
 - notwendige Bedingung, 103
 - ganzrationale, 75
 - gleichmäßig stetige, 66
 - global invertierbare, 91
 - Gradient, 76
 - Greensche, 40, 39–40, 42, 47, 51
 - implizite, 92–94
 - konvexe, 88, 91
 - kritische Stelle, 85, 86
 - lokal invertierbare, 91
 - lokale Extremalstelle, 85
 - Lyapunov–, 32, 36
 - Maximalstelle, 86, 89, 104
 - Minimalstelle, 86, 102
 - stetig differenzierbare, 81
 - stetige, 64, 64–67
 - streng konvexe, 88
 - strenge Lyapunov–, 32
 - Test–, 12, 46
 - unbeschränkt lineare, 68
 - vektorwertige, 81, 84
 - von mehreren Veränderlichen, 64–106
 - Zylinder–, 14
- Funktional
 - beschränkt lineares, 70
 - lineares, 67
- Funktionenfolge
 - Cauchy–Konvergenz, 73
 - gleichartig stetige, 15
 - gleichmäßig beschränkte, 15
 - gleichmäßig gleichartig stetige, 15
 - gleichmäßig konvergente, 73
- ganzrationale Funktion, 75
 Gaußsches Verfahren
 - des arithmetisch–geometrischen Mittels, 9
 - zur Berechnung elliptischer Integrale, 9, 61
- gedämpfte Schwingung, 14, 23, 32
 gedämpftes mathematisches Pendel, 36
 Gesamtkrümmung, 99
 Geschwindigkeit
 - Flucht–, 8
 - Licht–, 8
 - Schall–, 8
- gestörtes lineares System, 30
 gestörtes System, 33
 gewichtete $\mathcal{L}^2(I)$ –Norm, 38
 gewöhnliche Differentialgleichung, 3–63
 glatter Weg, 95
 - Länge, 95
- gleichartig stetige Funktionenfolge, 15
 gleichmäßig beschränkte Funktionenfolge, 15

- gleichmäßig gleichartig stetige Fktenfolge, **15**
- gleichmäßig konvergente Funktionenfolge, **73**
- gleichmäßig konvergente Teilfolge, 16
- gleichmäßig stetige Funktion, **66**
- gleichmäßige Konvergenz, 72–75
- Gleichung
 - Besselsche, **14**
 - Bewegungs–, 59, **61**
 - Fixpunkt–, 11, 52
 - Frenetsche, **98**
 - für das ungedämpfte Pendel, 18
 - für elektrischen Oszillator, **62**
 - für gedämpfte Schwingungen, **32**
 - Helmholtzsche Schwingungs–, **37**
 - Integral–, 10, 39, 50
 - Kegelschnitt–, 60
 - logistische, **27**
 - mit konstanten Koeffizienten, **23**
 - Newtonsche Bewegungs–, **7**
 - Normal–, **87**
 - Parallelogramm–, **69**
 - Parsevalsche, **47**
 - Satelliten–, **61**
 - von Volterra–Lotka, **28**
 - Wellen–, **37**
- Gleichungssystem, 92
 - Auflösung, 92
- global invertierbare Funktion, 91
- globale Lipschitzbedingung, **6, 10, 13, 17, 20**
- globale Lösung, **12**
- globale Umkehrbarkeit, 91
- globaler Eindeutigkeitspunkt, **29**
- Gradient, 25, **76**
 - anschauliche Deutung, **78**
- Gradientenfeld, 25, 79
- Gradientensystem, **33, 36**
- Gravitationsfeld, **59**
- Gravitationskonstante, 7, 59
- Green, 39
- Greensche Funktion, **40, 39–40, 42, 47, 51**
- Grenzpunkt, **29**
- Grenzpunktmenge, 29
- Grenzyklus, 28
- Gronwallsches Lemma, **12, 31**
- Hamiltonsches Integralprinzip, 10
- Hauptvektor, 21
- Helmholtzsche Schwingungsgleichung, **37**
- Hilbertraum, **69**
- Höhenlinien, **78**
- höhere Ableitung, 82–89
- homogene Differentialgleichung, **25**
- homogenes System, 18
- implizite Funktion, 92–94
- Impulsmoment, **59**
- Index, 36
 - eines singulären Punktes, **34**
 - eines Vektorfeldes, **34**
 - lokaler, **34**
- instabiler Punkt, **30, 31**
- Integral
 - elliptisches, 9, 97
 - Energie–, 9
 - erstes, **8, 25, 32, 35, 59**
- Integralgleichung, 10, 39, 50
- Integralkurve, **5**
- Integraloperator, 10, **39**
- Integralprinzip
 - Hamiltonsches, 10
