Gewöhnliche Differentialgleichungen

nach der Vorlesung im SS 2001
von Prof. Dr. Jürgen Elstrodt

21. Juli 2005

Verfasser:

Johannes Engel

Korrektoren:

Erik Müller

Christoph Meinerding

Auch allen anderen, die auf die eine oder andere Weise geholfen haben, sei hier ein Dank ausgesprochen.

Einleitung

0.1 Vorwort

Dieses Skript mag manche Funktion haben, allerdings mit Sicherheit nicht diejenige eines Lehrbuchs. Vielmehr möchte es eine Ergänzung zur Vorlesung sein und als Hilfe bei der Bearbeitung der Übungsaufgaben und der Klausurvorbereitung dienen. Es ist entstanden nach dem Manuskript der Vorlesung von Prof. Elstrodt im Sommersemester 2001, das Prof. Elstrodt mir freundlicherweise in der Vorlesungs-freien Zeit zur Verfügung gestellt hat.

0.2 Motivation

Unter einer Differentialgleichung versteht man eine Gleichung zwischen einer Funktion f, ihren Ableitungen $f', \ldots, f^{(n)}$ und der Variablen x vom Typ

$$F\big(x, f(x), f'(x), \ldots, f^{(n)}(x)\big) = 0.$$

Wir wollen das etwas präziser aussprechen:

0.2.1 Definition Es seien $I \subset \mathbb{R}$ ein Intervall, $G \subset \mathbb{R}^{n+2}$ $(n \geqslant 1)$, $f\colon I \to \mathbb{R}$ n-mal differenzierbar, und für $x \in I$ sei

$$\big(x, f(x), f'(x), \ldots, f^{(n)}(x)\big) \in G$$

f erfüllt eine *gewöhnliche Differentialgleichung* n-ter Ordnung , wenn es eine Funktion $F\colon G \to \mathbb{R}$, $F \neq 0$, gibt, so dass

$$F\big(x, f(x), f'(x), \ldots, f^{(n)}(x)\big) = 0 \qquad (x \in I) \tag{0.1}$$

Lässt sich (0.1) nach $f^{(n)}(x)$ auflösen in der Form

$$f^{(n)}(x) = H\big(x, f(x), f'(x), \ldots, f^{(n-1)}(x)\big) \tag{0.2}$$

mit passendem H, so heißt die Differentialgleichung *explizit* (von n-ter Ordnung).

Auf solche Differentialgleichungen wird man in vielen Bereichen der Mathematik und Physik geführt. Gerade die Physik ist hier eine reiche Quelle interessanter Beispiele: Die grundlegenden Gesetze der Physik sind ja von dem Typ, dass sie angeben, wie groß die Änderung einer physikalischen Größe ist, z. B.

Bewegungsgesetz Kraft = Impulsänderung = Masse × Beschleunigung nach dem 2. Newtonschen Gesetz.

Induktionsgesetz Induzierte Spannung = – Änderung des magnetischen Flusses, in Formeln

$$U_{\text{ind}} = \oint \mathfrak{E}\, d\Gamma = -\frac{d\Phi}{dt}$$

nach der sog. *Lenzschen Regel.*

Gesetz des radioaktiven Zerfalls Änderung der Teilchenzahl = $-\lambda\cdot$ Teilchenzahl. Dabei ist $\lambda > 0$ die Zerfallskonstante. (Schrödinger-Gleichung)

Das Ziel in der Physik besteht im allgemeinen darin, aus einer Gleichung dieses Typs und ggf. vorhandenen weiteren Bedingungen („Anfangsbedingungen") die jeweils interessierende physikalische Größe zu bestimmen. Wir betrachten diese Aufgabe hier unter rein mathematischen Aspekten und interessieren uns hier für die Aufgabe, die Lösungen von (0.1) zu bestimmen. Auch wird es eine wesentliche Aufgabe für uns sein, allgemeine Kriterien dafür zu finden, wann (0.1) eine Lösung hat (Existenzsätze), und unter welchen Bedingungen eine solche Lösung ggf. eindeutig bestimmt ist. Die Physiker wissen, dass ein vernünftig gestelltes Problem genau eine Lösung hat, und es ist Aufgabe der Mathematiker, diese plausible Grundüberzeugung durch handfeste allgemeine Kriterien zu untermauern (Eindeutigkeits-Satz). Natürlich interessieren wir uns auch für Lösungsverfahren für spezielle Differentialgleichungen.

Bezeichnungsweise: In der Theorie der gewöhnlichen Differentialgleichungen heißt die gesuchte Funktion meistens nicht f sondern $y = y(x)$, und man schreibt (0.1) in der Form

$$F\left(x, y, y', \dots, y^{(n)}\right) = 0 \qquad (x \in I)\,, \tag{0.3}$$

wobei man sich bei $y, y', \dots, y^{(n)}$ das Argument x eingetragen denken muss. Oft hat die Variable x die Bedeutung „Zeit" und wird dann in der Regel mit t bezeichnet. Natürlich braucht man sich bei der Betrachtung von Differentialgleichungen nicht auf Funktionen einer Variablen zu beschränken:

0.2.2 Definition Sei $M \subset \mathbb{R}^n$ offen, $y\colon M \to \mathbb{R}$ m-fach stetig partiell differenzierbar, $G \subset \mathbb{R}^p$ (p hinreichend groß), und für alle $x \in M$ sei

$$\left(x_1, \dots, x_n, \frac{\partial y}{\partial x_1}, \dots, \frac{\partial y}{\partial x_n}, \dots, \frac{\partial^{\nu_1+\dots+\nu_n}}{\partial x_1^{\nu_1}\cdots\partial x_n^{\nu_n}} y\right) \in G$$

$(\nu_1 + \dots + \nu_n \leqslant m)$. Genügt y einer Gleichung der Form

$$F\left(x_1, \dots, x_n, \frac{\partial y}{\partial x_1}, \dots, \frac{\partial y}{\partial x_n}, \dots, \frac{\partial^{\nu_1+\dots+\nu_n}}{\partial x_1^{\nu_1}\cdots\partial x_n^{\nu_n}} y\right) = 0$$

mit einer Funktion $F\colon G \to \mathbb{R}$, $F \neq 0$, so sagt man, y genüge einer *partiellen Differentialgleichung* der Ordnung m.

Beispiel: Für jede holomorphe Funktion $f: M \to \mathbb{C}$ ($M \subset \mathbb{R}^2 = \mathbb{C}$) genügen $u := \operatorname{Re} f$ und $v := \operatorname{Im} f$ der Potentialgleichung

$$\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) u = 0, \qquad \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) v = 0$$

Mit partiellen Differentialgleichungen werden wir uns hier gar nicht beschäftigen, da sie eine sehr umfangreiche Theorie eigener Art erfordern. Indirekt greift die Theorie der partiellen Differentialgleichungen aber auch in die Theorie der gewöhnlichen Differentialgleichungen ein, denn sie führt vermöge der Methode der Separation der Variablen auf viele interessante Beispiele von gewöhnlichen Differentialgleichungen und Probleme über gewöhnliche Differentialgleichungen.

Um einen ersten Eindruck zu geben, betrachten wir einige Beispiele:

0.2.3 Beispiel $y' = 0$ auf einem Intervall $I \subset \mathbb{R}$. Dann gibt es ein $c \in \mathbb{R}$ mit $y(x) = c$ für alle $x \in I$. Der aus Elstrodt (2002a) bekannte Beweis basiert auf dem Mittelwertsatz der Differentialrechnung. Die Differentialgleichung hat hier nicht nur eine sondern unendlich viele Lösungen. Kennt man aber die Lösung an einem Punkt $x_0 \in I$, etwa $y(x_0) = \alpha$, so ist y eindeutig festgelegt: $y(x) = \alpha$ für alle $x \in I$.

0.2.4 Beispiel (Freier Fall im homogenen Schwerefeld) Ein im Punkt $s_0 = 0$ ruhender Körper („Massenpunkt") wurde zur Zeit $t = 0$ losgelassen und falle unter dem Einfluss der Schwerkraft senkrecht nach unten im Vakuum („freier Fall"). $s(t)$ sei der Weg, den der Körper zur Zeit $t > 0$ zurückgelegt hat. $v(t)$ sei die Geschwindigkeit zur Zeit t. Dann gilt:

$$v(t) = \frac{ds}{dt}(t) =: \dot{s}(t).$$

(In der Physik wird nach Newton die Ableitung nach der Zeit durch einen über die Funktion geschriebenen Punkt bezeichnet.) $b(t)$ sei die Beschleunigung zur Zeit t, also

$$b(t) = \frac{dv}{dt}(t) = \dot{v}(t) = \ddot{s}(t).$$

Nach Voraussetzung ist nun

$$b(t) = g = \text{Fallbeschleunigung} \approx 9{,}80665\,\mathrm{m/s^2}.$$

Das gibt die Differentialgleichung

$$\ddot{s} = g. \tag{0.4}$$

$$\Rightarrow \frac{d}{dt}\left(\dot{s} - gt\right) = 0 \overset{\text{Beispiel 0.2.3}}{\Rightarrow} \dot{s} = gt + \alpha \qquad (t \geqslant 0)$$

mit geeignetem $\alpha \in \mathbb{R}$.

$$\Rightarrow \frac{d}{dt}\left(s - \frac{1}{2}gt^2 - \alpha t\right) = 0 \Longrightarrow s(t) = \frac{1}{2}gt^2 + \alpha t + \beta \qquad (t \geqslant 0)$$

mit geeignetem $\beta \in \mathbb{R}$. Das ist die allgemeine Lösung der Differentialgleichung (0.4). Welches sind nun die Werte von α und β? Offenbar ist $s(0) = \beta$ der Ort des Massenpunktes zur Zeit $t = 0$ und $\dot{s}(0) = \alpha$ die Geschwindigkeit des Massenpunktes zur Zeit $t = 0$. Wir betrachten den Fall der Anfangsbedingungen $s(0) = 0$ und $\dot{s}(0) = 0$. Dann wird $s(t) = \frac{1}{2}gt^2$ (Fallgesetz von Galileo Galilei). Man sieht durch Einsetzen: $s(1) \approx 5$, d. h. ein Sprung aus einer Höhe von 5 m dauert ungefähr 1 Sekunde, $s(2) \approx 20$, d. h. ein Sprung vom Mathematik-Gebäude dauert ca. 2 Sekunden.

0.2.5 Beispiel (Exponentielles Wachstum bzw. Abfallen) In vielen Naturgesetzen ist die Änderung einer Größe proportional zum Bestand. Das führt auf die Differentialgleichung

$$y' = \lambda y \qquad (x \in I) \tag{0.5}$$

mit einer festen Konstanten $\lambda \in \mathbb{R}$.

$$\Rightarrow \left(ye^{-\lambda x}\right)' = \left(y' - \lambda y\right) e^{-\lambda x} = 0 \qquad (x \in I)$$

$$\overset{\text{Beispiel 0.2.3}}{\Rightarrow} y(x) = \alpha e^{\lambda x} \qquad (x \in I)$$

mit einer Konstanten $\alpha \in \mathbb{R}$, und dies ist offenbar wirklich eine Lösung von (0.5). Damit haben wir alle Lösungen von (0.5). Ist z. B. $0 \in I$, so ist y durch den Anfangswerg $\alpha = y(0)$ eindeutig festgelegt. – Die Gleichung (0.5) kommt z. B. in der Natur vor bei folgenden Gelegenheiten:

(i) Barometerformel: $p(h)$ sei der Luftdruck in der Höhe h über dem Erdboden, $h \in I := [0,\infty[$, $\frac{dp}{dh} = -\rho g$, wobei $\rho = \rho(h)$ die Dichte der Luft in der Höhe h sei, g wie gehabt die Fallbeschleunigung. (Gerthsen und Meschede (2004): „Beim Anstieg um Δh ändert sich p um $-\rho g \Delta h$.") Nach der Gasgleichung ist aber $p \cdot V$ konstant (V entspricht dem Volumen), d. h. wegen $\rho = \frac{M}{V}$ (M entspricht der Masse) ist $\frac{p}{\rho} = \frac{p_0}{\rho_0}$ konstant.

$$\Rightarrow \frac{dp}{dh} = \underbrace{-\frac{p_0}{\rho_0} g}_{=\lambda=\text{const}} \cdot p \Longrightarrow p(h) = \alpha e^{-\frac{\rho_0}{p_0} gh} \qquad (h \geqslant 0)$$

mit geeignetem α. Hier ist $\alpha = p_0$ der Luftdruck am Boden, und wir erhalten die *Barometerformel*

$$p = p_0 e^{-\frac{\rho_0}{p_0} gh} \qquad (h \geqslant 0)$$

Diese Formel gilt bei einer Atmosphäre konstanter Temperatur, die in der Realität auf der Erde nicht vorliegt! Bei einer Temperatur von 0°C ($\Rightarrow \rho_0 = 1{,}292$ kg/m^3) und einem Normaldruck von ca. 1013,25 hPa liefert die Barometerformel: Bei Anstieg um 5,541 km halbiert sich der Luftdruck.

(ii) Gesetz des radioaktiven Zerfalls: Die Anzahl der in der Zeiteinheit zerfallenden Atome einer radioaktiven Substanz ist stets proportional zur Anzahl der vorhandenen (noch nicht zerfallenen) Atome $N(t)$:

$$\Delta N = -\lambda N \Delta t \qquad (\Delta t \text{ hinreichen klein})$$

Stoff	Halbwertszeit
^{238}U (Uran)	$4,47 \cdot 10^9$ Jahre
^{235}U (Uran)	$7,04 \cdot 10^8$ Jahre
^{226}Ra (Radium)	1599 Jahre
^{222}Rn (Radon)	3823 Tage
^{209}Po (Polonium)	105 Jahre
^{212}Po (Polonium)	$3,04 \cdot 10^{-7}$ Sekunden
^{214}Po (Polonium)	$0,16$ Sekunden
^{230}Th (Thonium)	$7,54 \cdot 10^4$ Jahre
^{228}Th (Thonium)	1912 Jahre
^{14}C (Kohlenstoff)	5730 Jahre
^{239}Pu (Plutonium)	24 360 Jahre

Tabelle 0.1: Einige Halbwertszeiten

wobei $\lambda > 0$ eine positive Konstante ist, die nur vom Material abhängt, nicht aber von äußeren Versuchsbedingungen wie Druck, Temperatur, chemischer Bindung etc. Wir idealisieren die Situation, nehmen N als differenzierbare reellwertige Funktion an und erhalten die Differentialgleichung

$$\frac{dN}{dt} = -\lambda N$$

mit der sog. Zerfallskonstante λ.

$$\Rightarrow N = N_0 e^{-\lambda(t-t_0)}, \quad N_0 = N(t_0),\ t_0 = \text{Anfangszeit}$$

Die Anzahl der Atome nimmt also exponentiell ab. Wie schnell die Abnahme erfolgt, hängt nur von der Materialkonstante λ ab. Wann ist nur noch die Hälfte des zur Zeit 0 vorhandenen Materials da? Ist $N = \frac{1}{2}N_0$, so ist $e^{-\lambda(t-t_0)} = \frac{1}{2}$. Sei $t - t_0 =: t_{\frac{1}{2}}$ die sog. *Halbwertszeit* , d. h. $\lambda t_{\frac{1}{2}} = \log 2 \approx 0,6931$ und damit $t_{\frac{1}{2}} = \frac{\log 2}{\lambda}$ und $\lambda = \frac{\log 2}{t_{\frac{1}{2}}}$. An Stelle der Zerfallskonstanten wird meistens die Halbwertszeit angegeben, und diese ist von Material zu Material sehr unterschiedlich (siehe Tabelle 0.1). Nach dem Kohlenstoff-Isotoph ^{14}C ist die sog. C^{14}-Methode zur Altersbestimmung benannt, für die Willard Frank Libby 1960 den Nobelpreis für Chemie erhielt. Nach 10 Halbwertszeiten ist vom ursprünglich vorhandenen Material noch $2^{-10} = \frac{1}{1024} \approx 1‰$vorhanden. – Die natürliche Radioaktivität des Urans kann z. B. zur Altersbestimmung der Erde verwandt werden: Man bestimmt die Anzahl N der in einem (möglichst alten) Stück Fels enthaltenen ^{238}U. Sodann braucht man N_0, die Anzahl der zur Entstehungszeit der Erde vorhandenen Atome ^{238}U. Diese Anzahl lässt sich über das (stabile) Zerfallsprodukt ^{206}Pb (Blei), das im Fels steckt, näherungsweise ermitteln. Da die Halbwertszeit bekannt ist, findet man: Alter der Erde $\approx 4,6 \cdot 10^9$ Jahre. Selbst bei der Altersbestimmung von Gemälden leistet das radioaktive Zerfallsgesetz gute Dienste; so konnte schon manche Fälschung als solche entlarvt werden. Das trifft z. B. auf die berühmten Fälschungen von Han van

Meegeren zu. Dieses Beispiel und viele andere Beispiele aus der Praxis findet man in Braun (1975).

Exponentielles Wachstum einer Bakterienpopulation: Sei $B(t)$ die Anzahl der Bakterien zur Zeit t. Für kleine Δt ist $\Delta B(t) \approx \alpha B(t)\Delta t$ mit $\alpha > 0$. Wenn wir idealisieren und $B(t)$ als differenzierbare Funktion annehmen, erhalten wir die Differentialgleichung

$$B'(t) = \alpha B(t) \qquad \left(t \in I \overset{\text{z.B.}}{=} [0,\infty[\right)$$

und diese hat die Lösung $B(t) = B_0 e^{\alpha t}$ mit $B_0 = B(0)$. Die Bakterienpopulation wächst also exponentiell rasch an, und das kann auch in der Natur sehr rasch gehen, wie man z. B. an Zeitrafferaufnahmen am Mikroskop eindrucksvoll sehen kann. Nach einer bestimmten Zeit t_2, der Verdoppelungszeit, werden doppelt so viele Bakterien vorhanden sein wie zur Zeit 0: $2B_0 = B_0 e^{\alpha t_2}$, d. h. $t_2 = \frac{\log 2}{2}$. Nach der Zeit $10t_2$ hat sich die Population um den Faktor $2^{10} = 1024$ vermehrt, also rund vertausendfacht. – Für die Menschheit auf der Erde ist $t_2 \approx 35$ Jahre. Das dürfte ungefähr hinkommen, denn ich habe noch auf der Schule um 1950 gelernt, dass es rund 2 Milliarden Menschen gebe. $t_2 = 35$ Jahre entspricht $\alpha = 0{,}02$ pro Jahr, also $B = B_0 e^{0{,}02t}$, wobei t in Jahren angegeben werden muss. Wegen $e^{0{,}02\cdot 46} \approx 2{,}5$ stimmt das einigermaßen, denn z. Zt. gibt es rund 6 Milliarden Menschen auf der Erde. Erst im Jahre 2500 wird es etwas eng, denn dann steht jedem Menschen noch genau 1 m^2 fester Erdoberfläche zur Verfügung – sofern das Wachstum anhält.

0.2.6 Beispiel (Harmonischer Oszillator) Bei mäßiger Entfernung des Massenpunktes aus der Ruhelage ist die rücktreibende Kraft $K = -Dx$, $x = x(t)$, $t \in \mathbb{R}$, $D =$ Federkonstante. Andererseits ist nach Newton $K = m\ddot{x} =$ Masse $\times$ Beschleunigung, und wir erhalten die Differentialgleichung des harmonischen Oszillators:

$$\ddot{x} + \omega^2 x = 0, \qquad \omega^2 = \frac{D}{m},\ \omega > 0,\ t \in \mathbb{R}$$

(freie Schwingung ohne Dämpfung). Diese Differentialgleichung lässt sich wie folgt elementar lösen: Dazu setzen wir

$$y := x - x(0)\cos\omega t - \frac{\dot{x}(0)}{\omega}\sin\omega t$$

und finden: $\ddot{y} + \omega^2 y = 0. \Rightarrow \frac{d}{dt}\left(\dot{y}^2 + \omega^2 y^2\right) = 0. \Rightarrow \dot{y}^2 + \omega^2 y^2 = C$ für alle $t \in \mathbb{R}$ mit passendem C. Wegen $y(0) = \dot{y}(0) = 0$ ist $C = 0$, also $\dot{y}^2 + \omega^2 y^2 = 0. \Rightarrow y = \dot{y} = 0$, also

$$x(t) = x(0)\cos\omega t + \frac{\dot{x}(0)}{\omega}\sin\omega t \qquad (t \in \mathbb{R},\ \omega = 2\pi\nu)$$

d. h. x ist eine harmonische Schwingung. Das kann man noch etwas anders fassen: Setze $A := \left(x(0)^2 + \left(\frac{\dot{x}(0)}{\omega}\right)^2\right)^{1/2}$. Dann gibt es genau ein $\alpha \in \mathbb{R}$, so dass $\frac{x(0)}{A} = \cos\alpha$ und $\frac{\dot{x}(0)}{\omega A} = \sin\alpha$, und dann ist

$$x(t) = A(\cos\alpha\cos\omega t + \sin\alpha\sin\omega t) = A\cos(\omega t - \alpha) \qquad (t \in \mathbb{R})$$

Damit wird vollends klar, dass x eine harmonische Schwingung der Amplitude A ist mit Phasenverschiebung α. Gedämpfte Schwingungen werden wir später betrachten.

0.2.7 Beispiel (Mathematisches Pendel) Sei $\varphi = \varphi(t)$ die Auslenkung eines Pendels der Länge ℓ. Dann beträgt der zurückgelegte Weg $\ell \cdot \varphi$, die wirkende Kraft also nach Newton $m\ell\ddot{\varphi}$. Andererseits beträgt die gesamte Schwerkraft mg, die wirksame Komponente also $mg\sin\varphi$. Wir kommen also zur Bewegungsgleichung:

$$\ddot{\varphi} = \frac{g}{\ell}\sin\varphi.$$

Diese ist unabhängig von m! Diese Differentialgleichung lässt sich nicht elementar lösen. Für kleine Schwingungen ist näherungsweise $\sin\varphi \approx \varphi$, und die Schwingung ist näherungsweise harmonisch. Allgemein führt die Lösung der Differentialgleichung auf elliptische Funktionen. Das ist ein durchaus typisches Phänomen: Viele Differentialgleichungen, die keine elementaren Lösungen haben, führen auf wichtige neue Funktionen, z. B. die hypergeometrischen Funktionen, Besselsche Funktionen, Legendre-Funktionen etc.

0.2.8 Beispiel Differentialgleichung der Schar der Parabeln mit Brennpunkt 0 und y-Achse als Symmetrieachse, Brennweite p:

$$y = \frac{x^2}{2p} - \frac{p}{2} \qquad (p > 0)$$

$\Rightarrow y' = \frac{x}{p}$.

$$\Rightarrow 2yy' = x\left(\frac{x^2}{p^2} - 1\right) = x\left(y'^2 - 1\right).$$

Für variables p genügen die Parabeln alle der Differentialgleichung

$$2yy' = x\left(y'^2 - 1\right).$$

0.2.9 Beispiel Differentialgleichung der Schar von Parabeln und Geraden des Typs $y = Ax^2 + Bx + C$, $y''' = 0$ mit $A, B, C \in \mathbb{R}$.

0.2.10 Beispiel Differentialgleichung der Parabelschar

$$\begin{aligned} y(x) &= (x-\alpha)^2 \qquad (\alpha \in \mathbb{R}) \\ y' &= 2(x-\alpha), \\ y'^2 &= 4y \end{aligned}$$

Diese Differentialgleichung hat neben den obigen Parabeln z. B. noch die Lösung $y = 0$.

1 Elementare Lösungsmethoden

Inhaltsangabe

1.1 Explizite Differentialgleichungen 1. Ordnung, Beispiele elementar integrierbarer Differentialgleichungen

Geometrische Deutung expliziter Differentialgleichungen 1. Ordnung: Sei $G \subset \mathbb{R}^2$, $I \subset \mathbb{R}$ ein Intervall (beliebigen Typs), $y\colon I \to \mathbb{R}$, und für alle $x \in I$ sei $\left(x, y(x)\right) \in G$. Ferner sei $f\colon G \to \mathbb{R}$, und y genüge der expliziten Differentialgleichung

$$y' = f\left(x, y\right). \tag{1.1}$$

Ist nun $x \in I$, so ist $\left(x, y(x)\right) \in G$, und die Funktion y hat in x die Steigung $y' = f\left(x, y\right)$. Wir tragen daher in $\left(x, y\right) \in G$ einen Einheitsvektor ab mit $\sphericalangle \alpha$ gegen die x-Achse, $\tan \alpha = f\left(x, y\right)$. Das tun wir in allen Punkten $\left(x, y\right) \in G$ und erhalten ein Richtungsfeld. Ist nun $\left(x_0, y_0\right) \in G$, so betrachten wir die Aufgabe: Löse die Differentialgleichung $y' = f\left(x, y\right)$ mit der Anfangsbedingung $y(x_0) = y_0$. Geometrisch bedeutet das: Bestimme eine „Kurve“ $y(x)$, welche durch $\left(x_0, y_0\right)$ verläuft und die Eigenschaft hat, dass in jedem Punkt $\left(x, y(x)\right)$ der Vektor des Richtungsfelds gleichzeitig Tangentenvektor an die Kurve ist.

Unter hinreichenden Glattheitsvoraussetzungen an f ist es durchaus plausibel, dass man durch jeden Anfangswert $\left(x_0, y_0\right)$ genau eine solche Lösungskurve legen kann.Es wird ein wesentliches Ziel des nächsten Kapitels sein, das streng nachzuweisen.

1.1.1 Beispiel $G = I \times \mathbb{R}$, und f sei nur abhängig von x. Dann hat die Differentialgleichung (1.1) die Form

$$y' = f(x) \qquad (x \in I) \tag{1.2}$$

Das Richtungsfeld hängt dann gar nicht von y ab. Sei f stetig. Dann liefert der Hauptsatz der Differential- und Integralrechnung: Für jedes $x_0 \in I$ ist $\varphi(x) := \int_{x_0}^{x} f(t)\, dt$ eine Lösung der

Differentialgleichung $\varphi' = f$, und alle Lösungen dieser Gleichung erhält man in der Form $y(x) = \varphi(x) + C$ mit $C \in \mathbb{R}$. Ist $y_0 \in \mathbb{R}$ beliebig vorgegeben, so existiert genau eine Lösung y von (1.2) mit $y(x_0) = y_0$, und zwar

$$y(x) = y_0 + \int_{x_0}^{x} f(t)\,dt \qquad (x \in I)$$

Diese Lösung existiert in ganz I: Das Anfangswertproblem ist eindeutig lösbar.

1.1.2 Beispiel $G = \mathbb{R} \times J$ und f nur abhängig von y: Dann hat (1.1) die Form

$$y' = f(y) \qquad (x \in \mathbb{R}) \tag{1.3}$$

mit in J stetigem f. Das Richtungsfeld hängt dann gar nicht ab von x. Zur Lösung nehmen wir an, es sei y eine Lösung, $x_0 \in \mathbb{R}$ und $f\big(y(x_0)\big) \neq 0$. Da f stetig ist, gibt es eine ganze Intervall-Umgebung U von x_0 mit $f\big(y(x)\big) \neq 0$ für alle $x \in U$, und auf U gilt dann

$$\frac{y'}{f(y)} = 1. \tag{1.4}$$

Integration von x_0 nach x liefert

$$\int_{x_0}^{x} \frac{y'}{f(y)}\,dx = \int_{y_0}^{y} \frac{dy}{f(y)} = x - x_0,$$

denn auf U ist entweder stets $f\big(y(x)\big) > 0$ oder stets < 0, d. h. y' stets > 0 oder stets < 0, d. h. y streng monoton, und die Substitutionsregel kann mit y angewandt werden. Bezeichnet also F eine Stammfunktion von $\frac{1}{f}$, so ist $F(y) - F(y_0) = x - x_0$, also $F(y) = x - x_0 + F(y_0)$. Mit $G := F^{-1}|_{F(U)}$ ist dann

$$y = G\big(x - x_0 + F(y_0)\big) \qquad (x \in U) \tag{1.5}$$

Probe: Die Funktion (1.5) erfüllt die Anfangsbedingung $y(x_0) = G\big(F(y_0)\big) = y_0$, und es gilt:

$$y' = G'\big(x - x_0 + F(y_0)\big) = \frac{1}{F'(y)} = \frac{1}{\frac{1}{f(y)}} = f(y)$$

Insbesondere haben wir gezeigt: Unter den obigen Annahmen ist das Anfangswertproblem $y' = f(y)$, $y(x_0) = y_0$ (lokal) eindeutig lösbar. Dazu folgendes konkretes

1.1.3 Beispiel $y' = \sqrt{|y|}$, $G = \mathbb{R}^2$. Symmetrie: Ist y eine Lösung auf $]\alpha, \beta[$, so ist auch $z(x) := -y(-x)$ eine Lösung auf $]-\beta, -\alpha[$. Wir betrachten daher zunächst nur positive Lösungen. Sei also $I \subset \mathbb{R}$ und $y|I > 0$. Nach Beispiel 1.1.2 ist $\frac{y'}{\sqrt{y}} = 1$, d. h. $\big(2\sqrt{y}\big)' = 1$, d. h. $2\sqrt{y} = x + C$ mit $C \in \mathbb{R}$, und das gilt nur für $x > -C$, da laut Vorgabe $y > 0$. $\Rightarrow y(x) = \frac{1}{4}(x + C)^2$ für $x > -C$. Diese Lösungen lassen sich zu Lösungen auf ganz $\mathbb{R}$ fortsetzen: Setzen wir

$$y_C(x) := \begin{cases} \frac{1}{4}(x + C)^2 & \text{für } x \geqslant -C, \\ -\frac{1}{4}(x + C)^2 & \text{für } x \leqslant -C \end{cases},$$

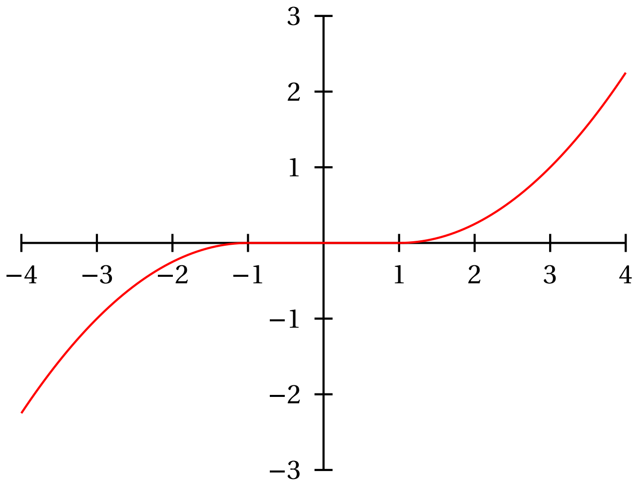

Abbildung 1.1: $y_{\alpha,\beta}$ mit $\alpha = -1$ und $\beta = 1$

so ist y_C für jedes $C \in \mathbb{R}$ eine Lösung der Differentialgleichung $y' = \sqrt{|y|}$. Insbesondere hat für alle $(x_0, y_0) \in \mathbb{R}^2$ das Anfangswertproblem $y' = \sqrt{|y|}$, $y(x_0) = y_0$ eine Lösung. Diese ist aber nicht eindeutig bestimmt: z. B. löst $y_{C=0}$ das Anfangswerteproblem $y(0) = 0$ ($x_0 = y_0 = 0$), aber auch $y = 0$ ist eine Lösung dieses Anfangswertproblems! Es gibt unendlich viele weitere Lösungen , denn für beliebige $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$ ist

$$y(x) := \begin{cases} \frac{1}{4}(x-\beta)^2 & \text{für } x \geqslant \beta, \\ 0 & \text{für } \alpha \leqslant x \leqslant \beta, \\ -\frac{1}{4}(x-\alpha)^2 & \text{für } x \leqslant \alpha \end{cases}$$

(siehe Abb. 1.1) eine Lösung, und diese erfüllt für beliebige $\alpha < 0 < \beta$ die Anfangsbedingung $y(0) = 0$. Also: Für die Differentialgleichung $y' = \sqrt{|y|}$ hat jedes Anfangswertproblem unendlich viele Lösungen.

Beispiel 1.1.1: $y' = f(x)$ und Beispiel 1.1.2: $y' = f(y)$ lassen sich verallgemeinern zu

1.1.4 Beispiel (Differentialgleichungen mit getrennten Variablen)

$$y' = f(x)\, g(y), \tag{1.6}$$

wobei $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ stetig seien und $I, J \subset \mathbb{R}$ Intervalle. Als Lösungsansatz bietet sich nach Beispiel 1.1.2 und Beispiel 1.1.3 $\frac{y'}{g(y)} = f(x)$ an ($y = y(x)$), also

$$\int_{y_0}^{y} \frac{d\eta}{g(\eta)} = \int_{x_0}^{x} f(\xi)\, d\xi, \tag{1.7}$$

d. h. $G(y) - G(y_0) = F(x) - F(x_0)$,

$$G(y) - G(y_0) = \int_{y_0}^{y} \frac{d\eta}{g(\eta)}, \qquad F(x) - F(x_0) = \int_{x_0}^{x} f(\xi)\, d\xi$$

Dann liefert Auflösung nach y:

$$y = H\left(F(x) - F(x_0) + G\left(y_0\right)\right), \qquad H = G^{-1}.$$

Dieser Ansatz klappt in der Umgebung jedes Punktes $y_0 \in J$ mit $g\left(y_0\right) \neq 0$:

1.1.5 Satz *Vorgelegt sei die Differentialgleichung* (1.6), *wobei* $f\colon I \to \mathbb{R}$, $g\colon J \to \mathbb{R}$ *stetig seien. Ferner seien* $\left(x_0, y_0\right) \in I \times J$ *und* g *nullstellenfrei in* J. *Dann gibt es eine (ggf. einseitige) Umgebung* U *von* x_0 *und eine Lösung* $y\colon U \to J$ *der Differentialgleichung* (1.6) *auf* U, *welche der Anfangsbedingung* $y(x_0) = y_0$ *genügt, und zwar ergibt sich* y *durch Auflösen von* (1.7) *nach* y.

Beweis Angenommen, es sei y eine Lösung von (1.6) in einer Umgebung U (ggf. einseitig) von x_0 mit $y(x_0) = y_0$. Für alle $\xi \in U$ gilt dann $\frac{y'(\xi)}{g(y(\xi))} = f(\xi)$, und Integration über $[x_0, x]$ (bzw. $[x, x_0]$) liefert:

$$\int_{x_0}^{x} \frac{y'(\xi)}{g\left(y(\xi)\right)}\, d\xi = \int_{x_0}^{x} f(\xi)\, d\xi.$$

Ist nun G eine Stammfunktion von $\frac{1}{g}$, so ist $\frac{d}{d\xi} G(x(\xi)) = \frac{y'(\xi)}{g(y(\xi))}$, also haben wir, wenn wir mit F eine Stammfunktion von f bezeichnen:

$$G\left(y(x)\right) - G\left(y_0\right) = F(x) - F(x_0).$$

Nun ist g stetig und nullstellenfrei, also $\frac{1}{g} > 0$ auf ganz J oder $\frac{1}{g} < 0$ auf ganz J. $\Rightarrow$ G ist stetig differenzierbar und streng monoton, hat also eine streng monotone Umkehrfunktion $H := G^{-1}$.

$$y(x) = H\left(F(x) - F(x_0) + G\left(y_0\right)\right) \qquad \text{auf } U$$

$\Rightarrow$ y ist eindeutig bestimmt. Dieses y ist wirklich eine Lösung, denn:

$$y'(x) = H'\big(\underbrace{F(x) - F(x_0) + G\left(y_0\right)}_{=G(y(x))}\big) F'(x) = \frac{1}{G'\left(y(x)\right)} f(x) = g\left(y(x)\right) f(x),$$

und es gilt

$$y(x_0) = H\left(F(x_0) - F(x_0) + G\left(y_0\right)\right) = H\left(G\left(y_0\right)\right) = y_0. \qquad \square$$

1.1.6 Beispiel $y' = e^y \sin x$. Die Voraussetzungen von Satz 1.1.5 sind erfüllt, und unser Ansatz liefert: $e^{-y} y' = \sin x$, d. h. $\left(-e^{-y}\right)' = (-\cos x)'$, also $e^{-y} = \cos x + C$ mit konstantem C, d. h. $y(x) = -\log(\cos x + C)$. Der Definitionsbereich D von y hängt stark ab von C:

$C > 1$ $\;D = \mathbb{R}$.

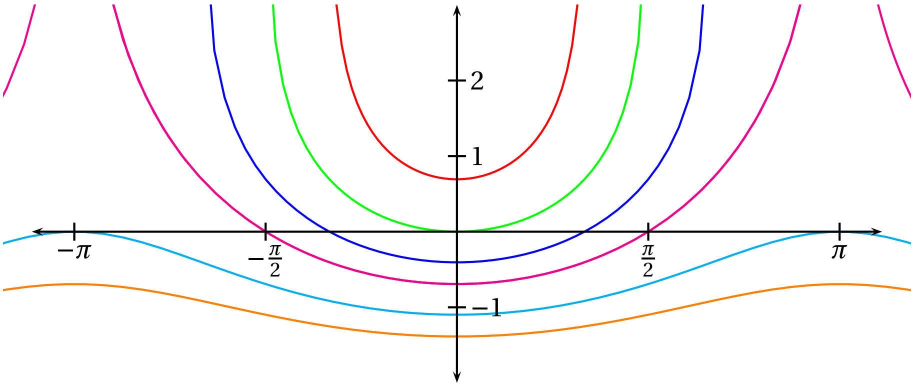

Abbildung 1.2: y_C mit $C=-\frac{1}{2}$, $C=0$, $C=\frac{1}{2}$, $C=1$, $C=2$, $C=3$

$-1 < C \leqslant 1$ Es gibt ein $\alpha \in \,]0,\pi]$ mit $C = -\cos\alpha$, und dann ist $\cos x + C = \cos x - \cos\alpha > 0$, falls $-\alpha < x - 2\pi k < \alpha$ mit $k \in \mathbb{Z}$. Dann ist also

$$D = \bigcup_{k\in\mathbb{Z}} \,]2\pi k - \alpha, 2\pi k + \alpha[\,.$$

$C \leqslant -1$ nicht sinnvoll, $D = \emptyset$.

Sind $(x_0, y_0) \in \mathbb{R}^2$ beliebig vorgegeben, so gilt wegen $\log(]0,\infty[) = \mathbb{R}$:

$$y(x_0) = y_0 \Leftrightarrow y_0 = -\log(x_0 + C) \Leftrightarrow e^{-y_0} = \cos x_0 + C \Leftrightarrow C = e^{-y_0} - \cos x_0,$$

und dieses C ist notwendig > -1, d. h. das zugehörige y ist sinnvoll, da $x_0 \in D$.
Ergebnis: Für alle $(x_0, y_0) \in \mathbb{R}^2$ ist das Anfangswertproblem

$$y' = e^y \sin x, \qquad y(x_0) = y_0$$

lokal eindeutig lösbar. Für $X = e^{-y_0} + \cos x_0 > 1$ ist die Lösung auf ganz $\mathbb{R}$ definiert, für $-1 < C \leqslant 1$ hat die Lösung den obigen Definitionsbereich D.

Obgleich die Differentialgleichung keine Probleme mit dem Definitionsbereich bereitet, sind die Lösungen nicht notwendig auf ganz $\mathbb{R}$ definiert. Selbst eine „kleine“ Änderung der Anfangswerte kann ganz erhebliche Änderungen bei der Lösung bewirken (siehe Abb. 1.2).
Oft lassen sich Differentialgleichungen durch einfache Transformationen auf bekannte Differentialgleichungen zurückführen:

1.1.7 Beispiel

$$y' = f(ax + by + c) \tag{1.8}$$

mit $a, b, c \in \mathbb{R}$. Sei gleich $b \neq 0$ (sonst liegt Beispiel 1.1.1 vor). Ansatz: $u(x) = ax + by(x) + c$.

$$\Rightarrow u' = a + by' = a + bf(u) \tag{1.9}$$

also eine Reduktion auf Beispiel 1.1.2. Umgekehrt: Jede Lösung u von (1.9) liefert eine Lösung y von (1.8).

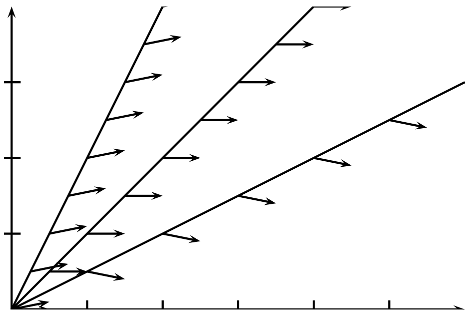

Abbildung 1.3: Richtungsfeld einer homogenen Differentialgleichung

1.1.8 Beispiel $y' = \left(x+y\right)^2$. $u := x+y \Rightarrow u' = 1+y' = 1+u^2$. $\Rightarrow (\arctan u)' = 1$, d. h. $\arctan u = x+C$ (nur möglich, solange $x+C \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$), also $u = \tan(x+C)$ und somit $y = \tan(x+C) - x$. Zur Probe rechnet man nach:

$$y' = \tan'(x+C) - 1 = 1 + \tan^2(x+C) - 1 = \left(x+y\right)^2$$

1.1.9 Beispiel (Homogene Differentialgleichungen) Die Funktion $f\left(x,y\right)$, f stetig, heißt *homogen* , falls für alle $\lambda \in \mathbb{R}$ gilt: $f\left(\lambda x, \lambda y\right) = f\left(x,y\right)$. Dann ist mit $\lambda := \frac{1}{x}$ für $x \neq 0$:

$$f\left(x,y\right) = f\left(1, \frac{y}{x}\right) =: g\left(\frac{y}{x}\right),$$

d. h. die Differentialgleichung hat die Form:

$$y' = f\left(\frac{y}{x}\right) \qquad \left(x \neq 0,\ g \text{ stetig}\right) \tag{1.10}$$

Ansatz: $u = \frac{y(x)}{x}$.

$$\Rightarrow y' = (ux)' = u + xu' = g(u),$$

d. h. die Differentialgleichung nimmt die Form an:

$$u' = \frac{g(u) - x}{x}, \tag{1.11}$$

so dass eine Reduktion auf eine Differentialgleichung mit getrennten Variablen vorliegt. Umgekehrt liefert jede Lösung von (1.11) vermöge $y = ux$ eine Lösung von (1.10).
Richtungsfeld einer homogenen Differentialgleichung: Auf allen Geraden durch 0 ist $\frac{y}{x} = c$ konstant, und auf dieser Geraden ist die Richtung festgelegt durch $\tan\alpha = g(c)$. Das sieht so aus wie in Abb. 1.3.

1.1.10 Beispiel Die Differentialgleichung

$$y' = \frac{y}{x} - \frac{x^2}{y^2} = g\left(\frac{y}{x}\right) \tag{1.12}$$

ist homogen. Setze $u = \frac{y}{x}$. Aus (1.11) folgt dann: $u' = \frac{g(u)-u}{x} = -\frac{1}{xu^2}$, d. h. $u^2u' = -\frac{1}{x}$, also $\left(\frac{1}{3}u^3\right)' = -\left(\log|x|\right)'$ für $x \neq 0$, d. h. $u^3 = -\log|x|^3 + C$ für $x \neq 0$, also $y^3 = x^3\left(-\log|x|^3 + C\right)$ für $x \neq 0$.
Für beliebiges $C \in \mathbb{R}$ und $x \neq 0$ ist das wirklich eine Lösung von (1.12), denn

$$3y^2y' = 3x^2\left(-\log|x|^3 + C\right) + x^3\left(-\frac{3}{x}\right) = 3\frac{y^3}{x} - 3x^2,$$

d. h. $y' = \frac{y}{x} - \frac{x^2}{y^2}$ (beachte: $y \neq 0$!).
Anfangsbedingungen: Der Anfangswert $x_0 = 0$ ist offenbar nicht sinnvoll. Sei $x_0 \neq 0$, $y_0 \in \mathbb{R}$: y erfüllt die Anfangsbedingung $y(x_0) = y_0$ genau dann, wenn $y_0^3 = x_0^3\left(-\log|x_0|^3 + C\right)$, d. h. wenn $C = \frac{y_0^3}{x_0^3} + \log|x_0|^3$. Das Anfangswertproblem (1.12) mit $y(x_0) = y_0$ ist also auf $]0,\infty[$ (im Fall $x_0 > 0$) bzw. auf $]-\infty,0[$ (im Fall $x_0 < 0$) eindeutig lösbar.

1.1.11 Beispiel $y' = f\left(\frac{ax+by+c}{\alpha x+\beta y+\gamma}\right)$ mit $a,b,c,\alpha,\beta,\gamma \in \mathbb{R}$. Zur Reduktion der Differentialgleichung unterscheiden wir 2 Fälle:

1. Fall: $\det\begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} \neq 0$. Wir wollen neue Variablen einführen gemäß $x = u + h$, $y = v + k$ mit konstanten $h, k \in \mathbb{R}$, die wie folgt gewählt werden:

$$\begin{aligned} ax + by + c &= au + bv + ah + bk + c, \\ \alpha x + \beta y + \gamma &= \alpha u + \beta v + \alpha h + \beta k + \gamma, \end{aligned}$$

und wegen $\det\begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} \neq 0$ können h, k so gewählt werden, dass $ah + bk + c = 0$ und $\alpha h + \beta k + \gamma = 0$. Dann hat die Differentialgleichung die Form $v' = f\left(\frac{au+bv}{\alpha u+\beta v}\right)$, und das ist eine homogene Differentialgleichung.

2. Fall: $\det\begin{pmatrix} a & b \\ \alpha & \beta \end{pmatrix} = 0$. Dann sind die beiden Zeilen dieser Matrix linear abhängig.

(i) $a = b = \alpha = \beta = 0$: Dann ist die Differentialgleichung trivial.

(ii) $\begin{pmatrix} a & b \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \end{pmatrix}$. Dann gibt es ein $\lambda \in \mathbb{R}$, so dass $\begin{pmatrix} \alpha & \beta \end{pmatrix} = \lambda\begin{pmatrix} a & b \end{pmatrix}$. Die Differentialgleichung hat also die Form $y' = f\left(\frac{ax+by+c}{\lambda(ax+by)+\gamma}\right)$.

(1) $b = 0$: klar

(2) $b \neq 0$: Setze $z := ax + by$, dann ist $z' = a + bf\left(\frac{z+c}{\lambda z+\gamma}\right)$, also haben wir eine Reduktion auf den Fall (1.3) aus Beispiel 1.1.2

(iii) $\begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$, aber $\begin{pmatrix} \alpha & \beta \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \end{pmatrix}$. Damit die Differentialgleichung interessant ist, setzen wir $c \neq 0$ voraus und setzen $g(t) := f\left(\frac{1}{t}\right)$ für $t \neq 0$, und haben die Differentialgleichung $y' = g\left(\frac{\alpha x+\beta y+\gamma}{c}\right)$, also eine Reduktion auf den Fall (2ii) mit $\lambda = 0$.

In allen Fällen ist die Differentialgleichung damit auf bekannte Fälle reduziert.

1.2 Lineare Differentialgleichungen erster Ordnung

1.2.1 Definition (Lineare Differentialgleichung erster Ordnung) Eine Differentialgleichung heißt linear von erster Ordnung, wenn sie die Form

$$f y' + g y = h \tag{1.13}$$

mit stetigen $f, g, h \colon I \to \mathbb{R}$ hat. Ist f nullstellenfrei, so kann gleich $f = 1$ angenommen werden, und dann hat (1.13) die *Normalform*

$$y' + g y = h \tag{1.14}$$

mit stetigen $g, h \colon I \to \mathbb{R}$. Ist in (1.14) $h = 0$, so heißt (1.14) homogen, sonst inhomogen.

1.2.2 Satz *Vorgelegt sei die lineare Differentialgleichung erster Ordnung* (1.14) *mit stetigen* $g, h \colon I \to \mathbb{R}$*. Dann existiert zu jedem* $(x_0, y_0) \in I \times \mathbb{R}$ *genau eine Lösung* $y \colon I \to \mathbb{R}$ *von* (1.14) *mit* $y(x_0) = y_0$*, und diese Lösung ist gegeben durch*

$$y(x) = \left(y_0 + \int_{x_0}^{x} h(t) \exp\left(\int_{x_0}^{t} g(u)\, du\right) dt\right) \exp\left(-\int_{x_0}^{x} g(t)\, dt\right) \qquad (x \in I) \tag{1.15}$$

Beweis (i) Lösung der homogenen Differentialgleichung $y' + g y = 0$: Sei y Lösung der Differentialgleichung $y' = -g y$. Man dividiert an dieser Stelle nicht durch y, da $y = 0$ eine Lösung ist, y also a priori nicht notwendig nullstellenfrei ist.

$$\Rightarrow \frac{d}{dx}\left(y(x) \exp\left(\int_{x_0}^{x} g(t)\, dt\right)\right) = \left(y'(x) + g(x)\, y(x)\right) \exp\left(\int_{x_0}^{x} g(t)\, dt\right) = 0$$

Die Funktion ist folglich konstant, also ist wegen der Anfangsbedingung $y(x_0) = y_0$:

$$y(x) = y_0 \exp\left(-\int_{x_0}^{x} g(t)\, dt\right) \qquad (x \in I)$$

Dieses ist wirklich eine Lösung der homogenen Differentialgleichung! Das ist (1.15) im Falle $h = 0$, also ist die Behauptung bewiesen für $h = 0$.