- Isochrone, **9**
- isolierter Punkt, 36
- Jacobi, 76
- Jacobi–Matrix, **76**
- Jordankurve, **34**
- Jordansche Normalform, 20, 33
- Kegelschnittgleichung, 60
- Kettenregel, **77**
- Knoten, **21, 30, 36**
- kompakter Träger, 12
- Kontingenzwinkel, **98**
- Kontraktion, 11
- konvergente Teilfolge, 15, 71
- Konvergenz
 - gleichmäßige, 72–75
- Konvergenzordnung, 17
- konvexe Funktion, **88, 91**
- konvexe Teilmenge, **70, 88**
- Koordinatensystem
 - mitrotierendes, **60**
- kritische Stelle einer Funktion, **85, 86**
- Krümmung, **97**
- Krümmungskreis, **98**
- Krümmungsvektor, **97**
- Kurven, 95–102
 - ebene, 98–99
 - im Raum, 99
- $\mathcal{L}^2(I)$ –Norm, 49
 - gewichtete, **38**
- labiler Punkt, **30, 31**
- Lagrangescher Multiplikator, **103**
- Laplace–Operator, **79**
- Lebensdauer einer Lösung, **13, 14**
- Legendresche Polynome, 51
- Lejeune–Dirichlet, 37
- Lemma
 - Gronwallsches, **12, 31**
- Lichtgeschwindigkeit, **8**
- Lindelöf, 10
- linear abhängige Lösungen, **19**
- lineare Abbildung, 67–72
- lineare Differentialgleichung, 14, 18, 23, 26, 33
- linearer Operator, 68
- lineares Funktional, **67**
- lineares System, 18–24, 30, 34
- linksseitige Ableitung, **29**

- Liouville, 37
- Lipschitzbedingung, 8, **12**, 13, 15
 - globale, **6**, **10**, 13, 17, 20
 - lokale, **6**
- Lipschitzkonstante, 17, 52
- lokal invertierbare Funktion, 91
- lokale Extremalstelle einer Funktion, 85
- lokale Lipschitzbedingung, **6**
- lokale Lösung, 6, **12**, 13
- lokale Umkehrbarkeit, 89–92
- lokaler Eindeutigkeitspunkt, **29**
- lokaler Index, **34**
- Lösung
 - allgemeine, 26
 - exakte, 4
 - globale, **12**
 - Lebensdauer, **13**, 14
 - lokale, 6, **12**, 13
 - maximale, **28**, 28–29
 - partikuläre, 19, 20, 24, 26
 - periodische, 34–35
 - singuläre, 39
 - stetig differenzierbare, 12
 - stetige, 11
- Lösungsmethoden, 18–28
- Lotka, 28
- Lyapunov, 30
- Lyapunov-Funktion, **32**, 36
 - strenge, **32**
- Matrix
 - Jacobi-, **76**
- maximale Lösung, **28**, 28–29
- Maximalstelle einer Funktion, 86, 89, 104
- Maximum, 66
- Mechanik
 - klassische, 3
- Mercer, 48
- metrischer Raum
 - nichtzusammenhängender, **66**
 - relativ abgeschlossen, **64**
 - relativ offen, **64**
 - wegzusammenhängender, **67**, **81**
 - zusammenhängender, **66**
- Mini-Max-Prinzip
 - Courantsches, **45**
- Minimalstelle einer Funktion, 86, 102
- Minimum, 66
- mitrotierendes Koordinatensystem, **60**
- Mittelwertsatz, **80**, 80–81, 83
 - der Integralrechnung, 62
- Multi-Indizes, **84**
- Multiplikator
 - Lagrangescher, **103**
- Nabla-Operator, **76**
- Nagumo, 13
- Nagumobedingung, **13**, 17, 62
- negativ definite quadr. Formen, **86**
- negativ semidefinite quadr. Formen, **86**
- negativer Attraktor, **21**
- Neumann, 37
- Neumannsche Randbedingung, **37**
- Neumannsche Randwertaufgabe, 39
- Neumannsche Reihe, **43**
- Newton, 3
- Newtonsche Bewegungsgleichung, 7
- nichtautonome Differentialgleichung, **6**, 10
- nichtautonomes System, **6**, 10, 22
- nichtlineare Differentialgleichung, 14
- nichtlineare Randwertaufgabe, 52
- nichtzusammenhängender metr. Raum, **66**
- Normale, **98**, **99**
- Normalform
 - Jordansche, 33
- Normalgleichung, **87**
- normierter Vektorraum, 67
- ω -Limesmenge, **29**
- Operator
 - abgeschlossener, 68