(ii) Allgemeiner Fall: Wir setzen („Variation der Konstanten“)

$$z(x) := y(x) \exp\left(\int_{x_0}^{x} g(u)\, du\right) \qquad (x \in I)$$

Dann ist die Anfangsbedingung $y(x_0) = y_0 \Leftrightarrow z(x_0) = y_0$.

$$\Rightarrow z' = \underbrace{(y' + g y)}_{= h \text{ wg } (1.14)} \exp\left(\int_{x_0}^{x} g(u)\, du\right) = h(x) \exp\left(\int_{x_0}^{x} g(u)\, du\right) \tag{1.16}$$

$$\Rightarrow z(x) = \int_{x_0}^{x} h(z) \exp\left(\int_{x_0}^{t} g(u)\, du\right) dt + y_0 \tag{1.17}$$

löst die Differentialgleichung (1.16) und erfüllt die Anfangsbedingung $z(x_0) = y_0$.

Ergebnis: Wenn eine Lösung y von (1.14) mit $y(x_0) = y_0$ existiert, so ist

$$z(x) = y(x)\exp\left(\int_{x_0}^{x} g(u)\,du\right)$$

Lösung von (1.16) mit $z(x_0) = y_0$, und z ist durch (1.17) gegeben, d. h. y ist durch (1.15) gegeben.

Umgekehrt: (1.15) ist wirklich eine Lösung von (1.14) mit $y(x_0) = y_0$; Probe: $y(x_0) = y_0$ ist klar.

$$\begin{aligned} &y'(x) + g(x)\,y(x) \\ &= \left(h(x)\exp\left(\int_{x_0}^{x} g(u)\,du\right)\right)\exp\left(-\int_{x_0}^{x} g(t)\,dt\right) + y(x)\left(-g(x)\right) + g(x)\,y(x) \\ &= h(x)\,. \end{aligned}$$

□

1.2.3 Bemerkung a) Sind y_1, y_2 zwei Lösungen der inhomogenen Differentialgleichung (1.14), so ist $y_1 - y_2$ eine Lösung der homogenen Differentialgleichung

$$y' + gy = 0. \tag{1.18}$$

b) Ist y_1 irgendeine „feste" Lösung von (1.14), so erhält man alle Lösungen y von (1.14) in der Form $y = y_1 + z$, wobei z eine Lösung von (1.18) ist.

c) Der Lösungsraum der homogenen Differentialgleichung (1.18) hat die Dimension 1 und wird von $\varphi(x) := \exp\left(-\int_{x_0}^{x} g(t)\,dt\right)$ aufgespannt. Insbesondere ist jede Lösung $y \neq 0$ von (1.18) nullstellenfrei auf I.

1.2.4 Beispiel (Fall mit (Luft)reibung im homogenen Schwerefeld) Zur Veranschaulichung der Situation siehe Abb. 1.4. $K = (0, -mg)$ sei die Schwerkraft, g die Fallbeschleunigung, $R = -\kappa v$ mit $\kappa > 0$ sei die Reibungskraft, $v = (\dot{x}, \dot{y})$ die Geschwindigkeit und $b = (\ddot{x}, \ddot{y})$ die Beschleunigung. Die Bewegungsgleichung $mb = K - \kappa v$ besagt dann: $m\ddot{x} = -\kappa\dot{x}$ und $m\ddot{y} = -\kappa\dot{y} - mg$, d. h. mit $\lambda := \frac{\kappa}{m} > 0$:

$$\ddot{x} + \lambda\dot{x} = 0, \qquad \ddot{y} + \lambda\dot{y} = -g$$

Mit der Anfangsgeschwindigkeit $\dot{x}(t_0) =: u_0$, $\dot{y}(t_0) =: v_0$ erhalten wir dann aus (1.15):

$$\begin{aligned} \dot{x}(t) &= u_0 e^{-\lambda(t-t_0)}, \\ \dot{y}(t) &= \left(v_0 - g\int_{t_0}^{t} e^{\lambda(\xi - t_0)}\,d\xi\right)e^{-\lambda(t-t_0)} \\ &= v_0 e^{-\lambda(t-t_0)} - \frac{g}{\lambda}\left(1 - e^{-\lambda(t-t_0)}\right), \end{aligned}$$

die Geschwindigkeit in x-Richtung nimmt also exponentiell ab.

$$\begin{aligned}
\Rightarrow x(t) &= \underbrace{x(t_0)}_{=:x_0} + \frac{u_0}{\lambda}\left(1 - e^{-\lambda(t-t_0)}\right) \\
y(t) &= y(t_0) + \frac{v_0}{\lambda}\left(1 - e^{-\lambda(t-t_0)}\right) - \frac{g}{\lambda}(t-t_0) + \frac{g}{\lambda^2}\left(1 - e^{-\lambda(t-t_0)}\right) \\
&= y(t_0) + \left(\frac{v_0}{\lambda} + \frac{g}{\lambda^2}\right)\left(1 - e^{-\lambda(t-t_0)}\right) - \frac{g}{\lambda}(t-t_0).
\end{aligned}$$

Das sind die Lösungen der Bewegungsgleichung.

$$\begin{aligned}
\Rightarrow \lim_{t\to\infty} x(t) &= x(t_0) + \frac{u_0}{\lambda} = x(t_0) + \frac{m u_0}{\kappa} \\
\lim_{t\to\infty} \dot{x}(t) &= 0, \\
\lim_{t\to\infty} \dot{y}(t) &= \frac{g}{\lambda}
\end{aligned}$$

d. h. die x Koordinate ist beschränkt, die Geschwindigkeit in x-Richtung nimmt exponentiell ab, und die Geschwindigkeit in y-Richtung ist für große t nahezu konstant. Das kann man z. B. schön bei Sinkversuchen in Flüssigkeiten beobachten.
Für $u_0 \neq 0$ erhalten wir:

$$\begin{aligned}
y(t) &= y_0 + \left(\frac{v_0}{\lambda} + \frac{g}{\lambda^2}\right)\underbrace{\left(1 - e^{-\lambda(t-t_0)}\right)}_{=(x-x_0)\frac{\lambda}{u_0}} - \frac{g}{\lambda}(t-t_0) \\
&= y_0 + \left(v_0 + \frac{g}{\lambda}\right)\frac{1}{u_0}(x(t)-x_0) + \frac{g}{\lambda^2}\log\left(1 - \frac{\lambda}{u_0}(x(t)-x_0)\right) \\
&= y_0 + \left(v_0 + \frac{g}{\lambda}\right)\frac{1}{u_0}(x(t)-x_0) - \frac{g}{\lambda^2}\left(\frac{\lambda}{u_0}(x(t)-x_0) + \frac{\lambda^2}{2u_0^2}(x(t)-x_0)^2 + \frac{\lambda^3}{3u_0^3}(x(t)-x_0)^3 + \dots\right) \\
&\xrightarrow{\kappa\to 0} y_0 + \frac{v_0}{u_0}(x(t)-x_0) - \frac{g}{2u_0^2}(x(t)-x_0)^2,
\end{aligned}$$

falls $\left|\frac{\lambda}{u_0}(x(t)-x_0)\right| < 1$, nach Definition von λ. Dabei hebt sich der erste Term der Reihe gegen den zweiten Summanden aus der Klammer weg.
Das stellt die bekannte Wurfparabel dar, d. h. im Falle $\kappa = 0$ (d. h. Fall ohne Reibung) fällt der Massenpunkt auf einer Parabel.

1.2.5 Beispiel (Wechselstromkreis) („quasi-stationär“, siehe Abb. 1.5) U sei die Spannung (bekannt), I die Stromstärke (gesucht), R ein Ohmscher Widerstand (konstant), L eine Selbstinduktion (konstant). $\Rightarrow U$ = gesamter Spannungsabfall = Spannungsabfall am Ohmschen Widerstand + induzierte Spannung $= RI + L\frac{dI}{dt}$. Als Anfangsbedingung haben wir $I(0) = I_0$. Das führt uns zu der Differentialgleichung

$$\frac{dI}{dt} + \frac{R}{L}I = \frac{U}{L}$$

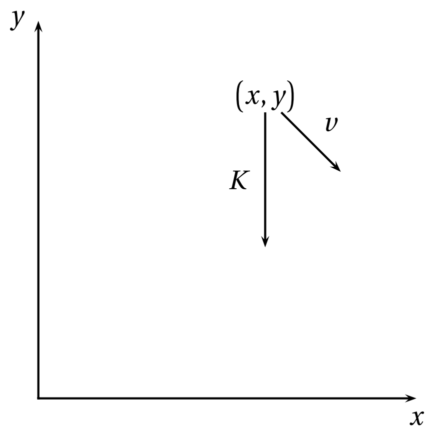

Abbildung 1.4: Fall im homogenen Schwerefeld

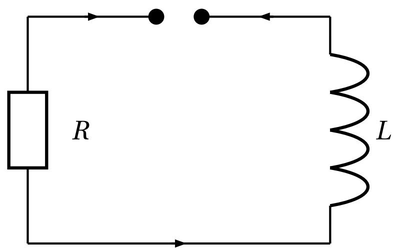

Abbildung 1.5: Stromkreis aus Beispiel 1.2.5

Mit $g := \frac{R}{L}$ und $h := \frac{U}{L}$ entspricht das der Form von (1.14), wir können also Satz 1.2.2 anwenden, und erhalten

$$\begin{aligned}
I(t) &= \left(I_0 + \int_0^t \frac{U(s)}{L} \exp\left(\int_0^s \frac{R}{L}\, du\right) ds\right) \exp\left(-\int_0^t \frac{R}{L}\, du\right) \\
&= \left(I_0 + \int_0^t \frac{U(s)}{L} e^{\frac{R}{L}s}\, ds\right) e^{-\frac{R}{L}t} \qquad (t \in \mathbb{R})
\end{aligned}$$

Betrachten wir nun den Spezialfall des Sinus-förmigen Wechselstroms: $U = U_0 \sin\omega t$ mit $\omega = 2\pi\nu > 0$.

$$\Rightarrow I(t) = \left(I_0 + \frac{U_0}{L} \int_0^t \sin\omega s \cdot e^{\frac{R}{L}s}\, ds\right) e^{-\frac{R}{L}t}$$

Wegen

$$\begin{aligned}
&\left(e^{\frac{R}{L}t}\left(\sin\omega t - \frac{\omega L}{R}\cos\omega t\right)\right)' \frac{RL}{R^2+\omega^2 L^2} \\
&= e^{\frac{R}{L}t}\left(\frac{R}{L}\sin\omega t - \omega\cos\omega t + \omega\cos\omega t + \frac{\omega^2 L}{R}\sin\omega t\right)\frac{RL}{R^2+\omega^2L^2} \\
&= e^{\frac{R}{L}t}\sin\omega t
\end{aligned}$$

folgt

$$\begin{aligned}
I(t) &= \left(I_0 + \frac{U_0}{L}\left[\frac{e^{\frac{R}{L}s}RL}{R^2+\omega^2L^2}\left(\sin\omega s - \frac{\omega L}{R}\cos\omega s\right)\right]_0^t\right) e^{-\frac{R}{L}t} \\
&= \left(I_0 + \frac{U_0\omega L}{R^2+\omega^2L^2}\right) e^{-\frac{R}{L}t} + \frac{U_0}{R^2+\omega^2L^2}\left(R\sin\omega t - \omega L\cos\omega t\right).
\end{aligned}$$

Es gibt ein (modulo 2π eindeutiges) $\varphi \in \mathbb{R}$ mit

$$\frac{R}{\sqrt{R^2+\omega^2L^2}} = \cos\varphi, \qquad \frac{\omega L}{\sqrt{R^2+\omega^2L^2}} = \sin\varphi$$

und mit dieser *Phasenverschiebung* φ erhalten wir[1]

$$I(t) = \left(I_0 + \frac{U_0\omega L}{R^2+\omega^2L^2}\right) e^{-\frac{R}{L}t} + \frac{U_0}{\sqrt{R^2+\omega^2L^2}}\sin\left(\omega t - \varphi\right)$$

Ergebnis: Bis auf den vom Einschaltvorgang herrührenden exponentiell abfallenden Anteil ist der Strom sinusförmig mit derselben Frequenz wie die Spannung, aber in der Phase verschoben. Die Amplitude des Stroms ist aber nicht gleich $\frac{U_0}{R}$, wie man vom Gleichstrom her erwarten möchte, sondern $\frac{U_0}{\sqrt{R^2+\omega^2L^2}}$.

[1] $\sin\left(\omega t - \varphi\right) = \sin\left(-\varphi\right)\cos\omega t + \sin\omega t\cos\left(-\varphi\right) = \sin\omega t\cos\varphi - \sin\varphi\cos\omega t$

1.2.6 Beispiel (Bernoullische Differentialgleichung) (benannt nach Johann Bernoulli, 1667–1748).

$$y' + g(x)\,y + h(x)\,y^\alpha = 0, \tag{1.19}$$

$g, h\colon I \to \mathbb{R}$ stetig. Für $\alpha \in \mathbb{R}\setminus\mathbb{Z}$ ist hier $y > 0$ vorauszusetzen, für $\alpha \in \mathbb{N}$ ist keine besondere Voraussetzung bzgl. y nötig, für $\alpha \in \mathbb{Z}$, $\alpha < 0$ ist y als nullstellenfrei vorauszusetzen.
Triviale Fälle:

$\alpha = 1$: Homogene lineare Differentialgleichung in Normalform (1.18): Lösung ist bekannt: (1.15).

$\alpha = 0$: Inhomogene lineare Differentialgleichung in Normalform (1.14): Lösung ist bekannt: (1.15)

Lösungsansatz: Sei gleich $\alpha \notin \{0, 1\}$, $x_0 \in I$ und $U \subset I$ eine (ggf. einseitige) Umgebung von x_0, $y\colon U \to \mathbb{R}$ sei eine Lösung von (1.19), insbesondere sei y^α sinnvoll (d. h. y genüge obigen Anforderungen); zusätzlich sei $y^{-\alpha}$ sinnvoll, d. h. $y|_U$ sei auch im Falle $\alpha \in \mathbb{N}$ nullstellenfrei. Multiplikation von (1.19) mit $(1-\alpha)\,y^{-\alpha}$ liefert dann:

$$(1-\alpha)\,y^{-\alpha}y' + g(x)\,(1-\alpha)\,y^{1-\alpha} + (1-\alpha)\,h(x) = 0,$$

d. h.

$$\left(y^{1-\alpha}\right)' + (1-\alpha)\,g(x)\,y^{1-\alpha} + (1-\alpha)\,h(x) = 0.$$

$\Rightarrow z := y^{1-\alpha}$ ist sinnvoll und genügt der linearen Differentialgleichung erster Ordnung

$$z' + (1-\alpha)\,g(x)\,z + (1-\alpha)\,h(x) = 0. \tag{1.20}$$

Umgekehrt: Ist $z\colon U \to \mathbb{R}$ eine Lösung von (1.20) und $y := z^{\frac{1}{1-\alpha}}$ sinnvoll, so ist

$$y' = \frac{1}{1-\alpha}\,\frac{y}{z}\,z' = -g(x)\,y - h(x)\,\frac{y}{z} = -g(x)\,y - h(x)\,y^\alpha,$$

also gilt (1.19).
Diskussion des Anfangswertproblems: Gegeben sei die Differentialgleichung (1.19) mit $\alpha \in \mathbb{R}$, $\alpha \neq 1$; gesucht ist eine Lösung $y > 0$ mit $y(x_0) = y_0$, wobei $x_0 \in I$, $y_0 > 0$ vorgegeben seien.
Lösung: Nach Satz 1.2.2 existiert eine Lösung z von (1.20) mit $z(x_0) = y_0^{1-\alpha} > 0$, und es gibt eine Intervallumgebung $U \subset I$ von x_0 mit $z|_U > 0$. Obige Überlegung liefert jetzt $y := z^{\frac{1}{1-\alpha}}$ ist auf U definiert und positiv, y löst (1.19) und es gilt $y(x_0) = y_0$. Das Anfangswertproblem ist also lokal eindeutig lösbar. Für spezielle α sind weitere Fälle zu diskutieren:

$\alpha = 0$: klar nach (1.15), jedes Anfangswertproblem ist eindeutig auf ganz I lösbar.

$\alpha \in \mathbb{Z}$, $\alpha \neq 1$: Auch Lösungen $y < 0$ sind zu diskutieren: (1.19) liefert:

$$(-y)' + g(x)\,(-y) + (-1)^{\alpha+1}\,h(x)\,(-y)^\alpha = 0. \tag{1.21}$$

D. h. mit y ist auch $-y$ Lösung einer Bernoullischen Differentialgleichung mit $(-1)^{\alpha+1}\,h$ an Stelle von h.

(i) α ungerade: Mit y ist auch $-y$ Lösung von (1.19). D. h.: Die negativen Lösungen von (1.19) erhält man aus den positiven durch Übergang $y \mapsto -y$.

(ii) α gerade: Ist $y < 0$ Lösung von (1.19), so ist $v := -y > 0$ eine Lösung von (1.21), und (1.21) ist vom Bernoulli-Typ, und die positiven Lösungen des Anfangswertproblems sind nach obigem bekannt.

$\alpha > 0$: Wegen $0^\alpha = 0$ ist auch $y = 0$ eine (triviale) Lösung von (1.19).

1.2.7 Beispiel

$$y' + \frac{1}{1+x} y + (1+x)\, y^4 = 0 \qquad (x \neq -1) \tag{1.22}$$

Die Differentialgleichung ist auch für $y < 0$ sinnvoll. Sei y eine nullstellenfreie Lösung im Intervall $I \subset \mathbb{R} \setminus \{-1\}$. $\Rightarrow z = y^{-3}$ ist Lösung der Differentialgleichung

$$z' - \frac{3}{1+x} z - 3\,(1+x) = 0 \qquad (x \neq -1)\,. \tag{1.23}$$

Mit (1.15) folgt dann für $-1 \notin [x_0, x]$ bzw. $[x, x_0]$ mit beliebigem $z_0 \in \mathbb{R}$:

$$\Rightarrow z\,(x) = \left(z_0 + \int_{x_0}^{x} 3\,(1+t)\, e^{-\int_{x_0}^{t} \frac{3}{1+u}\,du}\, dt \right) \exp\left(\int_{x_0}^{x} \frac{3}{1+t}\, dt \right)$$

Wir rechnen zunächst für $-1 < x_0 < x$ und erhalten:

$$\begin{aligned} z\,(x) &= \left(z_0 + \int_{x_0}^{x} 3\,(1+t)\, e^{-3\left(\log(1+t) - \log(1+x_0)\right)}\, dt \right) e^{3\left(\log(1+x) - \log(1+x_0)\right)} \\ &= \left(z_0 + \int_{x_0}^{x} 3\,(1+t)\, \frac{(1+x_0)^3}{(1+t)^3}\, dt \right) \frac{(1+x)^3}{(1+x_0)^3} \\ &= z_0 \frac{(1+x)^3}{(1+x_0)^3} + 3\,(1+x)^3 \int_{x_0}^{x} \frac{dt}{(1+t)^2} \\ &= z_0 \frac{(1+x)^3}{(1+x_0)^3} + 3\,(1+x)^3 \left[-(1+t)^{-1} \right]_{x_0}^{x}, \end{aligned}$$

also

$$z\,(x) = z_0 \frac{(1+x)^3}{(1+x_0)^3} + 3 \frac{(1+x)^3}{1+x_0} - 3\,(1+x)^2\,. \tag{1.24}$$

Das ist die eindeutig bestimmte Lösung des Anfangswertproblems $z' - \frac{3}{1+x} z - 3\,(1+x) = 0$, $z\,(x_0) = z_0$, wobei die Lösung bei beliebigem $z_0 \in \mathbb{R}$, $x_0 \neq 1$ existiert auf ganz $]-\infty, -1[$ im Fall $x_0 < -1$ bzw. $]-1, \infty[$ im Fall $x_0 > -1$. Diese Lösung wurde unter der Voraussetzung $-1 < x_0 < x$ hergeleitet, aber nun sehen wir: Für beliebiges $x_0 \neq -1$ ist (1.24) eine Lösung von (1.23) mit $z\,(x_0) = z_0$ für beliebiges $z_0 \in \mathbb{R}$ (triviale Probe). Die Lösung ist sogar auf ganz $\mathbb{R}$ sinnvoll, (1.23) nur für $x \neq -1$. Zur Lösung (1.24) von (1.23) gehört die Lösung $y = z^{-\frac{1}{3}}$:

$$y\,(x) = \frac{1}{\sqrt[3]{z_0 \left(\frac{1+x}{1+x_0} \right)^3 + 3 \frac{(1+x)^3}{1+x_0} - 3\,(1+x)^2}}, \tag{1.25}$$

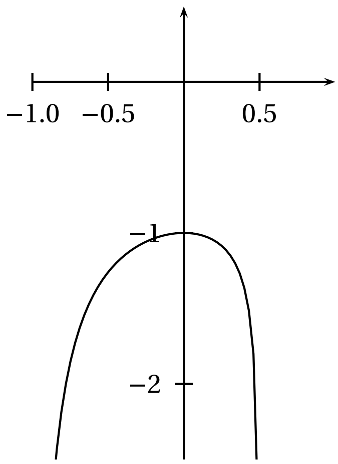

Abbildung 1.6: $y(x) = -\frac{1}{\sqrt[3]{(1+x)^2(1-2x)}}$

und diese ist sinnvoll, solange $x_= \neq -1$, $x \neq -1$ und der Nenner nicht verschwindet. $y(x_0) = z_= 0^{-\frac{1}{3}} =: y_0$, d. h. wähle $z_0 := y_0^3$, und das Anfangswertproblem der Differentialgleichung (1.19) mit dem Anfangswert $y(x_0) = y_0$ ist für $y_0 \neq 0$ lokal eindeutig lösbar. Für $y_0 = 0$ hat (1.22) die triviale Lösung $y = 0$; diese kann man als Grenzfall von (1.25) für $z_0 \to \infty$ auffassen. – Weiter gilt mit $x_0 \neq -1 \neq x$:

$$
\begin{aligned}
&\sqrt[3]{z_0\left(\frac{1+x}{1+x_0}\right)^3 + 3\frac{(1+x)^3}{1+x_0} - 3(1+x)^2} \neq 0 \\
&\quad \Leftrightarrow \underbrace{\left(\frac{z_0}{(1+x_0)^3} + 3\frac{1}{1+x_0}\right)}_{=:C}(1+x) \neq 3 \\
&\quad \Leftrightarrow C = 0 \text{ oder } C \neq 0 \wedge x \neq \frac{3}{C} - 1.
\end{aligned}
$$

Beispiel: $x_0 = 0$, $y_0 = -1$ (d. h. $z_0 = -1$). Dann haben wir die Lösung des Anfangswertproblems

$$
y(x) = \frac{1}{\sqrt[3]{2(1+x)^3 - 3(1+x)^2}} = \frac{1}{\sqrt[3]{(1+x)^2(2x-1)}} = -\frac{1}{\sqrt[3]{(1+x)^2(1-2x)}},
$$

und dies ist die Lösung des Anfangswertproblems mit maximalem Definitionsintervall $I = \left]-1, \frac{1}{2}\right[$ (siehe Abb. 1.6).

1.2.8 Beispiel (Riccatische Differentialgleichung) benannt nach Graf Jacopo Francesco Riccati, 1676–1754, der wichtige Arbeiten zur Theorie der Differentialgleichungen schrieb.

$$
y' + g(x)y + h(x)y^2 = k(x) \tag{1.26}
$$

mit stetigen $g, h, k\colon I \to \mathbb{R}$. Die Lösungen der Riccatischen Differentialgleichung lassen sich – von Spezialfällen mit z. B. $h = 0$ abgesehen – nicht elementar angeben. Die Differentialgleichung hat aber folgende merkwürdige Eigenschaft: Kennt man eine Lösung der Riccatischen Differentialgleichung, so sind die übrigen elementar berechenbar.

Beweis Sei φ eine „bekannte“ Lösung von (1.26), y irgendeine Lösung von (1.26), $u := y - \varphi$. Dann gilt mit (1.26)

$$u' + g(x)\,u + h(x)\underbrace{\left(y^2 - \varphi^2\right)}_{=u(u+2\varphi)} = 0,$$

und das schreibt sich als

$$u' + \left(g(x) + 2\varphi(x)\,h(x)\right)u + h(x)\,u^2 = 0. \tag{1.27}$$

Das ist eine Bernoullische Differentialgleichung vom Typ (1.19) mit $\alpha = 2$. Diese wird durch $z := u^{-1}$ (Annahme: u sei nullstellenfrei im betrachteten Intervall) überführt in die lineare Differentialgleichung

$$z' - \left(g(x) + 2\varphi(x)\,h(x)\right)z - h(x) = 0, \tag{1.28}$$

und diese kann man mit (1.15) lösen. – Wir haben schon gesehen: Jede nullstellenfreie Lösung z von (1.28) gibt die Lösung $u := z^{-1}$ von (1.27), und dann ist $y := u + \varphi$ eine Lösung von (1.26), da

$$\begin{aligned} y' &= u' + \varphi' = -\left(g + 2\varphi h\right)u - hu^2 - g\varphi - h\varphi^2 + k \\ &= -g\left(u + \varphi\right) - h\left(u + \varphi\right)^2 + k. \end{aligned}$$

□

1.2.9 Beispiel

$$y' - 2xy - y^2 = 2 \tag{1.29}$$

Diese Riccati-Differentialgleichung ($g(x) = -2x$, $h = -1$, $k = 2$) hat die spezielle Lösung $\varphi(x) = -\frac{1}{x}$ für $x \neq 0$. Wir schreiben (1.28) um und erhalten:

$$\begin{aligned} & z' + \left(2x - \frac{2}{x}\right)z + 1 = 0. \\ \overset{(1.15)}{\Rightarrow}\; z(x) &= \left(z_0 + \int_{x_0}^{x} (-1)\exp\left(\int_{x_0}^{t}\left(2u - \frac{2}{u}\right)du\right)dt\right)\exp\left(-\int_{x_0}^{x}\left(2t - \frac{2}{t}\right)dt\right) \\ &= \left(z_0 - \int_{x_0}^{x}\exp\left(t^2 - x_0^2 - \log\frac{t^2}{x_0^2}\right)dt\right)\exp\left(-\left(x^2 - x_0^2\right) + \log\frac{x^2}{x_0^2}\right) \\ &= \underbrace{z_0\frac{x^2}{x_0^2}e^{-\left(x^2 - x_0^2\right)}}_{\substack{\text{Lösung der hom. Gl.} \\ z' + \left(2x - \frac{2}{x}\right)z = 0}} \underbrace{-x^2 e^{-x^2}\int_{x_0}^{x}\frac{e^{t^2}}{t^2}\,dt}_{\substack{\text{Lösung der inhom. Gl.} \\ z' + \left(2x - \frac{2}{x}\right)z + 1 = 0}} \end{aligned}$$

Hier ist noch

$$
\begin{aligned}
x^2 e^{-x^2} \int_{x_0}^{x} \frac{e^{t^2}}{t^2}\, dt &= x^2 e^{-x^2} \int_{x_0}^{x} \left(-\frac{1}{t}\right)' e^{t^2}\, dt \\
&= x^2 e^{-x^2} \left[-\frac{e^{t^2}}{t}\right]_{x_0}^{x} + x^2 e^{-x^2} \int_{x_0}^{x} \frac{1}{t} 2t e^{t^2}\, dt \\
&= -x + \frac{x^2}{x_0} e^{-(x^2-x_0^2)} + 2x^2 e^{-x^2} \int_{x_0}^{x} e^{t^2}\, dt,
\end{aligned}
$$

und wir erhalten

$$
\begin{aligned}
z(x) &= \left(\frac{z_0}{x_0^2} - \frac{1}{x_0}\right) x^2 e^{-(x^2-x_0^2)} + x - 2x^2 e^{-x^2} \int_{x_0}^{x} e^{t^2}\, dt \\
&= C x^2 e^{-x^2} + x - 2x^2 e^{-x^2} \int_{x_0}^{x} e^{t^2}\, dt
\end{aligned}
$$

mit konstantem C.

$$
\begin{aligned}
\Rightarrow y &= u + \varphi \\
&= -\frac{1}{x} + z^{-1} \\
&= -\frac{1}{x} + \frac{1}{x + x^2 e^{-x^2} \left(C - 2\int_{x_0}^{x} e^{t^2}\, dt\right)},
\end{aligned}
$$

also ist

$$
y(x) = \frac{-e^{-x^2} \left(C - 2\int_{x_0}^{x} e^{t^2}\, dt\right)}{1 + x e^{-x^2} \left(C - 2\int_{x_0}^{x} e^{t^2}\, dt\right)} \tag{1.30}
$$

eine Lösung von (1.29). Diese Lösung ist überall dort definiert, wo der Nenner nicht verschwindet, speziell auch für $x = 0$! Besonders einfach wird die Sache für $x_0 = 0$: Für jedes $y_0 \in \mathbb{R}$ ist das Anfangswertproblem der Differentialgleichung (1.29) mit $y(0) = y_0$ lokal eindeutig lösbar, denn mit $C := -y_0$ leistet (1.30) das Gewünschte. Für beliebiges x_0, das keine Nullstelle des Nenners ist, gilt:

$$
\begin{aligned}
y(x_0) = y_0 &\Leftrightarrow \frac{-e^{-x_0^2} C}{1 + x_0 e^{-x_0^2} C} = y_0 \\
&\Leftrightarrow \frac{-C}{e^{x_0^2} + x_0 C} = y_0 \\
&\Leftrightarrow -C\left(1 + x_0 y_0\right) = y_0 e^{x_0^2} \\
&\Leftrightarrow C = -\frac{y_0 e^{x_0^2}}{1 + x_0 y_0},
\end{aligned}
$$

und das Anfangswertproblem ist lokal eindeutig lösbar genau dann, wenn $x_0 y_0 \neq -1$, denn dann ist mit obigem C automatisch

$$e^{x_0^2} + x_0 C = e^{x_0^2}\left(1 - \frac{x_0 y_0}{1 + x_0 y_0}\right) = \frac{e^{x_0^2}}{1 + x_0 y_0} \neq 0,$$

und man kann rechnen wie oben angeführt.

1.3 Exakte Differentialgleichungen

1.3.1 Beispiel (Differentialgleichung einer Kurvenschar) Gegeben seien ein Gebiet $G \subset \mathbb{R}^2$ und $f : G \to \mathbb{R}$ stetig. Dann bilden die Lösungen der Differentialgleichung $y' = f\left(x, y\right)$ für variable Anfangswerte $\left(x_0, y_0\right) \in G$ eine *Kurvenschar*, die G überdeckt: siehe Satz 2.1.7
Umgekehrt: Sei eine Kurvenschar gegeben bestehend aus differenzierbaren Funktionen, die G einfach überdeckt. Dann ist diese Kurvenschar die Menge der Lösungen einer geeigneten Differentialgleichung: Sei z. B. $\left(x_0, y_0\right) \in G$ und φ die Kurve zu $\left(x_0, y_0\right)$: $\varphi(x_0) = y_0$. Für $\left(x, y\right) \in G$ mit $\varphi(x) = y$ (d. h. für Punkte auf dem Graphen von φ) setzen wir $f\left(x, y\right) := \varphi'(x)$ (Beachte: φ ist differenzierbar nach Voraussetzung!). Dadurch wird eine (stetige) Funktion $f : G \to \mathbb{R}$ definiert, und nach Konstruktion ist jede Funktion der Kurvenschar Lösung der Differentialgleichung $y' = f\left(x, y\right)$. Dabei wird man als Kurven φ nicht nur Graphen von Funktionen der Variablen x betrachten, sondern gleich allgemeinere Kurven $\left(x(t), y(t)\right)$, wobei $t \in I$ ein geeigneter Parameter ist.

1.3.2 Beispiel $G = \mathbb{R}^2 \setminus \{0\}$, $x^2 + y^2 = r^2$ $(r > 0)$ bildet eine Schar konzentrischer Kreise. Die zugehörige Differentialgleichung lautet: $yy' + x = 0$. Wenn wir uns auf Funktionen der Variablen x beschränken wollen, haben wir z. B. folgende Möglichkeiten: $G^+ := \left\{\left(x, y\right) : y > 0\right\}$, Lösungen hier $y(x) = \sqrt{r^2 - x^2}$ für $-r < x < r$ $(r > 0$ Parameter) oder $G = \left\{\left(x, y\right) : y < 0\right\}$, Lösungen $y(x) = -\sqrt{r^2 - x^2}$ für $-r < x < r$ $(r > 0$ Parameter). Wenn wir im ursprünglichen G bleiben wollen, können wir nicht mehr Funktionen von x als Lösungen erwarten. Aber führen wir den Winkel als Parameter ein, $\left(x(t), y(t)\right) = r(\cos t, \sin t)$ $(0 \leqslant t \leqslant 2\pi)$, so bilden diese Kurven Lösungen der Differentialgleichung $yy' + xx' = 0$. Da es auf die Wahl des Kurvenparameters nicht ankommt, schreibt man das gern in der Form $y\,dy + x\,dx = 0$. Das kann man als Abkürzung für die vorangehende Differentialgleichung ansehen (das dt lässt man hier einfach weg). Auch kommt die Symmetrie in $\left(x, y\right)$ besser zum Ausdruck: In den Punkten der x-Achse wird man lokale Lösungen nicht in der Form $y(x)$ suchen, sondern in der Form $x\left(y\right)$ und y als Variable ansehen.
Allgemein: Hat man eine Kurvenschar der Form $F\left(x, y\right) = C$, $(\left(x, y\right) \in G)$, so genügt diese der Differentialgleichung

$$\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} = 0 \qquad \text{oder auch} \qquad \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0;$$

wenn wir Lösungen als Funktionen von x ausschreiben, haben wir die Differentialgleichung

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{dy}{dx} = 0.$$

Solche Differentialgleichungen nennt man *exakt*:

1.3.3 Definition (Exakte Differentialgleichung) Seien $G \subset \mathbb{R}^2$ ein Gebiet und $g, h\colon G \to \mathbb{R}$ stetig. Die Differentialgleichung

$$g(x,y)\,dx + h(x,y)\,dy = 0 \tag{1.31}$$

auf G heißt exakt, wenn es eine stetig differenzierbare Funktion $F\colon G \to \mathbb{R}$ gibt mit

$$F_x = g, \qquad F_y = h; \tag{1.32}$$

F heißt dann eine *Stammfunktion* der Differentialgleichung (1.31).

1.3.4 Folgerung Hat (1.31) eine Stammfunktion F im Gebiet G, so ist F bis auf eine Konstante eindeutig bestimmt.

Beweis Ist E eine zweite Stammfunktion, so hat $u := F - E$ die Eigenschaft: u ist stetig differenzierbar mit $u_x = u_y = 0$. Dann verschwinden wegen $u_\xi = \langle \operatorname{grad} u, \xi \rangle$ für beliebiges $\xi \in \mathbb{R}^2$ alle Richtungsableitungen von u, so dass u konstant ist. Siehe dazu auch Forster (1999). □

1.3.5 Satz *Ist die Differentialgleichung* (1.31) *exakt mit Stammfunktion F auf G, so erhält man durch (lokale) Auflösung der Gleichung $F(x,y) = C$ der sog.* Niveaukurve *sämtliche Lösungen $\binom{x(t)}{y(t)}$ von* (1.31)*. Zu jedem $(x_0, y_0) \in G$ mit $F_y(x_0,y_0) = h(x_0,y_0) \neq 0$ setze man $C := F(x_0,y_0)$. Dann gibt es offene Intervall-Umgebungen U von x_0 und V von y_0 mit $U \times V \subset G$ und eine stetig differenzierbare Lösung $y\colon U \to V$ der Differentialgleichung*

$$g(x,y) + h(x,y)\,y' = 0 \tag{1.33}$$

mit Anfangswert $y(x_0) = y_0$, so dass $F(x, y(x)) = C$ für alle $x \in U$ und

$$\{(x,y) \in G\colon F(x,y) = C\} \cap (U \times V) = \operatorname{Graph} y.$$

Bemerkung Ist $F_x(x_0,y_0) = g(x_0,y_0) \neq 0$, so existiert entsprechend eine Lösung $x(y)$ der Differentialgleichung

$$g(x,y)\frac{dx}{dy} + h(x,y) = 0 \tag{1.34}$$

mit $x(y_0) = x_0$, die in einer Umgebung von x_0 erklärt ist und die man durch Auflösung von $F(x(y), y) = F(x_0, y_0)$ erhält.

Beweis Sei $J \subset \mathbb{R}$ ein Intervall und $\binom{x}{y}\colon J \to G$ differenzierbar. Dann gilt laut Kettenregel:

$$g\dot{x} + h\dot{y} = F_x\dot{x} + F_y\dot{y} = \frac{d}{dt}F(x(t), y(t));$$

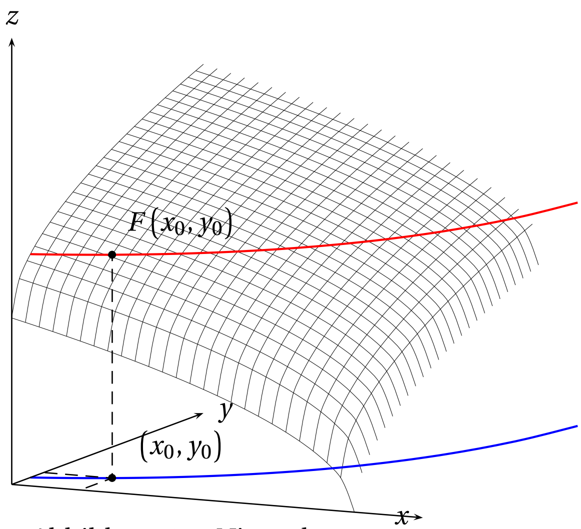

Abbildung 1.7: Niveaukurve

also ist $\binom{x}{y}$ Lösung von $g\dot{x} + h\dot{y} = 0$ genau dann, wenn $F\left(x(t), y(t)\right)$ konstant ist auf J. Sei weiter $\left(x_0, y_0\right) \in G$ und $F_y\left(x_0, y_0\right) = h\left(x_0, y_0\right) \neq 0$. Im Falle $F_x\left(x_0, y_0\right) = g\left(x_0, y_0\right) \neq 0$ schließt man ebenso mit $x\left(y\right)$ statt $y(x)$. Wegen $\frac{\partial F}{\partial y}\left(x_0, y_0\right) \neq 0$, existiert nach dem Satz über implizite Funktionen (siehe Elstrodt (2002a)) eine Umgebung U von x_0 und eine Umgebung V von y_0 und eine Funktion $y\colon U \to V$ mit $F\left(x, y(x)\right) = C$ auf U, und $\left\{\left(x, y\right) \in G\colon F\left(x, y\right) = C\right\} \cap (U \times V) = \operatorname{Graph} y$. Nach der Kettenregel gilt für $x \in U$:

$$0 = \frac{d}{dx} F\left(x, y(x)\right) = F_x\left(x, y(x)\right) + f_y\left(x, y(x)\right) y'(x) = g\left(x, y(x)\right) + h\left(x, y(x)\right) y'(x),$$

d. h. y ist eine Lösung von (1.33). □

1.3.6 Beispiele a) Die Differentialgleichung $2x\,dx + 2y\,dy = 0$ ist exakt, denn $F\left(x, y\right) = x^2 + y^2$ ist Stammfunktion. Die Niveaukurven $F\left(x, y\right) = R^2$ $(R > 0)$ sind Kreise. Diese können wir global parametrisieren in der Form $\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = R\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ für $t \in [0, 2\pi]$, und das ist eine Lösung der Differentialgleichung. Aber wir können die Kreise nicht global als Graphen von Funktionen in x darstellen: Wohl aber können wir die Kreise „lokal" als Funktionen in x darstellen in einer Umgebung des Anfangswertes $\left(x_0, y_0\right)$, wenn $F_y\left(x_0, y_0\right) = h\left(x_0, y_0\right) = 2y_0 \neq 0$ (siehe Abb. 1.8): Für $y_0 \neq 0$ ist

$$y(x) = \left(\operatorname{sgn} y_0\right) \sqrt{\left(x_0^2 + y_0^2\right) - x^2}$$

Lösung in $U = U_r(x_0)$, und dabei darf durchaus $y_0 > 0$ oder auch $y_0 < 0$ sein. Für $y_0 = 0$ schreibe man die Lösung lokal als $x = x\left(y\right)$! Das geht, da $F_x\left(x_0, y_0\right) 2x_0 \neq 0$, falls $y_0 = 0$.

b) Die Differentialgleichung

$$\left(y^2 e^{xy} + 3x^2 y\right) dx + \left(x^3 + \left(1 + xy\right) e^{xy}\right) dy = 0$$

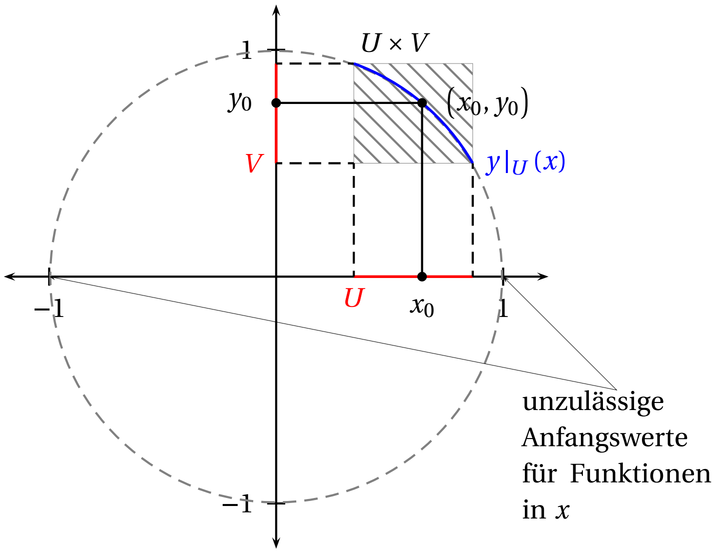

Abbildung 1.8: Parametrisierung von Kreisen

ist exakt mit der Stammfunktion $F(x,y) = y\left(e^{xy} + x^3\right)$. Die Auflösung von $F(x,y) = C$ nach y oder x ist hier nicht ohne weiteres in expliziter Form möglich. Aber numerisch ist eine solche Auflösung mit Computerhilfe in der Regel in befriedigender Weise möglich.

Im Prinzip kann man nach Satz 1.3.5 die Lösungen exakter Differentialgleichungen sofort angeben. Wie prüft man, ob eine Differentialgleichung exakt ist?

1.3.7 Satz (Notwendige Bedingung für Exaktheit) *Gegeben sei die Differentialgleichung*

$$g(x,y)\,dx + h(x,y)\,dy = 0 \qquad (x \in G) \tag{1.35}$$

mit dem Gebiet G und stetig differenzierbaren Funktionen $g, h\colon G \to \mathbb{R}$. *Ist* (1.35) *exakt, so gilt*

$$g_y = h_x \qquad (x \in G)\,. \tag{1.36}$$

Beweis Ist (1.35) exakt, so existiert eine stetig differenzierbare Funktion $F\colon G \to \mathbb{R}$ mit $F_x = g$ und $F_y = h$. g, h sind stetig differenzierbar, also ist F zweimal stetig partiell differenzierbar, und es gilt:

$$g_y = (F_x)_y = \left(F_y\right)_x = h_x. \qquad \square$$

Bemerkung Die Differentialform erster Ordnung ω auf G heißt *geschlossen*, falls $d\omega = 0$. Im Falle $\omega = g\,dx + h\,dy$ ist (vgl. Forster (1977)):

$$\begin{aligned} d\omega &= dg \wedge dx + dh \wedge dy \\ &= \left(g_x\,dx + g_y\,dy\right) \wedge dx + \left(h_x\,dx + h_y\,dy\right) \wedge d_y \\ &= \left(h_x - g_y\right)\,dx \wedge dy, \end{aligned}$$

dabei gilt $dx \wedge dx = dy \wedge dy = 0$, $dx \wedge dy = -dy \wedge dx$, und das ist eine Basis der Differentialformen zweiter Ordnung. $d\omega$ heißt die äußere Ableitung oder das Differential von ω. Siehe dazu auch Forster (1996), Königsberger (1903).
Wir haben also: ω geschlossen $\Leftrightarrow \exists \omega = dF$ mit $dF = F_x\, dx + F_y\, dy$. Satz 1.3.7 besagt in dieser Redeweise: ω exakt $\Rightarrow \omega$ geschlossen. Hintergrund dieser Implikation ist die allgemeine Rechenregel $d(dF) = 0$. Die Umkehrung ist hier im Allgemeinen falsch, gilt aber unter einer wichtigen Zusatzvoraussetzung: Ist G einfach zusammenhängend, so ist jede geschlossene Form auf G exakt.

1.3.8 Beispiel Die Differentialform

$$\omega = \underbrace{\left(3y + e^x\right)}_{=:g} dx + \underbrace{\left(3x + \cos y\right)}_{=:h} dy$$

ist geschlossen, denn $\frac{\partial}{\partial y}\left(3y + e^x\right) = 3$, $\frac{\partial}{\partial x}\left(3x + \cos y\right) = 3$. Ist ω exakt? Wenn ja, ist $\frac{\partial F}{\partial x} = g$, $\frac{\partial F}{\partial y} = h$, d. h. $\frac{\partial F}{\partial x} = 3y + e^x$, also $F\left(x, y\right) = 3xy + e^x + \varphi\left(y\right)$ mit einer differenzierbaren Funktion $\varphi\colon \mathbb{R} \to \mathbb{R}$. Die Bedingung $\frac{\partial F}{\partial y} = h\left(x, y\right) = 3x + \cos y \overset{!}{=} 3x + \varphi'\left(y\right)$ führt uns zu $\varphi\left(y\right) := \sin y$, so dass wir mit

$$F\left(x, y\right) = 3xy + e^x + \sin y$$

eine Stammfunktion von ω haben, und wir können die Differentialgleichung nach Satz 1.3.5 lösen.

1.3.9 Beispiel Die Differentialgleichung

$$\underbrace{12xy + 3}_{=:g} + \underbrace{6x^2}_{=:h} \frac{dy}{dx} = 0$$

erfüllt die notwendige Exaktheitsanforderung: $\frac{\partial g}{\partial y} = 12x = \frac{\partial h}{\partial x}$. Konstruktion einer Stammfunktion: Ist F Stammfunktion, so gilt $\frac{\partial F}{\partial x} = g = 12xy + 3$, also $F\left(x, y\right) = 6x^2 y + 3x + \varphi\left(y\right)$. Die Bedingung $\frac{\partial F}{\partial y} = 6x^2 + \varphi'\left(y\right) \overset{!}{=} h\left(x, y\right) = 6x^2$ liefert $\varphi' = 0$. Also ist $F\left(x, y\right) = 6x^2 y + 3x$ Stammfunktion. Sei $x_0 \neq 0$, $y_0 \in \mathbb{R}$. Dann ist die Lösung des Anfangswertproblems $y(x_0) = y_0$ lokal gegeben durch

$$F\left(x, y\right) = 6x^2 y + 3x = F\left(x_0, y_0\right) = 6x_0^2 y_0 + 3x_0,$$

d. h.

$$y = \frac{6x_0^2 y_0 + 3x_0 - 3x}{6x^2} = \frac{2x_0^2 y_0 + x_0 - x}{2x^2} \qquad (x \neq 0).$$

In der Tat ist dies eine Lösung des Anfangswertproblems, solange $x \neq 0$, und diese Lösung ist für die x mit $\operatorname{sgn} x = \operatorname{sgn} x_0$ eindeutig bestimmt.

1.3.10 Beispiel Das Gebiet $G = \mathbb{R}^2 \setminus \{0\}$ ist nicht einfach zusammenhängend. Die Differentialform

$$\omega = \underbrace{-\frac{y}{x^2+y^2}}_{=:g}\, dx + \underbrace{\frac{x}{x^2+y^2}}_{=:h}\, dy$$

ist geschlossen, denn

$$\frac{\partial}{\partial y}\left(-\frac{y}{x^2+y^2}\right) = -\frac{x^2-y^2}{\left(x^2+y^2\right)^2} = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right).$$

Aber ω ist nicht exakt! Für $x \neq 0$ ist $F\left(x,y\right) := \arctan\frac{y}{x}$ eine Stammfunktion, denn

$$\frac{\partial F}{\partial x} = \frac{1}{1+\left(\frac{y}{x}\right)^2}\left(-\frac{y}{x^2}\right) = g\left(x,y\right),$$
$$\frac{\partial F}{\partial y} = \frac{1}{1+\left(\frac{y}{x}\right)^2}\cdot\frac{1}{x} = h\left(x,y\right).$$

Aber ω ist nicht exakt! Angenommen, es gibt doch eine Stammfunktion F auf G. Dann kann wegen Folgerung 1.3.4 gleich angenommen werden, dass $F\left(x,y\right) = \arctan\frac{y}{x}$ für $x > 0$, und nach Folgerung 1.3.4 existiert ein $C \in \mathbb{R}$, so dass für $x < 0$ gilt: $F\left(x,y\right) = \arctan\frac{y}{cx} + C$. Zur Bestimmung von C lassen wir $\left(x,y\right) \to (0,1)$ gehen, und zwar einmal aus Richtung positiver x, und einmal aus Richtung negativer x. Wegen

$$\lim_{\substack{(x,y)\to(0,1)\\ x>0}} F\left(x,y\right) = \frac{\pi}{2},$$
$$\lim_{\substack{(x,y)\to(0,1)\\ x<0}} F\left(x,y\right) = -\frac{\pi}{2} + C$$

ist notwendig $C = \pi$. Mit diesem Wert für C ist in der Tat

$$F\left(x,y\right) = \begin{cases} \arctan\frac{y}{x} & \text{für } x > 0 \\ \frac{\pi}{2} & \text{für } x = 0,\ y > 0 \\ \arctan\frac{y}{x} + \pi & \text{für } x < 0 \end{cases} \tag{1.37}$$

eine Stammfunktion von ω auf dem kleineren Gebiet $\mathbb{R}^2 \setminus \left\{\binom{0}{y} : y < 0\right\}$. Aber es gibt keine auf $\mathbb{R}^2 \setminus \{0\}$ stetige Funktion F, die mit den Werten aus (1.37) verträglich ist! Es gilt nämlich bei $(0,-1)$:

$$\lim_{\substack{(x,y)\to(0,-1)\\ x>0}} F\left(x,y\right) = -\frac{\pi}{2},$$
$$\lim_{\substack{(x,y)\to(0,-1)\\ x<0}} F\left(x,y\right) = \frac{\pi}{2} + \pi = \frac{3\pi}{2},$$

und dasselbe gilt in allen Punkten $\left(0,y\right)$ mit $y < 0$. Also ist ω nicht exakt.