 - abschließbarer, 68
 - Differential-, 10, 68
 - Integral-, 10
 - Laplace-, **79**
 - linearer, 68
 - Nabla-, **76**
 - Spektraldarstellung, **49**
 - unbeschränkter, 68
- Operatorenrechnung, 23
- Orbit, **5**, 21, 32
 - periodischer, **5**, 9, 28, 34
- orientierte stückweise glatte Kurve, **96**
- orthogonale Projektion, **70**
- orthogonale Zerlegung, 72
- Parallelogrammgleichung, **69**
- parameterabhängige Differentialgleichung, 35
- Parameterdarstellung
 - einer m -dimensionalen Fläche, **100**
 - äquivalente, **100**
- Parsevalsche Gleichung, **47**
- partielle Ableitung, **76**, 81, 82
- partielle Stetigkeit, **65**
- partiell Integrieren, 38
- partikuläre Lösung, 19, 20, 24, 26
- Pendel
 - gedämpftes mathematisches, **36**
 - ungedämpftes, **8**, 18, 31
 - Zykloiden-, **9**
- Periode, **35**
- periodische Lösung, 34–35
- periodischer Orbit, **5**, 9, 28, 34
- Phasenfluß, **4**, 18, 20
 - einparametrische Gruppe, 18
- Phasengeschwindigkeit, **5**
- Phasenkurve, **5**
- Phasenraum, **4**

- Phaser, 4, 17
 Picard, 10
 Polarkoordinaten, 59
 Polarkoordinatenabbildung, 80
 Polygonzug, 16
 positiv definite quadr. Formen, **86**
 positiv semidefinite quadr. Formen, **86**
 positiver Attraktor, **21**
 Potential, 33
 Potenzreihenansatz, 26
 Produktregel, 77
 Projektion
 orthogonale, **70**
 Projektionssatz, **70**, 71
 Punkt
 asymptotisch stabiler, **30**, 31, 32, 36
 Bifurkations-, **36**
 instabiler, **30**, 31
 isolierter, 36
 labiler, **30**, 31
 Sattel-, **21**, 30, 31, 34
 singulärer, **6**, 30, 34
 stabiler, **30**
 stationärer, **5**
 Strudel-, **22**
 Verzweigungs-, 7, 36
 punkttrennende Unteralgebra, **74**
- quadratische Form, **86**, 104–105
 indefinite, **86**
 negativ definite, **86**
 negativ semidefinite, **86**
 positiv definite, **86**
 positiv semidefinite, **86**
 qualitative Aspekte, 28–36
 Quelle, **21**, 34, 36
- Räuber–Beute Modell, 31
 Randbedingung, 39
 Dirichletsche, **37**
 Neumannsche, **37**
 Randwertaufgabe, **37**, 40–41
 Dirichletsche, **37**
 Neumannsche, 39
 nichtlineare, 52
 Sturm–Liouvillesche, 37–39
 rechtsseitige Ableitung, **29**
 relativ abgeschlossen, **64**
 relativ offen, **64**
 relativ zu sich selbst kompakt, 65
 Resonanzfrequenz, 23
 Riccati, 14
 Riccatische Differentialgleichung, **14**, **26**
 Richtungsableitung, **76**, **78**
 Riesz, 70
 Rieszscher Darstellungssatz, 76
 RLC–Kreis, **62**
 Rotation, **79**
 Rotationsfläche, **100**
- Runge–Kutta–Verfahren, **17**
- Satellitenbahn, **61**
 Satellitengleichung, **61**
 Sattelpunkt, **21**, 30, 31, 34
 Satz
 Approximations-, **69**, 70, 73
 Auswahl-, **15**
 Banachscher Fixpunkt-, 90
 Mittelwert-, **80**, 80–81, 83
 Projektions-, **70**, 71
 Rieszscher Darstellungs-, **70**
 Umlauf-, **34**
 vom abgeschlossenen Graphen, **68**
 von Arzelà–Ascoli, **15**, 16, 43
 von der globalen Umkehrbarkeit, 91
 von der lokalen Umkehrbarkeit, 79, 80, 89, 93,
 101
 von Dini, **47**
 von Mercer, **48**, 52
 von Peano, **15**, **16**
 von Picard–Lindelöf, **10**, 26
 von Poincaré–Bendixson, **34**
 von Stone, **73**, 74
 von Stone–Weierstraß, **74**
 von Weierstraß, 74
 Zwischenwert-, **66**, 95
- Schallgeschwindigkeit, **8**
 Schmiegebene, **97**
 schwingende Saite, 51
 Schwingung
 gedämpfte, 14, 23, 32
 Schwingungsdauer, 8, 9