Erfüllt die Differentialgleichung

$$g(x,y) + h(x,y)\frac{dy}{dx} = 0$$

die notwendige Exaktheitsbedingung $\frac{\partial g}{\partial y} = \frac{\partial h}{\partial x}$ auf G, so braucht keine Stammfunktion zu existieren. Die nähere Untersuchung hat gezeigt, dass es wesentlich vom zugrunde liegenden Gebiet abhängt, ob eine Stammfunktion existiert oder nicht, und zwar fällt die Antwort stets dann positiv aus, wenn G einfach zusammenhängend ist. Diese Bedingung bedeutet anschaulich, dass G ein Gebiet „ohne Löcher" ist.
Einfach zusammenhängende Gebiete spielen in der Funktionentheorie eine wichtige Rolle und werden dort auf vielerlei ganz verschiedene Arten charakterisiert; z. B. gilt für beschränkte G: G ist einfach zusammenhängend genau dann, wenn $\mathbb{R}^2 \setminus G$ zusammenhängend ist. Siehe dazu auch Elstrodt (2003), Satz 22.5.
Wir wollen eine für unsere Zwecke besonders handliche Definition des einfachen Zusammenhangs mit Hilfe von Kurven in G vornehmen.

1.3.11 Definition a) Eine Kurve in G (bzw. ein Weg in G) ist eine stetige Abbildung

$$\gamma\colon [a,b] \to G, \quad t \mapsto \gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

wobei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall ist. $\gamma(a)$ heißt der Anfangspunkt von γ, $\gamma(b)$ der Endpunkt von γ, $[\gamma]$ oder $\gamma([a,b])$ die Spur von γ. γ heißt geschlossen genau dann, wenn $\gamma(a) = \gamma(b)$.

b) Seien $\gamma_0, \gamma_1\colon [0,1] \to G$ zwei geschlossene Kurven in G. Dann heißt γ_0 homotop zu γ_1, wenn es eine stetige Funktion

$$\varphi\colon [0,1] \times [0,1] \to G$$

gibt, so dass gilt:

(i) $\varphi(s,0) = \gamma_0(s)$ für $0 \leqslant s \leqslant 1$,
(ii) $\varphi(s,1) = \gamma_1(s)$ für $0 \leqslant s \leqslant 1$,
(iii) $\varphi(0,t) = \varphi(1,t)$ für $0 \leqslant t \leqslant 1$, d. h. $\gamma_t := \varphi(\cdot,t)\colon [0,1] \to G$ ist eine geschlossene stetige Kurve für alle $t \in [0,1]$.

Schreibweise: $\gamma_0 \sim \gamma_1$ (dies ist eine Äquivalenz-Relation!).
Anschauliche Vorstellung: $\gamma_0 = \varphi(\cdot,0)$ wird im Laufe der Zeit $t \in [0,1]$ über die geschlossenen Kurven γ_t stetig in γ_1 deformiert, ohne dass eine der geschlossenen Kurven γ_t „zerrissen" wird.

1.3.12 Beispiel $G = \mathbb{R}^2 \setminus \{0\}$, $0 < r < R$:

$$\gamma_0(s) := re^{2\pi i s} = \begin{pmatrix} r\cos 2\pi s \\ r\sin 2\pi s \end{pmatrix} \qquad (0 \leqslant s \leqslant 1),$$

$$\gamma_1(s) := Re^{2\pi i s} = \begin{pmatrix} R\cos 2\pi s \\ R\sin 2\pi s \end{pmatrix} \qquad (0 \leqslant s \leqslant 1).$$

$\Rightarrow \gamma_0 \sim \gamma_1$, denn

$$\varphi\colon [0,1]\times[0,1] \to G, \quad \varphi(s,t) := (r + t(R-r))\, e^{2\pi i s} \qquad (0 \leqslant s, t \leqslant 1)$$

leistet das Verlangte.

1.3.13 Definition a) Sei $a \in G$. Dann heißt $\gamma_1\colon [0,1] \to G$, $\gamma_1(s) := a$ für $0 \leqslant s \leqslant 1$ ein konstanter Weg.

b) Sei $\gamma\colon [0,1] \to G$ ein stetiger geschlossener Weg. Dann heißt γ *nullhomotop* in G, wenn γ zu einem konstanten Weg in G homotop ist. Man sagt in diesem Fall auch, γ sei stetig auf einen Punkt zusammenziehbar.

1.3.14 Beispiel a) $G = \mathbb{R}^2$, dann ist $\gamma(s) := e^{2\pi i s}$ für $0 \leqslant s \leqslant 1$ nullhomotop in G, denn $\varphi(s,t) := (1-t)\, e^{2\pi i s}$ leistet das Verlangte.

b) $G = \mathbb{R}^2 \setminus \{0\}$, dann ist $\gamma(s) := e^{2\pi i s}$ für $0 \leqslant s \leqslant 1$ nicht nullhomotop.

1.3.15 Definition (Einfacher Zusammenhang) G heißt einfach zusammenhängend, wenn jede geschlossene Kurve in G nullhomotop in G ist.

1.3.16 Definition (Sterngebiet) G heißt sternförmig oder *Sterngebiet*, wenn es ein $a \in G$ gibt, so dass für alle $b \in G$ gilt: $\{a + t(b-a) : 0 \leqslant t \leqslant 1\} \subset G$.

1.3.17 Beispiel a) $\{kx\colon k \in \mathbb{R}\} \subset \mathbb{R}^2$ ist ein Sterngebiet.

b) Ein „Stern“ mit Mittelpunkt a im $\mathbb{R}^2$ ist ein Sterngebiet.

c) $\mathbb{R}^2 \setminus \{(x,0) : x \leqslant 0\}$ ist ein Sterngebiet.

1.3.18 Satz *Jedes Sterngebiet ist einfach zusammenhängend.*

Beweis Sei $\gamma_0\colon [0,1] \to G$ geschlossen und a wie in Definition 1.3.16, $\gamma_1(s) = a$ für $0 \leqslant s \leqslant 1$. Dann leistet

$$\varphi(s,t) := a + (1-t)\big(\gamma_0(s) - a\big) \in G \qquad (0 \leqslant s, t \leqslant 1)$$

das Verlangte:

$$\begin{aligned}
&\varphi(s,0) = \gamma_0(s) \qquad (0 \leqslant s \leqslant 1),\\
&\varphi(s,1) = \gamma_1(s) \qquad (0 \leqslant s \leqslant 1),\\
&\varphi(0,t) = a + (1-t)\big(\gamma_0(0) - a\big) = a + (1-t)\big(\gamma_0(1) - a\big) = \varphi(1,t) \qquad (0 \leqslant t \leqslant 1).
\end{aligned}$$

□

1.3.19 Definition (Stückweise stetige Differenzierbarkeit) Die Kurve $\gamma\colon \big[\alpha,\beta\big] \to \mathbb{R}^2$ heißt *stückweise stetig differenzierbar*, wenn es eine Unterteilung $\alpha = \alpha_0 < \ldots < \alpha_n = \beta$ gibt, so dass $\gamma|_{[\alpha_{j-1},\alpha_j]}$ für $1 \leqslant j \leqslant n$ stetig differenzierbar ist (mit einseitiger Differenzierbarkeit in den Endpunkten).

1.3.20 Beispiel Jeder Streckenzu ist bei passender Parametrisierung stückweise stetig differenzierbar.

1.3.21 Definition (Kurvenintegral) Seien $g, h\colon G \to \mathbb{R}$ stetig und $\gamma := \binom{x}{y}\colon [\alpha, \beta] \to G$ stückweise stetig differenzierbar. Dann wird definiert:

$$\begin{aligned}\int_\gamma (g\,dx + h\,dy) &:= \int_\alpha^\beta \left(g\big(x(t), y(t)\big)\frac{dx(t)}{dt} + h\big(x(t), y(t)\big)\frac{dy(t)}{dt}\right) dt \\ &:= \sum_{\nu=1}^{n} \int_{\alpha_{\nu-1}}^{\alpha_\nu} \left(g\big(x(t), y(t)\big)\frac{dx(t)}{dt} + h\big(x(t), y(t)\big)\frac{dy(t)}{dt}\right) dt\end{aligned}$$

1.3.22 Beispiel Sei $\gamma(t) := a + t(b-a) \in G$, $a, b \in G$, $0 \leqslant t \leqslant 1$, wobei $a = \binom{a_1}{a_2}$, $b = \binom{b_1}{b_2}$. Dann ist

$$\int_\gamma (g\,dx + h\,dy) = \int_0^1 \Big(g\big(\gamma(t)\big)(b_1 - a_1) + h\big(\gamma(t)\big)(b_2 - a_2)\Big)\, dt.$$

1.3.23 Satz (Rechenregeln) *a) Bezeichnet $\gamma^-(t) := \gamma\big(\alpha + \beta - t\big)$ für $\alpha \leqslant t \leqslant \beta$ die „rückwärts durchlaufene" Kurve γ, so ist*

$$\int_{\gamma^-} (g\,dx + h\,dy) = -\int_\gamma (g\,dx + h\,dy).$$

b) Ist $\varphi\colon [\alpha', \beta'] \to [\alpha, \beta]$ bijektiv und stückweise stetig differenzierbar mit $\varphi'(t) > 0$ für alle $t \in [\alpha', \beta']$ (ggf. einseitige Ableitungen an den Unstetigkeitsstellen von φ'), so heißt $\gamma_\varphi := \gamma \circ \varphi\colon [\alpha', \beta'] \to G$ eine Umparametrisierung *von γ. Es gilt die Substitutionsregel:*

$$\int_{\gamma_\varphi} (g\,dx + h\,dy) = \int_\gamma (g\,dx + h\,dy).$$

c) Existiert eine stetig differenzierbare Stammfunktion F, so dass $F_x = g$ und $F_y = h$, so gilt:

$$\int_\gamma (g\,dx + h\,dy) = F\big(\gamma(\beta)\big) - F\big(\gamma(\alpha)\big)$$

mit den Bezeichnungen wie in Definition 1.3.21.

Beweis hier nur zu c):

$$\begin{aligned}\int_\gamma (g\,dx + h\,dy) &= \sum_{\nu=1}^{n} \int_{\alpha_{\nu-1}}^{\alpha_\nu} \underbrace{\left(F_x\big(x(t), y(t)\big)\frac{dx(t)}{dt} + F_y\big(x(t), y(t)\big)\frac{dy(t)}{dt}\right)}_{=\frac{dF(\gamma(t))}{dt}} \\ &= \sum_{\nu=1}^{n} \Big(F\big(\gamma(\alpha_\nu)\big) - F\big(\gamma(\alpha_{\nu-1})\big)\Big) \\ &= F\big(\gamma(\beta)\big) - F\big(\gamma(\alpha)\big).\end{aligned}$$

□

Im folgenden bereiten wir den Hauptsatz über Stammfunktionen (Satz 1.3.27) vor. Der Beweis beruht maßgeblich auf

1.3.24 Lemma (Poincaré) benannt nach Henri Poincaré, 1854–1912, französischer Mathematiker, schrieb über 500 Arbeiten und 30 Bücher, Vetter von Raymond Poincaré, dem 9. Präsidenten der französischen Republik (1913–1920).
Es seien G ein Sterngebiet und die Differentialgleichung

$$g\,dx + h\,dy = 0, \tag{1.38}$$

$g, h\colon G \to \mathbb{R}$ stetig differenzierbar, erfülle die notwendige Exaktheitsbedingung

$$g_y = h_x. \tag{1.39}$$

Dann existiert eine Stammfunktion $F\colon G \to \mathbb{R}$ von (1.38) mit $g = F_x$ und $h = F_y$, und zwar folgende: Sei $a \in G$ so beschaffen, dass

$$\sigma(t) := a + t\left(\begin{pmatrix} x \\ y \end{pmatrix} - a\right) \in G \qquad \left(\begin{pmatrix} x \\ y \end{pmatrix} \in G,\ t \in [0,1]\right). \tag{1.40}$$

Dann ist

$$F(x,y) := \int_\sigma \left(g(\xi,\eta)\,d\xi + h(\xi,\eta)\,d\eta\right) \qquad \left(\begin{pmatrix} x \\ y \end{pmatrix} \in G\right)$$

eine Stammfunktion von (1.38) auf G.

Beweis Nach Definition ist mit $a = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$:

$$F(x,y) = \int_0^1 \left(g\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right)(x-\alpha) + h\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right)(y-\beta)\right) dt.$$

Nach der in Elstrodt (2002a) entwickelten Theorie ist die Differentiation unter dem Integral zulässig und liefert unter Benutzung der Kettenregel:

$$\begin{aligned}
\frac{\partial}{\partial x} F(x,y) &= \int_0^1 \Big(g_x\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right) t(x-\alpha) + g\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right) \\
&\qquad + h_x\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right) t(y-\beta)\Big)\, dt \\
&\overset{(1.39)}{=} \int_0^1 \Big(\big(g_x\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right)(x-\alpha) \\
&\qquad + g_y\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right)(y-\beta)\big)\, t \\
&\qquad + g\left(\alpha + t(x-\alpha), \beta + t(y-\beta)\right)\Big)\, dt \\
&= \int_0^1 \left(t\frac{d}{dt} g\left(a + t\left(\begin{pmatrix} x \\ y \end{pmatrix} - a\right)\right) + g\left(a + t\left(\begin{pmatrix} x \\ y \end{pmatrix} - 1\right)\right)\right) dt \\
&= \int_0^1 \frac{d}{dt}\left(tg\left(a + t\left(\begin{pmatrix} x \\ y \end{pmatrix} - 1\right)\right)\right) dt \\
&= \left[tg\left(a + t\left(\begin{pmatrix} x \\ y \end{pmatrix} - a\right)\right)\right]_0^1 = g(x,y)
\end{aligned}$$

und analog folgt $\frac{\partial F}{\partial y} = h$. Dito gilt auch der n-dimensionale Fall, siehe dazu Königsberger (1903), S. 193. $\square$

1.3.25 Satz (Homotopieinvarianz des Kurvenintegrals) *Es seien γ_0, γ_1 zwei (frei) homotope geschlossene Streckenzüge im (nicht notwendig einfach zusammenhängenden) Gebiet G und $g, h\colon G \to \mathbb{R}$ stetig differenzierbar mit $g_y = h_x$. Dann gilt:*

$$\int_{\gamma_0} \left(g\,dx + h\,dy\right) = \int_{\gamma_1} \left(g\,dx + h\,dy\right)$$

unt entsprechendes für stückweise stetig differenzierbare frei homotope geschlossene Kurven.

Beweis Sei $I = [0,1]$ und $\varphi\colon I \times I \to G$ eine Homotopie, die γ_0 in γ_1 überführt wie in Definition 1.3.11 b). Dann ist $K := \varphi(I \times I) \subset G$ kompakt. Sei

$$r := \begin{cases} \frac{1}{2}\inf\left\{\|x-y\| : x \in K,\ y \in G^c\right\}, & \text{falls } G^c \neq \emptyset, \\ r > 0 \text{ beliebig}, & \text{falls, } G = \mathbb{R}^2 \end{cases}.$$

$\Rightarrow r > 0$, und für alle $(s,t) \in I \times I$ ist $K_r\left(\varphi(s,t)\right) \subset G$. φ ist stetig, also gleichmäßig stetig auf dem Kompaktum $I \times I$, d. h.

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall \begin{pmatrix} s \\ t \end{pmatrix}, \begin{pmatrix} s' \\ t' \end{pmatrix} \in I, \left\| \begin{pmatrix} s \\ t \end{pmatrix} - \begin{pmatrix} s' \\ t' \end{pmatrix} \right\|_2 < \delta \quad \left|\varphi(s,t) - \varphi\left(s',t'\right)\right| < \varepsilon.$$

Wir wählen $\varepsilon := r$ und dürfen oBdA gleich annehmen, dass das zugehörige δ die Form $\delta = \frac{2}{n}$ hat mit geeignetem $n \in \mathbb{N}$. Dann ist also mit diesem n:

$$\forall \begin{pmatrix} s \\ t \end{pmatrix}, \begin{pmatrix} s' \\ t' \end{pmatrix} \in I, \left\| \begin{pmatrix} s \\ t \end{pmatrix} - \begin{pmatrix} s' \\ t' \end{pmatrix} \right\|_2 < \frac{2}{n} \quad \left|\varphi(s,t) - \varphi\left(s',t'\right)\right| < r. \tag{1.41}$$

Wir spannen zwischen γ_0 und γ_1 ein Netz von Streckenzügen: $z_{jk} := \varphi\left(\frac{j}{n}, \frac{k}{n}\right) \in K \subset G$ für $0 \leqslant j,k \leqslant n$ gemäß (1.41). Die $z_{j,0}$ liegen also alle auf γ_0, und die $z_{j,n}$ alle auf γ_1. Nach Wahl von r und (1.41) lieft dann der Streckenzug $\sigma_{jk}\colon z_{jk} \to z_{j+1,k} \to z_{j+1,k+1} \to z_{j,k+1} \to z_{jk}$ ganz in $K_r\left(z_{jk}\right) \subset G$. σ_{jk} ist geschlossen und liegt ganz in einem Sterngebiet ($K_r\left(z_{jk}\right)$), also liefert Lemma 1.3.24 zusammen mit Satz 1.3.23 c):

$$\int_{\sigma_{jk}} \left(g\,dx + h\,dy\right) = 0 \qquad \left(0 \leqslant j,k \leqslant n\right)$$

Es sei nun $\delta_k\colon z_{0,k} \to z_{1,k} \to \ldots \to z_{n,k} = z_{0,k}$ für $k = 0, \ldots, n$. Dann ist nach Satz 1.3.23 a):

$$0 = \sum_{j=0}^{n} \int_{\sigma_{jk}} \left(g\,dx + h\,dy\right) \overset{1.3.23\,\text{a)}}{=} \int_{\delta_k} \left(g\,dx + h\,dy\right) - \int_{\delta_{k+1}} \left(g\,dx + h\,dy\right),$$

d. h. für alle $k = 0, \ldots, n-1$ ist

$$\int_{\delta_k} \left(g\,dx + h\,dy\right) = \int_{\delta_{k+1}} \left(g\,dx + h\,dy\right),$$

also

$$\int_{\delta_0} \left(g\,dx + h\,dy\right) = \int_{\delta_n} \left(g\,dx + h\,dy\right).$$

Wir müssen nur noch zeigen $\int_{\delta_0} = \int_{\gamma_0}$ und $\int_{\delta_n} = \int_{\gamma_1}$. Wir machen uns klar, dass δ_0 ein Streckenzug ist, der Punkte $z_{j,0}$ auf γ_0 verbindet. Wir definieren für $j = 0,\dots,n-1$ Wege $\tau_j\colon z_{j,0} \to z_{j+1,0} \to z_{j,0}$, wobei der „Hinweg" entlang γ_0 verlaufe und der „Rückweg" auf der Verbindungsstrecke der Punkte, die ein Teil von δ_0 ist. Nach (1.41) liegt dann τ_j ganz in $K_r\left(z_{j,0}\right)$, und Lemma 1.3.24 liefert wegen Satz 1.3.23 c): $\int_{\tau_j}\left(g\,dx + h\,dy\right) = 0$, also auch $\sum_{j=0}^{n-1}\int_{\tau_j}\left(g\,dx + h\,dy\right) = 0$.

$$\Rightarrow \int_{\gamma_0} \left(g\,dx + h\,dy\right) = \int_{\delta_0} \left(g\,dx + h\,dy\right).$$

Ebenso schließt man mit δ_n und γ_1. □

1.3.26 Korollar Ist $\gamma\colon [0,1] \to G$ nullhomotop bzgl. G und $g, h\colon G \to \mathbb{R}$ stetig differenzierbar mit $g_y = h_x$, so gilt:

$$\int_{\gamma} \left(g\,dx + h\,dy\right) = 0.$$

Beweis $\gamma \sim \gamma_1$, wobei γ_1 ein konstanter Weg ist. Dann liefert Satz 1.3.25:

$$\int_{\gamma} \left(g\,dx + h\,dy\right) = \int_{\gamma_1} \left(g\,dx + h\,dy\right) = 0,$$

da γ_1 konstant ist. □

1.3.27 Satz (Hauptsatz über Stammfunktionen) *Es seien $G \subset \mathbb{R}^2$ ein einfach zusammenhängendes Gebiet und $g, h\colon G \to \mathbb{R}$ stetig differenzierbar mit der notwendigen Exaktheitsbedingung $g_y = h_x$. Dann hat die Differentialgleichung*

$$g\left(x,y\right) + h\left(x,y\right)\frac{dy}{dx} = 0 \tag{1.42}$$

eine Stammfunktion F, und zwar folgende: Man wähle einen festen Punkt $\left(x_0,y_0\right) \in G$ und verbinde $\left(x,y\right) \in G$ mit $\left(x_0,y_0\right)$ durch einen ganz in G verlaufenden Streckenzug $\gamma_{(x_0,y_0)}^{(x,y)}$. Dann hängt

$$F\left(x,y\right) := \int_{\gamma_{(x_0,y_0)}^{(x,y)}} \left(g\left(\xi,\eta\right)\,d\xi + h\left(\xi,\eta\right)\,d\eta\right) \tag{1.43}$$

nicht ab von der Auswahl von $\gamma_{(x_0,y_0)}^{(x,y)}$ und ist eine Stammfunktion von (1.42)

BEWEIS 1. Schritt: F ist wohldefiniert.

Begründung: Wir gerbinden $(x_0, y_0) \in G$ mit $(x, y) \in G$ durch zwei stückweise stetig differenzierbare Kurven $\gamma_{(x_0,y_0)}^{(x,y)}$ und $\tilde{\gamma}_{(x_0,y_0)}^{(x,y)}$ in G und müssen zeigen:

$$\int_{\gamma_{(x_0,y_0)}^{(x,y)}} \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right) = \int_{\tilde{\gamma}_{(x_0,y_0)}^{(x,y)}} \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right)$$

Dazu beachten wir: Durch Aneinanderhängen von $\gamma_{(x_0,y_0)}^{(x,y)}$ und $\left(\tilde{\gamma}_{(x_0,y_0)}^{(x,y)}\right)^-$ entsteht eine geschlossene stückweise stetig differenzierbare Kurve δ in G. G ist einfach zusammenhängend, δ also nullhomotop. Dann liefert Korollar 1.3.26:

$$\int_\delta \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right) = 0$$

Mit Satz 1.3.23 a) folgt daraus

$$\int_{\gamma_{(x_0,y_0)}^{(x,y)}} \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right) = \int_{\tilde{\gamma}_{(x_0,y_0)}^{(x,y)}} \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right).$$

2. Schritt: F ist eine Stammfunktion. Sei $a \in G$, $r > 0$, $K_r(a) \subset G$. Wir zeigen: $F|_{K_r(a)}$ ist Stammfunktion.

Begründung: Sei $(x, y) \in K_r(a)$, $a = \binom{\alpha}{\beta}$. Nach dem 1. Schritt ist es belanglos, auf welchem Wege wir (x, y) mit (x_0, y_0) verbinden. Wir wählen den Weg wie folgt: Sei δ irgendein stückweise stetig differenzierbarer Weg in G von (x_0, y_0) nach a und $\sigma(t) := a + t\left(\binom{x}{y} - a\right)$ für $0 \leqslant t \leqslant t$. Wir wählen $\gamma_{(x_0,y_0)}^{(x,y)} := \delta \perp \sigma$ als Verkettung von δ und σ und haben

$$F(x, y) = \int_\delta \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right) + \int_\sigma \left(g(\xi,\eta)\, d\xi + h(\xi,\eta)\, d\eta\right).$$

Betrachten wir das in Abhängigkeit von $\binom{x}{y} \in K_r(a)$, so ist das δ-Integral eine Konstante, aber das σ-Integral nach Lemma 1.3.24 eine Stammfunktion auf $K_r(a)$. Also ist $F|_{K_r(a)}$ eine Stammfunktion. □

1.3.28 Beispiel Wir betrachten die sog. *Windungsform*

$$-\frac{y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dy = 0$$

auf $G := \mathbb{R}^2 \setminus \{(x, 0) : x \leqslant 0\}$. Wir wissen, dass die notwendige Exaktheitsbedingung erfüllt ist, und wir wissen nach Satz 1.3.27, dass auf G (Sterngebiet!) eine Stammfunktion existiert. Wir können das Integral explizit ausrechnen: Dabei nutzen wir aus, dass wir den Weg beliebig wählen dürfen. Wir wählen $(x_0, y_0) = (1, 0)$, setzen für $r > 0$ und $|\vartheta| < \pi$ $(x, y) = (r\cos\vartheta, r\sin\vartheta)$ und wählen den in Abb. 1.9 zu sehenden Weg.

$$F(x, y) = \int_{\gamma_{(1,0)}^{(x,y)}} \left(-\frac{\eta}{\xi^2 + \eta^2}\frac{d\xi}{dt} + \frac{\xi}{\xi^2 + \eta^2}\frac{d\eta}{dt}\right) dt.$$

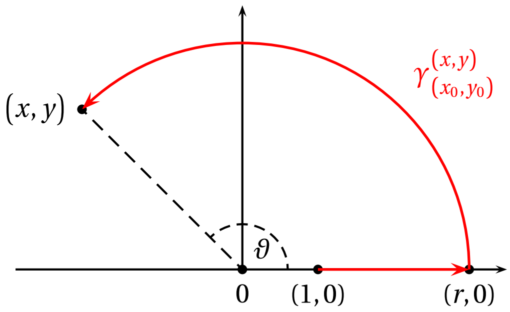

Abbildung 1.9: Der Weg aus Beispiel 1.3.28

Das Integral von $(1,0)$ nach $(r,0)$ liefert den Beitrag 0, denn auf dieser Strecke ist $\eta(t) = \frac{d\eta}{dt} = 0$. Das Integral über den Kreisbogen parametrisieren wir durch $\xi(t) = r\cos t$ und $\eta(t) = r\sin t$ für $t \in [0,\vartheta]$, bzw. $[\vartheta, 0]$ und erhalten

$$F(x,y) = \int_0^{\vartheta} \left(-\frac{r\sin t}{r^2\cos^2 t + r^2\sin^2 t}(-r\sin t) + \frac{r\cos t}{r^2\cos^2 t + r^2\sin^2 t}(r\cos t) \right) dt = \int_0^{\vartheta} 1 = \vartheta.$$

(Das stimmt auch für $-\pi < \vartheta < 0$). Das stimmt mit unserem früheren Beispiel 1.3.10 überein.

1.3.29 Definition (Integrierende Faktoren, Multiplikatoren) Oft ist die Differentialgleichung $g + hy' = 0$, $g, h\colon G \to \mathbb{R}$ stetig, nicht exakt, kann aber durch Multiplikation mit einer geeigneten stetigen Funktion $M\colon G \to \mathbb{R}$ exakt gemacht werden, wobei M (möglichst) nullstellenfrei sein soll: Die stetige Funktion $M\colon G \to \mathbb{R}$ (M nullstellenfrei) heißt ein *integrierender Faktor* oder ein *(Eulerscher) Multiplikator* für die Differentialgleichung $g + hy' = 0$ in G, $g, h\colon G \to \mathbb{R}$ stetig, falls die Differentialgleichung $Mg + Mhy' = 0$ exakt ist. Sind M, g, h stetig differenzierbar, so lautet die notwendige Exaktheitsbedingung $\left(Mg\right)_y = (Mh)_x$, d. h.

$$\begin{aligned} M_y g + M g_y &= M_x h + M h_x, \\ g M_y - h M_x &= M\left(h_x - g_y\right). \end{aligned}$$

Diese partielle Differentialgleichung ist im Allgemeinen nicht leicht zu lösen, manchmal findet man Lösungen in der Form $M = M(x)$: Dann hat man die Differentialgleichung

$$\frac{M_x}{M} = \frac{g_y - h_x}{h},$$

d. h. im Falle $M > 0$: $\left(\log M\right)' = \frac{g_y - h_x}{h}$.
Ein solches M existiert genau dann, wenn die rechte Seite nur von x (und nicht von y) abhängt. Ebenso: Ein Multiplikator $M = M(y) > 0$ existiert genau dann, wenn $\frac{h_x - g_y}{g}$ nur abhängt von y, und dann ist

$$\frac{d}{dy} \log M(y) = \frac{h_x - g_y}{g}.$$

1.3.30 Beispiel Die Differentialgleichung

$$\underbrace{\left(2x^2+2xy^2+1\right)y}_{=:g}+\underbrace{\left(3y^2+x\right)}_{=:h}y'=0$$

ist nicht exakt, da $g_y = 2x^2+6xy^2+1 \neq h_x = 1$, aber

$$\frac{g_y-h_x}{h}=\frac{2x\left(x+3y^2\right)}{3y^2+x}=2x.$$

Es gibt also einen Multiplikator $M(x)$, und zwar $M(x)=e^{x^2}$. Wir erhalten die äquivalente Differentialgleichung

$$\underbrace{e^{x^2}\left(2x^2+2xy^2+1\right)y}_{=:\tilde{g}}+\underbrace{e^{x^2}\left(3y^2+x\right)}_{=:\tilde{h}}y'=0.$$

Man rechnet leicht nach, dass

$$\tilde{g}_y = e^{x^2}\left(2x^2+6xy^2+1\right)=e^{x^2}\left(2x\left(3y^2+x\right)+1\right)=\tilde{h}_x.$$

In $\mathbb{R}^2$ hat die Differentialgleichung eine Stammfunktion F:

$$\begin{aligned}
F_x &= \tilde{g} = e^{x^2}\left(2x^2+2xy^2+1\right)y,\\
F_y &= \tilde{h} = e^{x^2}\left(3y^2+x\right).\\
\Rightarrow F(x,y) &= e^{x^2}\left(y^3+xy\right)+\varphi(x)\\
\Rightarrow F_x &= 2xe^{x^2}\left(y^3+xy\right)+e^{x^2}y+\varphi'(x)
\end{aligned}$$

$\Rightarrow \varphi'=0$ durch Koeffizientenvergleich, wir können also $\varphi(x)=0$ wählen.

$$F(x,y)=e^{x^2}\left(y^3+xy\right)$$

ist also eine Stammfunktion auf $\mathbb{R}^2$. Die Gleichung $F(x,y)=F(x_0,y_0)$ lässt sich allgemein weder nach x noch nach y auflösen, in speziellen Fällen aber doch, nämlich wenn $y_0^3+x_0y_0=0$ ist.

1.4 Systeme linearer Differentialgleichungen mit konstanten Koeffizienten

Sei $I\subset\mathbb{R}$ ein Intervall, $y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}: I\to\mathbb{R}^n$ differenzierbar, $A\in M(n\times n,\mathbb{R})$. Eine Differentialgleichung des Typs

$$y' = Ay \qquad (x\in I) \tag{1.44}$$

heißt ein *System linearer Differentialgleichungen mit konstanten Koeffizienten*; ausführlich: Sei $A = \left(a_{jk}\right)_{1\leqslant j,k\leqslant n}$:

$$
\begin{aligned}
y_1' &= a_{1,1}y_1 + a_{1,2}y_2 + \ldots + a_{1,n}y_n,\\
&\vdots\\
y_n' &= a_{n,1}y_1 + a_{n,2}y_2 + \ldots + a_{n,n}y_n.
\end{aligned}
$$

Anfangswertproblem: Sind $x_0 \in I$, $b \in \mathbb{R}^n$ gegeben, so bestimme man eine Lösung von (1.44) mit $y(x_0) = b$. Wir werden sehen: Dieses Anfangswertproblem ist eindeutig lösbar, die Lösung existiert auf ganz I und kann explizit angegeben werden.

Bemerkung Auf den ersten Blick erscheinen solche System recht abstrakt. Wir werden ihre Nützlichkeit aber noch sehen, denn die Lösungstheorie der linearen Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + a_0y = 0 \qquad \left(y\colon I \to \mathbb{R}\right)$$

ergibt sich ganz leicht aus der Lösungstheorie für das System (1.44).

Lösungsansatz: Im Fall $n = 1$ hat man $y' = ay$, $y\colon I \to \mathbb{R}$, $a \in \mathbb{R}$ und die allgemeine Lösung $y = y_0 e^{a(x-x_0)}$ für $x \in I$.
Idee: Wen man e^A für Matrizen definieren könnte unter Erhalt der wichtigsten Rechenregeln, so wird $y = e^{A(x-x_0)}b$ mit $b \in \mathbb{R}^n$ eine Lösung von (1.44) sein mit $y(x_0) = b$.
Ansatz: $e^A := \sum_{k=0}^{\infty} \frac{1}{k!}A^k$. Wir behandeln im folgenden die Fälle $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ parallel. Wir begeben uns auf einen kleinen Exkurs über die Exponentialfunktion von Matrizen:

1.4.1 Lemma Für alle $A \in M(n\times n,\mathbb{K})$ konvergiert die Reihe $e^A := \sum_{m=0}^{\infty} \frac{1}{m!}A^m$ im Sinne elementweiser Konvergenz der Folge der Teilsummen $\left(\sum_{m=0}^{p} \frac{1}{m!}A^m\right)_{p\geqslant 0}$, d. h. im Sinne der Norm $\|X\| := \left(\sum_{j,k=1}^{n} \left|x_{j,k}\right|^2\right)^{\frac{1}{2}}$ auf dem $\mathbb{K}$-Vektorraum der $(n\times n)$-Matrizen. Für jedes $\mu > 0$ ist die Konvergenz gleichmäßig auf der Menge

$$\mathcal{M}_\mu := \left\{A = \left(\alpha_{j,k}\right) \in M(n\times n,\mathbb{K}) : \left|\alpha_{j,k}\right| \leqslant \mu \text{ für alle } j,k = 1,\ldots,n\right\}.$$

Beweis Sei $\mu > 0$ und $A \in \mathcal{M}_\mu$; $A^m = \left(\alpha_{j,k}^{(m)}\right)_{j,k=1,\ldots,n}$ für $m \geqslant 0$. Dann ist $\left|\alpha_{j,k}^{(0)}\right| = \left|\delta_{j,k}\right| \leqslant 1 = (n\mu)^0$ und $\left|\alpha_{j,k}^{(1)}\right| \leqslant \mu \leqslant (n\mu)^1$; induktiv schließen wir weiter: Ist $m \geqslant 0$ und bekannt, dass $\left|\alpha_{j,k}^{(m)}\right| \leqslant (n\mu)^m$ für alle $j,k = 1,\ldots,n$, so ist

$$\left|\alpha_{j,k}^{(m+1)}\right| = \left|\sum_{\ell=1}^{n} \underbrace{\alpha_{j,\ell}^{(m)}}_{\leqslant (n\mu)^m} \underbrace{\alpha_{\ell,k}}_{\leqslant\mu}\right| \leqslant \sum_{\ell=1}^{n} (n\mu)^m \mu = (n\mu)^{m+1}$$

$\Rightarrow \sum_{m=0}^{\infty} \frac{\alpha_{j,k}^{(m)}}{m!}$ konvergiert, denn $\left|\frac{\alpha_{j,k}^{(m)}}{m!}\right| \leqslant \frac{(n\mu)^m}{m!}$, und das ist der m-te Term der Reihe

$$\sum_{m=0}^{\infty} \frac{(n\mu)^m}{m!} = \exp(n\mu).$$

Der Weierstraßsche Majorantentest liefert jetzt die Behauptung. □

1.4.2 Lemma a) Ist $M \in \mathrm{GL}_n(\mathbb{K})$, so gilt: $e^{M^{-1}AM} = M^{-1}e^A M$.

b) $\exp(\operatorname{diag}(\lambda_1, \ldots, \lambda_n)) = \operatorname{diag}\left(e^{\lambda_1}, \ldots, e^{\lambda_n}\right)$.

c) Sind $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ die mit algebraischer Vielfachheit gezählten Eigenwerte von A, so sind $e^{\lambda_1}, \ldots, e^{\lambda_n}$ die mit algebraischer Vielfachheit gezählten Eigenwerte von e^A.

Beweis a) klar, denn

$$\sum_{m=0}^{p} \frac{1}{m!}\left(M^{-1}AM\right)^m = M^{-1}\left(\sum_{m=0}^{p} \frac{1}{m!}A^m\right)M,$$

und $p \to \infty$ liefert a).

b) klar

c) Über $\mathbb{C}$ ist A zu einer oberen Dreiecksmatrix mit $\lambda_1, \ldots, \lambda_n$ auf der Hauptdiagonalen ähnlich: Es gibt ein $M \in \mathrm{GL}_n(\mathbb{C})$ mit

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Dann ist

$$\left(M^{-1}AM\right)^m = \begin{pmatrix} \lambda_1^m & & * \\ & \ddots & \\ 0 & & \lambda_n^m \end{pmatrix},$$

also nach a):

$$M^{-1}e^A M = e^{M^{-1}AM} = \begin{pmatrix} e^{\lambda_1} & & * \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}.$$

Das liefert die Behauptung. □

1.4.3 Korollar $\det e^A = e^{\operatorname{Spur} A}$; insbesondere ist $e^A \in \mathrm{GL}_n(\mathbb{K})$ für alle $A \in M(n \times n, \mathbb{K})$.

Beweis Nach Lemma 1.4.2 c) ist

$$\det e^A = e^{\lambda_1+\ldots+\lambda_n} = e^{\operatorname{Spur} A}. \qquad \square$$

1.4.4 Beispiele a) Sei A diagonalisierbar, $A = M^{-1}DM$, $D = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$. Nach Lemma 1.4.2 b) ist $e^D = \operatorname{diag}\left(e^{\lambda_1},\ldots,e^{\lambda_n}\right)$, also $e^A = M^{-1}\operatorname{diag}\left(e^{\lambda_1},\ldots,e^{\lambda_n}\right)M$.

b) Zerfällt das charakteristische Polynom von A über $\mathbb{K}$ in Linearfaktoren (über $\mathbb{C}$ gilt das immer), so ist $A = M^{-1}(D+N)M$ mit $D = \operatorname{diag}(\lambda_1,\ldots,\lambda_n)$ und nilpotentem N, $DN = ND$ (z. B. Jordan-Normalform). $\Rightarrow e^A = M^{-1}e^D e^N M$, denn nach Satz 1.4.5 ist $e^{D+N} = e^D e^N$, und hier ist $e^N = \sum_{m=0}^n \frac{1}{m!}N^m$, da $N^k = 0$ für alle $k > n$. Ist also die Jordan-Normalform bekannt, so ist $e^A = M^{-1}e^D e^N M$ elementar berechenbar. Ein bequemes Verfahren zur Berechnung von e^A findet sich bei Gantmacher (1965).

c) $A = \left(\begin{smallmatrix} 5 & -4 \\ 4 & -3\end{smallmatrix}\right)$ hat das charakteristische Polynom

$$p_A = |A - XE_2| = \begin{vmatrix} 5-X & -4 \\ 4 & -3-X \end{vmatrix} = X^2 - 2X + 1 = (X-1)^2,$$

ist aber nicht diagonalisierbar, da sonst A zu E_2 ähnlich sein, also gleich E_2 sein müsste. $\Rightarrow A = E_2 + N$, wobei $N = \left(\begin{smallmatrix} 4 & -4 \\ 4 & -4\end{smallmatrix}\right)$, $N^2 = 0$.

$$\Rightarrow e^A = e^{E_2}e^N = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix}(E_2 + N) = eA.$$

Für $x \in \mathbb{R}$ hat Ax die Minimalzerlegung $Ax = (xE_2) + (xN)$, also

$$e^{Ax} = \begin{pmatrix} e^x & 0 \\ 0 & e^x \end{pmatrix}(E_2 + xN) = e^x(E_2 + xN).$$

1.4.5 Satz *Sind $A, B \in M(n \times n, \mathbb{K})$ vertauschbar, d. h. $AB = BA$, so gilt: $e^{A+B} = e^A e^B$.*

Beweis

$$(A+B)^m = \sum_{j=0}^m \binom{m}{j} A^j B^{m-j},$$

da $AB = BA$. Daher gilt für alle $p \in \mathbb{N}$:

$$\sum_{m=0}^{2p} \underbrace{\frac{1}{m!}(A+B)^m}_{\sum_{k+\ell=m}\frac{1}{k!}A^k\frac{1}{\ell!}B^\ell} = \left(\sum_{k=0}^p \frac{1}{k!}A^k\right)\left(\sum_{\ell=0}^p \frac{1}{\ell!}B^\ell\right) + R_p, \tag{1.45}$$

wobei $R_p = \sum_{\substack{k,\ell\geqslant 0 \\ \max(k,\ell)>p \\ k+\ell\leqslant 2p}} \frac{1}{k!}A^k + \frac{1}{\ell!}B^\ell$. Die Anzahl der Summanden in dieser Darstellung von R_p beträgt

$$2\sum_{k=p+1}^{2p}\left(2p-k+1\right) = 2\sum_{\nu=1}^p \nu = p\left(p+1\right).$$

Sei S ein Matrixelement im Term $\frac{1}{k!}A^k\frac{1}{\ell!}B^\ell$, n die Anzahl der Summanden im Produkt A^kB^ℓ, μ wie in Lemma 1.4.1. Dann gilt

$$|S| \leqslant n\frac{(n\mu)^k}{k!}\frac{(n\mu)^\ell}{\ell!} \overset{\max(k,\ell)>p}{\leqslant} n\frac{(n\mu)^{2p}}{p!}.$$

Dann sind alle Koeffizienten von R_p bereits $\leqslant np(p+1)\frac{\left((n\mu)^2\right)^p}{p!} \xrightarrow{p\to\infty} 0$ (vgl. Konvergenz-Beweis für die Exponential-Reihe). $\Rightarrow R_p \xrightarrow{p\to\infty} 0 \in M(n\times n,\mathbb{K})$. Lassen wir in (1.45) nun $p\to\infty$ gehen, folgt die Behauptung. □

1.4.6 Korollar $e^{-A} = \left(e^A\right)^{-1}$ für alle $A \in M(n\times n,\mathbb{K})$.

Beweis

$$e^Ae^{-A} \overset{\text{Satz 1.4.5}}{=} e^{A-A} = e^0 = E_n.$$ □

1.4.7 Satz *Für jedes $A \in M(n\times n,\mathbb{K})$ ist die Funktion $\mathbb{R} \ni x \mapsto e^{Ax}$ elementweise beliebig oft differenzierbar, und es gilt:*

$$\frac{d}{dx}e^{Ax} = Ae^{Ax} \qquad (x \in \mathbb{R}).$$

Beweis In der Notation des Beweises von Lemma 1.4.1 ist

$$e^{Ax} = \left(\sum_{m=0}^{\infty}\frac{\alpha_{j,k}^{(m)}}{m!}x^m\right)_{j,k=1,\dots,n},$$

und hier konvergieren die Potenzreihen für alle $x \in \mathbb{R}$, dürfen also beliebig oft termweise differenziert werden. Das liefert

$$\frac{d}{dx}e^{Ax} = \left(\sum_{m=1}^{\infty}\frac{\alpha_{j,k}^{(m)}}{m!}mx^{m-1}\right)_{j,k=1,\dots,n} = \left(\sum_{m=0}^{\infty}\frac{\alpha_{j,k}^{(m+1)}}{m!}x^m\right)_{j,k=1,\dots,n} = Ae^{Ax}.$$ □

Kehren wir nun zurück zur Betrachtung von Differentialgleichungen:

1.4.8 Satz *Vorgelegt sei die Differentialgleichung*

$$y' = Ay \tag{1.46}$$

auf dem Intervall I mit $A \in M(n\times n,\mathbb{K})$, $y\colon I \to \mathbb{K}^n$, $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$. Dann gilt:

a) *Für jedes $x_0 \in I$, $y_0 \in \mathbb{K}^n$ hat das Anfangswertproblem $y(x_0) = y_0$ für die Differentialgleichung (1.46) genau eine Lösung und zwar*

$$y(x) = e^{A(x-x_0)}y_0 \qquad (x \in I). \tag{1.47}$$

Speziell gilt die Alternative: Entweder ist $y = 0$ (im Fall $y_0 = 0$), oder es ist $y(x) \neq 0$ für alle $x \in I$ (im Fall $y_0 \neq 0$). Beachte: Auch im ersten Fall können gewisse Koordinatenfunktionen von $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ Nullstellen haben, aber es gibt keine gemeinsame Nullstelle aller Koordinatenfunktionen.

b) *Für jedes $x_0 \in I$ ist $\mathbb{K}^n \ni y_0 \mapsto y$, wobei y Lösung von* (1.47) *sei, ein Isomorphismus des $\mathbb{K}$-Vektorraums $\mathbb{K}^n$ auf den $\mathbb{K}$-Vektorraum $\mathscr{L}$ der Lösungen von* (1.46). *Insbesondere hat $\mathscr{L}$ die Dimension n über $\mathbb{K}$.*

Beweis a) Nach Satz 1.4.7 gilt: $\frac{d}{dx} e^{A(x-x_0)} = Ae^{A(x-x_0)}$. Also ist $y(x) = e^{A(x-x_0)} y_0$ für alle $x \in I$ Lösung von (1.46), und zwar erfüllt diese Lösung die Anfangsbedingung $y(x_0) = y_0$. Sei y irgendeine Lösung von (1.46) mit $y(x_0) = y_0$.

$$\Rightarrow \frac{d}{dx}\left(e^{-A(x-x_0)} y\right) = -Ae^{-A(x-x_0)} y + e^{-A(x-x_0)} y' = -Ae^{-A(x-x_0)} y + e^{-A(x-x_0)} Ay = 0.$$

$e^{-A(x-x_0)} y$ ist also koordinatenweise konstant gleich dem Wert an der Stelle x_0, und der ist y_0, also: $y(x) = e^{A(x-x_0)} y_0$ nach Korollar 1.4.6.
Alternativ zeigt man $y_0 = 0 \Rightarrow y = 0$. $y_0 \neq 0 \Rightarrow y(x) \neq 0$ für alle $x \in I$, da $e^{A(x-x_0)} \in \mathrm{GL}_n(\mathbb{K})$.

b) Die Abbildung $\mathbb{K}^n \ni y_0 \mapsto y$, wobei y Lösung von (1.47) sei, ist nach a) injektiv mit Umkehrabbildung $\mathscr{L} \ni y \mapsto y(x_0) \in \mathbb{K}^n$. Das liefert die Behauptung. □

Ist die Jordansche Normalform von A bekannt, so kann man $e^{A(x-x_0)}$ sofort hinschreiben und (1.46) bei beliebiger rechter Seite durch (1.47) lösen.

1.4.9 Beispiel $I = \mathbb{R}$, $A = \left(\begin{smallmatrix} 5 & -4 \\ 4 & -3 \end{smallmatrix}\right)$, die Differentialgleichung $y' = Ay$, d. h.

$$y_1' = 5y_1 - 4y_2$$
$$y_2' = 4y_1 - 3y_2$$

Oben haben wir schon e^{Ax} ausgerechnet, und wir finden: Für $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^2$ ist

$$y(x)\, e^{(x-x_0)} (E_2 + (x-x_0) N)\, y_0 \qquad (N := A - E_2)$$

die Lösung des Anfangswertproblems, d. h.

$$y(x) = \left(y_0 + (x-x_0)\left(Ny_0\right)\right) e^{x-x_0}.$$

Probe: $y(x_0) = y_0$ ist klar.

$$y'(x) = \underbrace{\left(Ny_0\right) e^{x-x_0}}_{=Ny \text{ wegen } N^2=0} + \underbrace{\left(y_= + (x-x_0)\, Ny_0\right) e^{x-x_0}}_{=y(x)} = (E_2 + N)\, y = Ay.$$

1.4.10 Beispiel Ist $A \in M(n \times n, \mathbb{K})$ nilppotent, $A^\mu = 0$, so hat $y' = Ay$, $y(x_0) = y_0$ die Lösung

$$y(x) = e^{A(x-x_0)} y_0 = \left(\sum_{\nu=0}^{\mu-1} \frac{(x-x_0)^\nu}{\nu!} A^\nu \right) y_0.$$

Probe: $y' = \left(\sum_{\nu=1}^{\mu-1} \frac{(x-x_0)^{\nu-1}}{(\nu-1)!} A^\nu \right) y_0 = Ay$. Hier erhalten wir also reine Polynomfunktionen als Lösungen.