 Schwingungsgleichung
 Helmholtzsche, **37**
 Schwingungstheorie, 24
 Senke, **21**, 34
 separable Differentialgleichung, **24**
 Separation der Variablen, 27
 singuläre Lösung, 39
 singulärer Punkt, **6**, 30, 34
 Singularität, 13
 Skalarprodukt, **38**
 im \mathbb{R}^n , **69**
 soziale Reibungsterme, **28**
 Spektraldarstellung
 eines Operators, **49**, 51
 spezielle Lösungsmethoden, 18–28
 Spirale, **22**
 stabiler Punkt, **30**
 Stabilität, 30–33
 stationärer Punkt, **5**
 stereographische Projektion, **101**
 stetig differenzierbare Funktion, 81
 stetig differenzierbare Lösung, 12
 stetige Funktion, **64**, 64–67
 stetige Lösung, 11
 Stetigkeit
 partielle, **65**

- Stone, 73
- Streifen, **10**
- streng konvexe Funktion, **88**
- strenge Lyapunov-Funktion, **32**
- Strudel, 30
- Strudelpunkt, **22**
- stückweise glatter Weg, **95**
 - Länge, **95**
 - Äquivalenzklassen, **96**
- Sturm, 37
- Sturm–Liouvillesche Randwertaufgabe, 37–39
- Summe
 - direkte, 70
- Support, **64**
- Supremumsnorm, **37**, 42, 48, 68, 72
 - $C_1([a, b], \mathbb{R}^n)$ –, **96**
- System
 - autonomes, **6**, 24
 - erster Ordnung, 22
 - gestörtes, 33
 - gestörtes lineares, 30
 - Gradienten–, **33**, 36
 - homogenes, 18
 - lineares, 18–24, 30, 34
 - mit konstanten Koeffizienten, **20**
 - nichtautonomes, **6**, 10, 22
 - vollständiges, 38, **46**
 - Volterra–Lotka–, **28**
- Tangente, **98**
- Tangentenvektor, **97**
- Tangentialraum, **100**
- Taylorreihe, 84
- Taylorische Formel, 75, 83
 - Integraldarstellung des Restgliedes, 84
- Testfunktion, 12, 46
- Trajektorie, **5**
- Transformation der Variablen, **25**, 27
- Träger, **64**
 - kompakter, 12
- Umkehrabbildung, 80, 91, 93, 96, 101
 - Existenzfrage, 89
- Umkehrbarkeit
 - globale, 91
 - lokale, 89–92
- Umlaufsatz, **34**
- unbeschränkt lineare Abbildung, **68**
- unbeschränkter Operator, 68
- ungedämpftes Pendel, **8**, 18, 31
- Ungleichung
 - Besselsche, **46**
 - Cauchy–Schwarzsche, 87
 - Dreiecks–, 96
- Unteralgebra
 - punktetrennende, **74**
- Variablentransformation, **25**, 27
- Variation der Konstanten, **20**, 23
- Vektorfeld, **5**, 8, 18, 28
 - divergenzfreies, 35
 - exaktes, **25**
 - Index, **34**
 - singulärer Punkt, **6**
- Vektorraum
 - der beschränkten Abbildungen, 72
 - der differenzierbaren Abbildungen, 77
 - normierter, 67
 - stetiger Funktionen, 66
 - vollständiger, 73
- vektorwertige Funktion, 81, 84
- verbessertes Euler–Verfahren, **17**
- Verhulst, 27
- Verzweigungen, 35–36
- Verzweigungspunkt, **7**, 36
- vollständiger metrischer Raum, 11
- vollständiges System, 38, **46**
- Volterra, 28
- Volterra–Lotka–System, **28**
- Wachstumsmodelle, 27–28
- Weg, **67**, **95**
 - äquivalenter, 96
 - differenzierbarer, 81, **95**
 - glatter, **95**
 - Länge, **95**
 - stückweise glatter, **95**
 - Länge, **95**
 - Äquivalenzklassen, **96**
- wegzusammenhängender metr. Raum, **67**, **81**
- Wellengleichung, **37**
- Windung, **99**
- Wronski, 19
- Wronskische Determinante, **19**, 22, 40
- Zentrum, **22**, 30, 32
- Zerlegung
 - orthogonale, 72
- Zusammenhang, **66**
- zusammenhängender metrischer Raum, **66**
- Zustand, **5**
- Zustandsraum, **4**
- Zweikörperproblem, **59**
- Zwischenwertsatz, **66**, 95
- Zykloide, **9**
- Zykloidenpendel, **9**
- Zylinderfunktion, **14**