1.4.11 Satz (Basis des Lösungsraums) *Wir wollen eine Basis des Lösungeraums $\mathscr{L}$ der Differentialgleichung $y' = Ay$ bestimmen, ohne die Jordan-Form von A auszurechnen. Sei $A \in M(n \times n, \mathbb{K})$, $p_A := \det(A - XE_n)$ zerfalle über $\mathbb{K}$ in Linearfaktoren (automatisch für $\mathbb{K} = \mathbb{C}$), $\lambda_1, \dots, \lambda_k$ seien die verschiedenen Eigenwerte von A. Dann ist*

$$p_A = (-1)^n \prod_{j=1}^{k} \left(X - \lambda_j\right)^{m_j},$$

$$m_A = \prod_{j=1}^{k} \left(X - \lambda_j\right)^{\mu_j} = \textit{Minimalpolynom von } A.$$

Setzen wir $U_j := \ker\left(A - \lambda_j E_n\right)^{\mu_j}$ für $j = 1, \dots, k$, so ist nach dem Satz von der Minimalzerlegung (Jordan-Form) $\dim U_j = m_j$ für $j = 1, \dots, k$ und $\mathbb{K}^n = \bigoplus_{j=1}^k U_j$. Für $j = 1, \dots, k$ ist U_j A-invariant, und $\left(A - \lambda_j E_n\right)|_{U_j}$ ist nilpotent vom Index μ_j. Für jedes $c \in U_j$ ist daher

$$\begin{aligned} e^{A(x-x_0)} c &= e^{\left(A-\lambda_j E_n\right)(x-x_0)} \underbrace{e^{\lambda_j (x-x_0) E_n}}_{= e^{\lambda_j (x-x_0)} E_n} c \\ &= e^{\lambda_j (x-x_=)} e^{\left(A-\lambda_j E_n\right)(x-x_0)} c \\ &= e^{\lambda_j (x-x_0)} \sum_{\nu=0}^{\mu_j - 1} \frac{(x-x_0)^\nu}{\nu!} \left(A - \lambda_j E_n\right)^\nu c. \end{aligned}$$

Wenn die Anfangswerte eine Basis von $\mathbb{K}$ durchlaufen, durchläuft das entsprechende y eine Basis von $\mathscr{L}$ nach Satz 1.4.8.
Ergebnis: Durchläuft c eine Basis B_j von U_j, so durchläuft

$$y_c(x) := \left(\sum_{\nu=0}^{\mu_j - 1} \frac{(x-x_0)^\nu}{\nu!} \left(A - \lambda_j E_n\right)^\nu c \right) e^{\lambda_j (x-x_0)} \qquad \left(c \in B_j,\ j = 1, \dots, k\right)$$

eine Basis von $\mathscr{L}$. Die effektive Bestimmung der c erfordert die Bestimmung einer Basis von Lösungen für ein homogenes lineares Gleichungssystem. Diese Basis von $\mathscr{L}$ ist also durch die Anfangsbedingung $y_c(x_0) = c$ eindeutig festgelegt. Dieses Verfahren klappt immer im Falle $\mathbb{K} = \mathbb{C}$, und im Falle $\mathbb{K} = \mathbb{R}$ sicher dann, wenn p_A über $\mathbb{R}$ zerfällt. Im Falle $A \in M(n \times n, \mathbb{R})$ kann man auch nach obigem Muster verfahren, wenn man über $\mathbb{C}$ zerlegt und eventuell komplexwertige Lösungen zulässt. Wenn man reellwertige Lösungen haben möchte, geht man

wie folgt vor:
Wir setzen nun voraus: Sei $A \in M(n \times n, \mathbb{R})$, $\mathbb{K} = \mathbb{R}$,

$$p_A = (-1)^n \prod_{j=1}^{k} \left(X - \lambda_j\right)^{m_j},$$

$$m_A = \prod_{j=1}^{k} \left(X - \lambda_j\right)^{\mu_j}$$

mit $\lambda_1, \dots, \lambda_r \in \mathbb{R}$, $\lambda_{r+1}, \dots, \lambda_k \in \mathbb{C}$, *wobei* $k - r = 2s$, $\operatorname{Im}\lambda_{r+1}, \dots, \operatorname{Im}\lambda_{r+s} > 0$, $\lambda_{r+s+\nu} = \overline{\lambda_{r+\nu}}$ *für* $\nu = 1, \dots, s$, *also* $\operatorname{Im}\lambda_{r+s+1}, \dots, \operatorname{Im}\lambda_{r+2s} < 0$. *Für* $j = 1, \dots, r$ *ist immer noch* $\dim_{\mathbb{R}} \ker\left(A - \lambda_j E_n\right)^{\mu_j} = m_j$, *und setzen wir* B_j *als Basis von* $\ker_{\mathbb{R}^n}\left(A - \lambda_j E_n\right)^{\mu_j}$ *für* $j = 1, \dots, r$, *so sind die zugehörigen* y_c $(c \in B_j,\ j = 1, \dots, r)$ *reell und natürlich linear unabhängig. Für* $j = r+1, \dots, r+s$ *sei* B_j *eine Basis von* $\ker_{\mathbb{C}^n}\left(A - \lambda_j E_n\right)^{\mu_j}$, *wobei wir* $\left(A - \lambda_j E_n\right)^{\mu_j}$ *als Endomorphismus von* $\mathbb{C}^n$ *auffassen,* $c \in B_j$ *und* y_c *wie oben. Dann sind die Funktionen* $\operatorname{Re} y_c \colon I \to \mathbb{R}^n$ *und* $\operatorname{Im} y_c \colon I \to \mathbb{R}^n$ *reellwertig und*

$$\frac{d}{dx}\left(\operatorname{Re} y_c(x)\right) = \operatorname{Re}\left(\frac{d}{dx} y_c(x)\right) = \operatorname{Re}\left(A y_c(x)\right) = A\left(\operatorname{Re} y_c(x)\right)$$

und ebenso mit $\operatorname{Im} y_c$.
Behauptung: Die Funktionen $y_c(x)$ $(c \in B_j,\ j = 1, \dots, r)$ *zusammen mit den Funktionen* $\operatorname{Re} y_c(x)$, $\operatorname{Im} y_c(x)$ $(c \in B_j,\ j = r+1, \dots, r+s)$ *bilden eine Basis von* $\mathscr{L}$ *über* $\mathbb{R}$.

Beweis Es ist nur zu zeigen, dass die obigen $r + 2s = n$ Funktionen linear unabhängig sind über $\mathbb{R}$. Ist nun $c \in \ker_{\mathbb{C}^n}\left(A - \lambda_j E_n\right)^{\mu_j}$ für $r + 1 \leqslant j \leqslant r + s$, so ist

$$\overline{c} \in \ker_{\mathbb{C}^n}\left(A - \overline{\lambda_j} E_n\right)^{\mu_j} = \ker_{\mathbb{C}^n}\left(A - \lambda_{j+s} E_n\right)^{\mu_j}$$

(beachte: $\mu_j = \mu_{j+s}$!), und $B_{j+s} := \left\{\overline{c} \colon c \in B_j\right\}$ für $j = r, \dots, r+s$ ist eine Basis von U_{j+s}. Für die obigen y_c gilt bei dieser Basiswahl: $\overline{y_c} = y_{\overline{c}}$, wobei y_c zu λ_j, $y_{\overline{c}}$ zu $\overline{\lambda_{j+s}}$ gehört. Sei $B_j := \left\{c_{j,\nu} \colon \nu = 1, \dots, m_j\right\}$ für $j = 1, \dots, r, r+1, \dots, r+s$. Angenommen, mit $\alpha_{j,\nu}, \alpha'_{j,\nu} \in \mathbb{R}$ ist

$$\sum_{j=1}^{r} \sum_{\nu=1}^{m_j} \alpha_{j,\nu} y_{c_{j,\nu}} + \sum_{j=r+1}^{r+s} \sum_{\nu=1}^{m_j} \alpha_{j,\nu} \operatorname{Re} y_{c_{j,\nu}} + \sum_{j=r+1}^{r+s} \sum_{\nu=1}^{m_j} \alpha'_{j,\nu} \operatorname{Im} y_{c_{j,\nu}} = 0.$$

Dann liefert $\operatorname{Re} y_c = \frac{1}{2}\left(y_c + y_{\overline{c}}\right)$ und $\operatorname{Im} y_c = \frac{1}{2i}\left(y_c - y_{\overline{c}}\right)$:

$$\sum_{j=1}^{r} \sum_{\nu=1}^{m_j} \alpha_{j,\nu} y_{c_{j,\nu}} + \sum_{j=r+1}^{r+s} \sum_{\nu=1}^{m_j} \frac{1}{2}\left(\alpha_{j,\nu} + \frac{1}{i}\alpha'_{j,\nu}\right) y_{c_{j,\nu}} + \sum_{j=r+1}^{r+s} \sum_{\nu=1}^{m_j} \frac{1}{2}\left(\alpha_{j,\nu} - \frac{1}{i}\alpha'_{j,\nu}\right) y_{\overline{c_{j,\nu}}} = 0.$$

Die hier auftretenden $y_{c_{j,\nu}}$ etc. sind nach dem schon Bewiesenen linear unabhängig über $\mathbb{C}$, also $\alpha_{j,\nu} = 0$ für $j = 1, \dots, r$, $\nu = 1, \dots, m_j$ und $\alpha_{j,\nu} pmi \alpha'_{j,\nu} = 0$ für $j = r+1, \dots, r+s$, $\nu = 1, \dots, m_j$. Also sind alle $\alpha_{j,\nu} = 0$ und alle $\alpha'_{j,\nu} = 0$. Das liefert die Behauptung. □

1.4.12 Beispiel $A = \begin{pmatrix} 1 & 1 & 5 & 1 \\ -1 & -1 & -1 & 3 \\ 5 & 1 & 1 & 1 \\ -1 & 3 & -1 & -1 \end{pmatrix}$. $\Rightarrow p_A = (X-4)^2\,(X+4)^2$ und $m_A = (X-4)^2\,(X+4)$ zerfallen über $\mathbb{R}$, so dass die obige Konstruktion unnötig ist. $\lambda_1 = 4$, $\lambda_2 = -4$, $m_1 = m_2 = 2$, $\mu_1 = \mu_2 = 2$, $U_1 = \ker(A-4E_4)^2$. Wir berechnen

$$(A-4E_4)^2 = \begin{pmatrix} 32 & 0 & -32 & 0 \\ 0 & 32 & 0 & -32 \\ -32 & 0 & 32 & 0 \\ 0 & -32 & 0 & 32 \end{pmatrix}$$

hat den Rang 2, also $\dim U_1 = 2 = m_1$.

$$U_2 = \ker(A+4E_4) = \ker \begin{pmatrix} 5 & 1 & 5 & 1 \\ -1 & 3 & -1 & 3 \\ 5 & 1 & 5 & 1 \\ -1 & 3 & -1 & 3 \end{pmatrix}.$$

Die Matrix hat den Rang 2, also $\dim U_2 = 2 = m_2$. Die Basis B_1 von U_1 sieht wie folgt aus:

$$B_1 = \left\{ c_{1,1} := \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix},\ c_{1,2} := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

wobei $c_{1,1}$ Eigenvektor von A zum Eigenwert 4 ist:

$$(A-4E_4)\,c_{1,1} = \begin{pmatrix} -3 & 1 & 5 & 1 \\ -1 & -5 & -1 & 3 \\ 5 & 1 & -3 & 1 \\ -1 & 3 & -1 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = 0.$$

Weiter gilt

$$(A-4E_4)\,c_{1,2} = \begin{pmatrix} -3 & 1 & 5 & 1 \\ -1 & -5 & -1 & 3 \\ 5 & 1 & -3 & 1 \\ -1 & 3 & -1 & -5 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix} = 2c_{1,1}.$$

Die zugehörigen Lösungen lauten:

$$y_{c_{1,1}}(x) = e^{4(x-x_0)}\,c_{1,1}$$
$$y_{c_{1,2}}(x) = e^{4(x-x_0)}\,(E_4 + (x-x_0)(A-4E_4))\,c_{1,2} = e^{4(x-x_0)}\left(c_{1,2} + 2\,(x-x_0)\,c_{1,1}\right).$$

Die Basis B_2 von U_2 sieht so aus:

$$B_2 := \left\{ c_{2,1} := \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix},\ c_{2,2} := \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}.$$

Die zugehörigen Lösungen:

$$y_{2,1}(x) = c_{2,1} e^{-4(x-x_0)},$$
$$y_{2,2}(x) = c_{2,2} e^{-4(x-x_0)}.$$

Insgesamt ist $y_{\mu,\nu}$ $(\mu,\nu = 1,2)$ eine Basis des Lösungsraums von $y' = Ay$.

Bisher haben wir nur die homogene Gleichung $y' = Ay$ betrachtet; nun wenden wir uns der inhomogenen Gleichung $y' = Ay + f$ zu:

1.4.13 Satz *Es seien $A \in M(n \times n, \mathbb{K})$, $I \subset \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{K}^n$ stetig, $x_0 \in I$ und $y_0 \in \mathbb{K}^n$. Dann hat die Differentialgleichung*

$$y' = Ay + f \tag{1.48}$$

genau eine stetig differenzierbare Lösung $y\colon I \to \mathbb{K}^n$ mit

$$y(x_0) = y_0, \tag{1.49}$$

nämlich

$$y(x) = e^{A(x-x_0)} y_0 + \int_{x_0}^{x} e^{A(x-t)} f(t)\, dt. \tag{1.50}$$

Beweis Sei y eine Lösung von (1.48) mit (1.49), $z(x) := e^{-A(x-x_0)} y(x)$. Dann ist

$$z'(x) = -Ae^{-A(x-x_0)} y(x) + e^{-A(x-x_0)} y'(x) = e^{-A(x-x_0)} f(x),$$

$z(x_0) = y_0$.

$$\Rightarrow z(x) = y_0 + \int_{x_0}^{x} e^{-A(t-x_0)} f(t)\, dt.$$
$$\overset{\text{Satz 1.4.5}}{\Rightarrow} y(x) = e^{A(x-x_0)} y_0 + \int_{x_0}^{x} e^{A(x-t)} f(t)\, dt,$$

d. h. y ist durch (1.50) gegeben. – Umgekehrt: Die Probe bestätigt: (1.50) ist Lösung von (1.48) mit (1.49). □

Ein schönes Beispiel ist der freie Fall im Vakuum in der Nähe der Erdoberfläche unter Berücksichtigung der Erdrotation, zu finden in Gantmacher (1965), S. 111 ff.

1.5 Lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten

1.5.1 Definition (Differentialgleichung n-ter Ordnung) Sind $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{K}$, $I \subset \mathbb{R}$ ein Intervall, $y\colon I \to \mathbb{K}$, $g\colon I \to \mathbb{K}$, so heißt

$$y^{(n)} + \alpha_{n-1} y^{(n-1)} + \ldots + \alpha_1 y' + \alpha_0 y = g \tag{1.51}$$

eine inhomogene lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten. Ist $g = 0$, also

$$y^{(n)} + \alpha_{n-1} y^{(n-1)} + \ldots + \alpha_1 y' + \alpha_0 y = 0, \tag{1.52}$$

so heißt die Differentialgleichung homogen.

1.5.2 Folgerung a) Die Menge $V := \{y\colon I \to \mathbb{K}\colon y\ n\text{-mal differenzierbar mit } (1.52)\}$ ist ein $\mathbb{K}$-Vektorraum; alle $y \in V$ sind beliebig oft differenzierbar.

b) Ist (1.51) lösbar, so ist die Menge der Lösungen von (1.51) eine Restklasse modulo V, d. h. sind y_1, y_2 zwei Lösungen von (1.51), so ist $y_1 - y_2 \in V$, und ist y_1 eine Lösung von (1.51), so ist $\{y_1 + y\colon y \in V\}$ die Menge aller Lösungen von (1.51).

Welche Dimension hat V? Diese Frage lässt sich am einfachsten mit Hilfe eines passenden linearen Differentialgleichungs-Systems beantworten.

1.5.3 Konstruktion Homogene lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten und zugeordnete homogene Systeme linearer Differentialgleichungen erster Ordnung mit konstanten Koeffizienten.
Sei $y\colon I \to \mathbb{K}$ eine n-mal differenzierbare Lösung von (1.52). Wir setzen

$$z := \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \colon I \to \mathbb{K}^n, \qquad z_k := y^{(k-1)} \quad (k = 1, \ldots, n)$$

Dann liefert (1.52): $z_1' = z_2, \ldots, z_{n-2}' = z_{n-1}$ und $z_n' = -\alpha_0 z_1 - \ldots - \alpha_{n-1} z_n$, oder anders geschrieben

$$z' = Az \qquad \text{mit} \quad A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{n-1} \end{pmatrix} \tag{1.53}$$

Ersichtlich ist die Abbildung $y \mapsto z$, wobei y Lösung von (1.52) und z Lösung von (1.53) sei, linear und injektiv. Umgekehrt: Ist $z\colon I \to \mathbb{K}^n$ eine Lösung von (1.53), so ist $y := z_1\colon I \to \mathbb{K}$ eine n-mal differenzierbare Lösung von (1.52). Die Abbildung

$$\varphi\colon V \to \mathcal{L} := \{z\colon I \to \mathbb{K}^n \text{ differenzierbar}\colon z' = Az\}, \qquad \varphi(y) := z$$

ist offenbar ein Vektorraum-Isomorphismus. Aber wir wissen: $\dim_{\mathbb{K}} \mathcal{L} = n$ nach Satz 1.4.8 b).□

1.5.4 Satz *a) Der $\mathbb{K}$-Vektorraum V der Lösungen der homogenen linearen Differentialgleichung* (1.52) *mit konstanten Koeffizienten hat die Dimension n:* $\dim_{\mathbb{K}} V = n$.

b) *Sind* $x_0 \in I$, $y_0, y_1, \ldots, y_{n-1} \in \mathbb{R}$ *beliebig vorgegeben, so gibt es genau eine Lösung* $y\colon I \to \mathbb{R}$ *von* (1.52) *mit* $y(x_0) = x_0$, $y'(x_0) = y_1, \ldots, y^{(n-1)}(x_0) = y_{n-1}$.

Beweis a) $\dim_{\mathbb{K}} V \overset{\text{Konstruktion 1.5.3}}{=} \dim_{\mathbb{K}} \mathscr{L} \overset{\text{Satz 1.4.8 b)}}{=} n.$

b) Zu $z_0 := \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix} \in \mathbb{K}^n$ existiert nach Satz 1.4.8 a) genau ein $z \in \mathscr{L}$ mit $z(x_0) = z_0$. Dann leistet $y := \varphi^{-1}(z) = z_1$ das Verlangte und ist wegen $V \simeq \mathscr{L}$ nach Satz 1.4.8 eindeutig bestimmt. □

1.5.5 Satz (Basissatz) *Vorgelegt sei die Differentialgleichung* (1.52) *über* $\mathbb{K}$, $y\colon I \to \mathbb{K}$ *(man kann* $I = \mathbb{R}$ *wählen), und das sog.* charakteristische Polynom *von* (1.52) $p(X) := X^n + \alpha_{n-1}X^{n-1} + \ldots + \alpha_1 X + \alpha_0$ *zerfalle über* $\mathbb{K}$ *in Linearfaktoren:*

$$p(X) = \prod_{j=1}^{k} \left(X - \lambda_j\right)^{m_j},$$

($\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ *verschieden,* $m_1, \ldots, m_k \in \mathbb{N}$, $m_1 + \ldots + m_k = n$*). Dann bilden für jedes* $x_0 \in I$ *die Funktionen*

$$y_{j,\nu}(x) := (x - x_0)^{\nu} e^{\lambda_j (x - x_0)} \qquad \left(j = 1, \ldots, k,\ \nu = 0, \ldots, m_j - 1\right) \tag{1.54}$$

ein Hauptsystem des Lösungsraums V *von* (1.52). *Insbesondere ist damit für* $\mathbb{K} = \mathbb{C}$ *eine Basis des Lösungsraums von* (1.52) *bekannt.*

Beweis Wir betrachten den Isomorphismus

$$\varphi\colon V \to \mathscr{L} := \left\{z\colon I \to \mathbb{K}^n\colon z' = Az\right\}, \qquad y \mapsto z = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

aus Konstruktion 1.5.3 mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{n-1} \end{pmatrix}$$

Mit Argumenten der Linearen Algebra sieht man nun leicht, dass

$$p_A(X) = (-1)^n \left(X^n + \alpha_{n-1}X^{n-1} + \ldots + \alpha_1 X + \alpha_0\right) = (-1)^n p(X);$$

daher der Name *charakteristisches Polynom* für p. Nach Voraussetzung zerfällt p_A, und wir können die Basiskonstruktion aus Satz 1.4.11 für die Differentialgleichung $z' = Az$ anwenden. Dazu brauchen wir das Minimalpolynom m_A von A: Wiederum mit Methoden der

Linearen Algebra finden wir: $p_A = (-1)^n\, m_A$ für obiges A, also $m_A = p$. Wir wollen aber diese Information gar nicht benutzen, da sie sich bei unseren Argumenten automatisch ergibt. p_A zerfällt, also auch m_A, etwa

$$m_a(X) = \prod_{j=1}^{k} \left(X - \lambda_j\right)^{\mu_j}.$$

Satz 1.4.11 besagt: $\mathscr{L}$ hat eine Basis von Funktionen des Typs

$$z_c(x) = \left(\sum_{\nu=0}^{\mu_j-1} \frac{(x-x_0)^\nu}{\nu!} \left(A - \lambda_j E_n\right)^\nu c\right) e^{\lambda_j (x-x_0)},$$

wobei $c \in B_j$ eine Basis von U_j durchläuft für $j = 1,\dots,k$. Die Elemente von V sind beliebige Linearkombinationen der ersten Koordinatenfunktionen solcher z_c. Also ist jedes $y \in V$ eine Linearkombination des Typs

$$y = \sum_{j=1}^{k} \sum_{\nu_0}^{\mu_j-1} \alpha_{j\nu} y_{j\nu} \tag{1.55}$$

mit den $y_{j\nu}$ aus (1.54). Wir setzen $q := \sum_{j=1}^{k} \mu_j \leqslant n$ und

$$W := \left\{ a = \begin{pmatrix} \alpha_{1,0} \\ \vdots \\ \alpha_{1,\mu_1-1} \\ \alpha_{2,0} \\ \vdots \\ \alpha_{2,\mu_2-1} \\ \vdots \\ \alpha_{k,0} \\ \vdots \\ \alpha_{k,\mu_k-1} \end{pmatrix} \in \mathbb{K}^q : \ \sum_{j=1}^{k} \sum_{\nu=0}^{\mu_j-1} \alpha_{j\nu} y_{j\nu} \in V \right\}$$

Dann ist $W \subset \mathbb{K}^q$ ein Untervektorraum, und

$$\psi\colon W \to V, \qquad a \mapsto \sum_{j=1}^{k} \sum_{\nu=0}^{\mu_j-1} \alpha_{j\nu} y_{j\nu}$$

ist eine surjektive lineare Abbildung, denn jedes $y \in V$ hat ja eine Darstellung (1.55). Wegen $\dim_{\mathbb{K}} V = n$ ist $\dim_{\mathbb{K}} \psi(W) \geqslant n$, denn $\dim_{\mathbb{K}} V + \dim_{\mathbb{K}} \ker\psi = \dim_{\mathbb{K}} W$. Aus $W \subset \mathbb{K}^q$ folgt $q \geqslant \dim W \geqslant \dim\psi(W) \geqslant n$, also $q \geqslant n$. Wegen $\mu_j \leqslant m_j$ ist aber ohnehin $q = \sum_{j=1}^{k} \mu_j \leqslant \sum_{j=1}^{k} m_j = n$, d. h. $q = n$, $\mu_j = m_j$ für $j = 1,\dots,k$, $p_A = (-1)^n\, m_A$, $\dim W = q = n$, also $W = \mathbb{K}^n$. Nun ist $\psi\colon \mathbb{K}^n \to V$ surjektiv, $\dim \mathbb{K}^n = \dim V = n$, also ist ψ ein Isomorphismus, der eine beliebige Basis von $\mathbb{K}^n$ in eine Basis von V überführt. Wählt man in $\mathbb{K}^n$ die Standardbasis $e_1,\dots,e_n$, so sind $\psi(e_1),\dots,\psi(e_n)$ gerade die $y_{j\nu}$ für $j = 1,\dots,k$ und $\nu = 0,\dots,m_j - 1$. $\square$

1.5.6 Satz (Basissatz, reelle Version) *Gegeben sei die Differentialgleichung* (1.52) *mit Koeffizienten* $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{R}$*, und das charakteristische Polynom*

$$p(X) = X^n + \alpha_{n-1} X^{n-1} + \ldots + \alpha_1 X + \alpha_0 \in \mathbb{R}[X]$$

von (1.52) *habe über* $\mathbb{C}$ *die Zerlegung*

$$p(X) = \prod_{j=1}^{k} \left(X - \lambda_j\right)^{m_j}$$

mit verschiedenen $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$*, wobei* $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$*,* $\lambda_j = \alpha_j + i\beta_j$ *und* $\lambda_{j+s} = \alpha_j - i\beta_j$ *mit* $\alpha_j \in \mathbb{R}$*,* $\beta_j > 0$ *für* $j = r+1, \ldots, r+s$*. Dann hat der Lösungsraum* $V_\mathbb{R}$ *von* (1.52) *über* $\mathbb{R}$ *für jedes* $x_0 \in I$ *das Hauptsystem*

$$\begin{aligned} &(x - x_0)^\nu\, e^{\lambda_j (x - x_0)} && \left(j = 1, \ldots, r,\ \nu = 0, \ldots, m_j - 1\right) \\ &(x - x_0)^\nu\, e^{\alpha_j (x - x_0)} \cos \beta_j (x - x_0) && \\ &(x - x_0)^\nu\, e^{\alpha_j (x - x_0)} \sin \beta_j (x - x_0) && \left(j = r+1, \ldots, r+s,\ \nu = 0, \ldots, m_j - 1\right) \end{aligned} \tag{1.56}$$

Beweis Nach Satz 1.5.5 sind die Funktionen (1.56) wegen $e^{\alpha + i\beta} = e^\alpha \left(\cos\beta + i \sin\beta\right)$ eine Basis von $V_\mathbb{C}$ über $\mathbb{C}$. Sei $y \in V_\mathbb{R}$. Dann ist y eine Linearkombination der Funktionen (1.56) mit Koeffizienten aus $\mathbb{C}$. Übergang zum Realteil: y ist Linearkombination der Funktionen (1.56) mit reellen Koeffizienten. $\Rightarrow$ (1.56) ist ein Erzeugendensystem von $V_\mathbb{R}$ und hat genau $n = \dim_\mathbb{R} V_\mathbb{R}$ Elemente. Also ist (1.56) eine Basis von $V_\mathbb{R}$. ☐

1.5.7 Beispiel Die Differentialgleichung $y'' + \alpha y = 0$ haben wir schon in Beispiel 0.2.6 (im Fall $\alpha = \omega^2 > 0$) und in Aufgabe 5 (für den Fall $\alpha = -\lambda^2 < 0$) betrachtet. Jetzt geht das so: Das charakteristische Polynom $p(X) := X^2 + \alpha$ zerfällt über $\mathbb{R}$ genau dann, wenn $\alpha \leqslant 0$.

(i) $\alpha = -\lambda^2 < 0$. Dann ist $p(X) = (X - \lambda)(X + \lambda)$. Dann liefert Satz 1.5.5 mit $\mathbb{K} = \mathbb{R}$: Es gibt ein Hauptsystem, und zwar $e^{\lambda x}$, $e^{-\lambda x}$ (oder $\cosh \lambda x$, $\sinh \lambda x$).

(ii) $\alpha = 0 \Rightarrow p(X) = X^2$. In diesem Fall liefert Satz 1.5.5 mit $\mathbb{K} = \mathbb{R}$ das Hauptsystem $1, x$

(iii) $\alpha = \omega^2 > 0 \Rightarrow p(X) = (X - i\omega)(X + i\omega)$ Jetzt liefert Satz 1.5.6 für $\mathbb{K} = \mathbb{R}$ das Hauptsystem über $\mathbb{R}$: $\cos \omega x$, $\sin \omega x$. Alternativ liefert Satz 1.5.5 für $\mathbb{K} = \mathbb{C}$ das Hauptsystem über $\mathbb{C}$: $e^{i\omega x}$, $e^{-i\omega x}$.

1.5.8 Beispiel (Homogene lineare Differentialgleichung 2. Ordnung) (mit konstanten Koeffizienten über $\mathbb{R}$): $y'' + 2ay' + by = 0$ $(a, b \in \mathbb{R})$. Charakteristisches Polynom:

$$p(X) = X^2 + 2aX + b = (X + a)^2 + b - a^2.$$

3 Fälle können auftreten:

(i) $a^2 - b^2 > 0 \Rightarrow p(X) = (X - \lambda_1)(X - \lambda_2)$ mit $\lambda_1 = -a + \sqrt{a^2 - b}$, $\lambda_2 = -a - \sqrt{a^2 - b} \in \mathbb{R}$, Hauptsystem: $e^{\lambda_1 x}$, $e^{\lambda_2 x}$.

(ii) $a^2 - b = 0 \Rightarrow p(X) = (X+a)^2$, $\lambda_1 = -a$, $m_1 = 2$, Hauptsystem: $e^{\lambda_1 x}$, $xe^{\lambda_1 x}$.

(iii) $a^2 - b < 0 \Rightarrow p(X) = (X-\lambda_1)(X-\lambda_2)$ mit $\lambda_1 = -a + i\sqrt{b-a^2}$, $\operatorname{Im}\lambda_1 > 0$, $\lambda_2 = \overline{\lambda_1}$, im Sinne von Satz 1.5.6 setze man $\alpha_1 := -a$, $\beta_1 := \sqrt{b-a^2}$, Hauptsystem: $e^{\alpha_1 x}\cos\beta_1 x$, $e^{\alpha_1 x}\sin\beta_1 x$. Man kann hier auch überall $x - x_0$ an Stelle von x schreiben, wenn das bequemer ist.

1.5.9 Beispiel (Gedämpfte Schwingung) Wir befinden uns in der Situation von Beispiel 0.2.6 (siehe Abb. 1.10): Rücktreibende Federkraft $K_{\text{Rück}} = -Dx$, mit der Federkonstante $D > 0$, Reibungskraft: $K_{\text{Reib}} = -2R\dot{x}$ (für geringe Geschwindigkeiten ist diese also proportional zur Geschwindigkeit), $R \geqslant 0$; Masse: $m\ddot{x} = -Dx - 2R\dot{x}$, also:

$$\ddot{x} + \frac{2R}{m}\dot{x} + \frac{D}{m}x = 0.$$

Beispiel 1.5.8 liefert die Hauptsysteme $a = \frac{R}{m}$, $b = \frac{D}{m}$.

(i) $a^2 > b$ (d. h. die Reibung ist groß im Verhältnis zur rücktreibenden Kraft: $R^2 > mD$):

$$\lambda_1 = -a + \sqrt{a^2 - b} = \frac{1}{m}\left(-R + \sqrt{R^2 - mD}\right) < 0,$$
$$\lambda_2 = \frac{1}{m}\left(-R - \sqrt{R^2 - mD}\right) < 0$$

Hauptsystem $e^{\lambda_1(t-t_0)}$, $e^{\lambda_2(t-t_0)}$. Anfangsbedingung $x(t_0) = x_0$, $\dot{x}_0(t_0) = v_0$:

$$x(t) = \alpha e^{\lambda_1(t-t_0)} + \beta e^{\lambda_2(t-t_0)},$$

wobei $x_0 = \alpha + \beta$, $v_0 = \lambda_1\alpha + \lambda_2\beta$. Also ist $\alpha = \frac{v_0 - \lambda_2 x_0}{\lambda_1 - \lambda_2}$, $\beta = \frac{v_0 - \lambda_1 x_0}{\lambda_2 - \lambda_1}$.

$$\Rightarrow x(t) = \frac{v_0 - \lambda_1 x_0}{\lambda_1 - \lambda_2} e^{\lambda_1(t-t_0)} - \frac{v_0 - \lambda_1 x_0}{\lambda_1 - \lambda_2} e^{\lambda_2(t-t_0)}$$

ist die Lösung der Differentialgleichung mit $x(t_0) = x_0$, $\dot{x}(t_0) = v_0$.
In unserem Beispiel sind $R > 0$, $-\frac{R \pm \sqrt{R^2 - mD}}{m} < 0$, d. h. der Massenpunkt bewegt sich exponentiell in die Ruheposition zurück: Situation bei gut eingestellten Schwingtüren oder bei einer Schwingung in zähem Medium.

(ii) $a^2 = b$, d. h. $R^2 = mD$. $\Rightarrow \lambda_1 = -\frac{R}{m}$, $m_1 = 2$, Hauptsystem: $e^{\lambda_1(t-t_0)}$, $(t-t_0)e^{\lambda_1(t-t_0)}$. Anfangsbedingung: $x(t_0) = x_0$, $\dot{x}(t_0) = v_0$:

$$x(t) = \alpha e^{\lambda_1(t-t_0)} + \beta(t-t_0)e^{\lambda_1(t-t_0)},$$

$x_0 = \alpha$, $v_0 = \lambda_1\alpha + \beta$, also $\alpha = x_0$, $\beta = v_0 - \lambda - 1x_0$,

$$x(t) = x_0 e^{\lambda_1(t-t_0)} + (v_0 - \lambda_1 x_0)(t-t_0)e^{\lambda_1(t-t_0)}.$$

In jedem Fall kehrt der Massenpunkt asymptotisch in die Ausgangslage zurück, ohne zu schwingen. Allerdings kann es – je nach Wahl von x_0 und v_0 – vorkommen, dass die Rückkehr in die Ausgangslage „einseitig“ stattfindet, oder dass der Massenpunkt zunächst über 0 hinaus zurückschnellt und dann in die Ausgangslage zurückkriecht.

(iii) $a^2 < b$, d. h. $R^2 < mD$ (Reibung klein im Verhältnis zur rücktreibenden Kraft). Dann ist eine echte gedämpfte Schwingung zu erwarten, da sich im Grenzfall $R = 0$ bekanntlich eine ungedämpfte Schwingung einstellt:

$$\begin{aligned} \lambda_1 &= -a + i\sqrt{b-a^2} & \alpha_1 &= -\frac{R}{m}, \\ &= \alpha_1 + i\beta_1 & \beta_1 &= \frac{1}{m}\sqrt{mD-R^2} > 0, \end{aligned}$$

Hauptsystem:

$$e^{\alpha_1(t-t_0)}\cos\beta_1(t-t_0), \qquad e^{\alpha_1(t-t_0)}\sin\beta_1(t-t_0).$$

Anfangsbedingung: $x(t_0) = x_0$, $\dot{x}(t_0) = v_0$,

$$x(t) = \alpha e^{\alpha_1(t-t_0)}\cos\beta_1(t-t_0) + \beta e^{\alpha_1(t-t_0)}\sin\beta_1(t-t_0),$$

$x_0 = \alpha$, $v_0 = \alpha\alpha_1 + \beta\beta_1$, also

$$\alpha = x_0, \qquad \beta = \frac{v_0 - \alpha_1 x_0}{\beta_1} = \frac{v_0 + \frac{R}{m}x_0}{\frac{1}{m}\sqrt{mD-R^2}},$$

und damit

$$x(t) = x_0 e^{-\frac{R}{m}(t-t_0)}\cos\frac{\sqrt{mD-R^2}}{m}(t-t_0) + \frac{mv_0 + Rx_0}{\sqrt{mD-R^2}}e^{-\frac{R}{m}(t-t_0)}\sin\frac{\sqrt{mD-R^2}}{m}(t-t_0).$$

Exponentiell gedämpfte Schwingung; $R = 0$: $\frac{D}{m} = \omega^2$, $\omega > 0$:

$$x = x_0\cos\omega(t-t_0) + \frac{v_0}{\omega}\sin\omega(t-t_0)$$

ist die Lösung aus Beispiel 0.2.6. Man kann auch hier die maximale Amplitude und die Phasenverschiebung berechnen wie in Beispiel 0.2.6 und für x eine Darstellung des Typs

$$x(t) = Ae^{-\frac{R}{m}(t-t_0)}\cos\big(\omega(t-t_0) - \varphi\big)$$

herleiten.

Nun betrachten wir inhomogene lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten:

1.5.10 Satz *Vorgelegt sei die inhomogene lineare Differentialgleichung (über* $\mathbb{K}$*)* (1.51). *mit stetigem* $g\colon I \to \mathbb{K}$. *Dann existiert zu beliebig vorgegebenen* $x_0 \in I$, $y_0, \dots, y_{n-1} \in \mathbb{K}$ *genau eine Lösung* $y\colon I \to \mathbb{K}$ *(n-mal stetig differenzierbar) von* (1.51) *mit*

$$y(x_0) = y_0, \dots, y^{(n-1)}(x_0) = y_{n-1}. \tag{1.57}$$

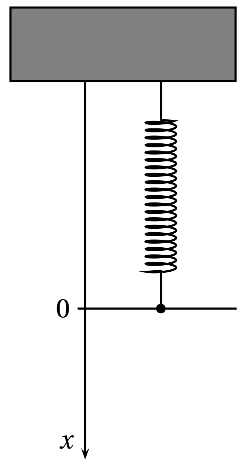

Abbildung 1.10: Gedämpfte Schwingung einer Feder aus Beispiel 1.5.9

Beweis Sei y eine n-mal differenzierbare Lösung von (1.51). Wir setzen

$$z\colon I \to \mathbb{K}^n, \qquad z := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix},$$

und haben:

$$z' = \begin{pmatrix} y' \\ \vdots \\ y^{(n-1)} \\ y^{(n)} \end{pmatrix} = \begin{pmatrix} y' \\ \vdots \\ y^{(n-1)} \\ -\alpha_0 y - \ldots - \alpha_{n-1} y^{(n-1)} + g \end{pmatrix} = Az + f$$

mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{n-1} \end{pmatrix}, \qquad f := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ g \end{pmatrix} : I \to \mathbb{K}^n \text{ stetig}$$

Umgekehrt: Ist $z\colon I \to \mathbb{K}^n$ differenzierbar mit $z' = Az + f$, so ist $y := z_1$ eine n-mal differenzierbare Lösung von (1.51). Nun hat nach Satz 1.4.13 die Differentialgleichung $z' = Az + f$ für jedes $x_0 \in I$, $z_0 \in \mathbb{K}^n$ genau eine Lösung z mit $z(x_0) = z_0$, nämlich

$$z(x) := e^{A(x-x_0)} z_0 + \int_{x_0}^{x} e^{A(x-t)} f(t)\, dt.$$

Wählen wir $z_0 := \left(y_0, \ldots, y_{n-1}\right)^\top$, A und f wie oben, so folgt die Behauptung, und wir haben die Lösung sogar explizit angegeben. $\square$

Effektive Bestimmung der Lösungen: Gegeben seien die inhomogenen Differentialgleichungen (alles über $\mathbb{K}$) (1.51) mit $g\colon I \to \mathbb{K}$ stetig und die zugehörige homogene Differentialgleichung (1.52). Sei $u_1, \ldots, u_n$ eine Basis des Lösungsraums V von (1.52). Dann brauchen wir nach Folgerung 1.5.2 b) nur eine sog. partikuläre Lösung u_0 von (1.51), und erhalten in der Gestalt

$$u = u_0 + \lambda_1 u_1 + \ldots + \lambda_n u_n \qquad (\lambda_1, \ldots, \lambda_n \in \mathbb{K})$$

alle Lösungen von (1.51) über $\mathbb{K}$. Da das Anfangswertproblem für (1.51) für jeden Anfangswert (1.57) nach Satz 1.5.10 eindeutig lösbar ist, ist damit bei bekannten Funktionen $u_0, u_1, \ldots, u_n$ die Lösung des Anfangswertproblems auf die Lösung eines linearen Gleichungssystems reduziert. Eine Basis $u_1, \ldots, u_n$ ist bekannt aus Satz 1.5.5 und Satz 1.5.6. Bleibt das Problem: Wie findet man ein u_0? Die Antwort findet sich in Konstruktion 1.5.12. Vorab ein

1.5.11 Lemma Ist $u_1, \ldots, u_n\colon I \to \mathbb{K}$ eine Basis des Lösungsraums V von (1.52) über $\mathbb{K}$, so gilt für alle $x_0 \in I$:

$$W(x) := \begin{vmatrix} u_1(x) & \ldots & u_n(x) \\ u_1'(x) & \ldots & u_n'(x) \\ \vdots & & \vdots \\ u_1^{(n-1)}(x) & \ldots & u_n^{(n-1)}(x) \end{vmatrix} = W(x_0)\, e^{\alpha_{n-1}(x-x_0)} \neq 0 \qquad (x \in I)$$

$W(x)$ nennt man die *Wronskische Determinante*, benannt nach Graf Hoëné Wronski, 1778–1853.

Beweis W ist differenzierbar, und nach der Leibniz-Formel für die Determinante

$$\det M = \sum_{\sigma \in S_n} \prod_{i=1}^{n} M_{i,\sigma(i)} \operatorname{sgn} \sigma$$

gilt:

$$\begin{aligned} W'(x) &= \underbrace{\begin{vmatrix} u_1' & \ldots & u_n' \\ u_1' & \ldots & u_n' \\ \vdots & & \vdots \\ u_1^{(n-1)} & \ldots & u_n^{(n-1)} \end{vmatrix} + \begin{vmatrix} u_1 & \ldots & u_n \\ u_1'' & \ldots & u_n'' \\ \vdots & & \vdots \\ u_1^{(n-1)} & \ldots & u_n^{(n-1)} \end{vmatrix} + \ldots +}_{\text{alle gleich 0!}} \begin{vmatrix} u_1 & \ldots & u_n \\ \vdots & & \vdots \\ u_1^{(n-2)} & \ldots & u_n^{(n-2)} \\ u_1^{(n)} & \ldots & u_n^{(n)} \end{vmatrix} \\ &= \begin{vmatrix} u_1 & \ldots & u_n \\ u_1' & \ldots & u_n' \\ \vdots & & \vdots \\ -\alpha_{n-1} u_1^{(n-1)} - \ldots - \alpha_0 u_1 & \ldots & -\alpha_{n-1} u_n^{(n-1)} - \ldots - \alpha_0 u_n \end{vmatrix} \\ &= -\alpha_{n-1} W. \end{aligned}$$

$\Rightarrow W(x) = W(x_0)\, e^{-\alpha_{n-1}(x-x_0)}$ für alle $x \in I$. Da das Anfangswertproblem der Differentialgleichung (1.52) mit den Anfangsbedingungen $y(x_0) = y_0,\ y'(x_0) = y_1,\ \dots,\ y^{(n-1)}(x_0) = y_{n-1}$ für beliebiges $\left(y_0, \dots, y_{n-1}\right)^\top \in \mathbb{K}^n$ lösbar ist, sind die Vektoren

$$\begin{pmatrix} u_1(x_0) \\ u_1'(x_0) \\ \vdots \\ u_1^{(n-1)}(x_0) \end{pmatrix}, \dots, \begin{pmatrix} u_n(x_0) \\ u_n'(x_0) \\ \vdots \\ u_n^{(n-1)}(x_0) \end{pmatrix}$$

linear unabhängig, also ist $W(x_0) \neq 0$. Das gilt für beliebiges $x_0 \in I$. Da wir nur $W(x) \neq 0$ für alle x brauchen, ist für den engeren Zweck der folgenden effektiven Lösungsbestimmung die genaue Bestimmung der Determinante $W(x)$ hier eigentlich nicht nötig, kommt aber von selbst mit heraus. □

1.5.12 Konstruktion (Effektive Bestimmung einer partikulären Lösung) Gegeben seien Konstanten $\alpha_0, \dots, \alpha_{n-1} \in \mathbb{K}$, $g\colon I \to \mathbb{K}$ stetig, $p(X) := X^n + \alpha_{n-1}X^{n-1} + \ldots + \alpha_0$, $D := \frac{d}{dx}$; dann ist die inhomogene Differentialgleichung (1.51) äquivalent zu $p(D)\, y = g$, die zugehörige homogene Differentialgleichung (1.52) zu $p(D)\, y = 0$. Sei $u_1, \dots, u_n$ ein Hauptsystem von (1.52). Zur Bestimmung einer partikulären Lösung u von (1.51) machen wir den Ansatz

$$u = \sum_{j=1}^{n} c_j u_j$$

mit differenzierbaren Funktionen $c_j\colon I \to \mathbb{K}$ für $j = 1, \dots, n$ (Methode der Variation der Konstanten). Dann gilt:

$$n-1 \text{ Gleichungen} \begin{cases} u' = \sum_{j=1}^{n} c_j u_j', & \text{falls } \sum_{j=1}^{n} c_j' u_j = 0, \\ u'' = \sum_{j=1}^{n} c_j u_j'', & \text{falls } \sum_{j=1}^{n} c_j' u_j' = 0, \\ \vdots & \\ u^{(n-1)} = \sum_{j=1}^{n} c_j u_j', & \text{falls } \sum_{j=1}^{n} c_j' u_j^{(n-2)} = 0, \end{cases}$$

und es folgt weiter:

$$u^{(n)} = \sum_{j=1}^{n} c_j u_j^{(n)} + \sum_{j=1}^{n} c_j' u_j^{(n-1)},$$

also nach Multiplikation mit $\alpha_0, \dots, \alpha_{n-1}$ und Summation

$$p(D)\, u = \sum_{j=1}^{n} c_j \underbrace{p(D)\, u_j}_{=0!} + \sum_{j=1}^{n} c_j' u_j^{(n-1)},$$

d. h. $p(D)\, u = g$, falls $\sum_{j=1}^{n} c_j' u_j^{(n-1)} = g$.
Ergebnis: Sind die differenzierbaren Funktionen $c_1, \dots, c_n\colon I \to \mathbb{K}$ so gewählt, dass

$$\sum_{j=1}^{n} c_j' u_j^{(i-1)} = 0 \qquad (i = 1, \dots, n), \tag{1.58}$$

so ist $u = \sum_{j=1}^{n} c_j u_j$ eine partikuläre Lösung der Differentialgleichung (1.51). Wir zeigen weiter: (1.58) ist lösbar! Wir fassen (1.58) auf als lineares Gleichungssystem für $c_1', \ldots, c_n'$ (und zwar für jedes feste $x \in I$): Nach Lemma 1.5.11 ist die Determinante

$$W(x) := \begin{vmatrix} u_1(x) & \ldots & u_n(x) \\ u_1'(x) & \ldots & u_n'(x) \\ \vdots & & \vdots \\ u_1^{(n-1)}(x) & \ldots & u_n^{(n-1)}(x) \end{vmatrix} = W(x_0)\, e^{\alpha_{n-1}(x-x_0)} \neq 0 \qquad (x \in I).$$

Dann liefert die Cramersche Regel: Es gibt eindeutig bestimmte stetige Funktionen $c_1', \ldots, c_n' \colon I \to \mathbb{K}$, so dass (1.58) gilt. Zu diesen (stetigen!) $c_1', \ldots, c_n'$ wählen wir Stammfunktionen $c_1, \ldots, c_n \colon I \to \mathbb{K}$ und bilden $u := \sum_{j=1}^{n} c_j u_j$. Dann ist u eine partikuläre Lösung von (1.51). □

Bemerkung Eine Abänderung von $c_1, \ldots, c_n$ um irgendwelche Konstanten muss nach dieser Überlegung erlaubt sein und ist es offenbar auch, denn sie bedeutet eine Abänderung von u um eine Lösung des homogenen Systems.

Ergebnis: Sind $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{K}$ und $g \colon I \to \mathbb{K}$ stetig, $u_1, \ldots, u_n$ eine Basis des Lösungsraums der Differentialgleichung (1.52), so gibt es eine partikuläre Lösung u der inhomogenen Differentialgleichung (1.51) von der Form $u = \sum_{j=1}^{n} c_j u_j$ mit einfach stetig differenzierbaren Funktionen $c_1, \ldots, c_n \colon I \to \mathbb{K}$. Ein solches System von Funktionen $c_1, \ldots, c_n$ erhält man durch Auflösen des linearen Gleichungssystems (1.58) nach $c_1', \ldots, c_n'$ und anschließende Wahl von Stammfunktionen $c_1, \ldots, c_n$ zu $c_1', \ldots, c_n'$.

1.5.13 Beispiel (Lineare Differentialgleichungen zweiter Ordnung) Man vgl. hierzu auch Beispiel 1.5.8.

$$y'' + ay' + by = g(t),$$

$g \colon I \to \mathbb{R}$, $a, b \in \mathbb{R}$. Wir betrachten z. B. den Fall $a^2 - b > 0$ und haben nach Beispiel 1.5.8 das Hauptsystem $u_1 = e^{\lambda_1 t}$, $u_2 = e^{\lambda_2 t}$ mit $\lambda_1 = -a + \sqrt{a^2 - b}$ und $\lambda_2 = -a - \sqrt{a^2 - b}$. Wir setzen wir oben unsere partikuläre Lösung an in der Form $u_0 = c_1 u_1 + c_2 u_2$ mit $c_1, c_2 \colon I \to \mathbb{R}$ und haben als Gleichungssystem (1.58):

$$\begin{aligned} c_1' e^{\lambda_1 t} + c_2' e^{\lambda_2 t} &= 0, \\ \lambda_1 c_1' e^{\lambda_1 t} + \lambda_2 c_2' e^{\lambda_2 t} &= g(t), \end{aligned}$$

und das wird gelöst von

$$\begin{aligned} c_1' &= \frac{1}{\lambda_1 - \lambda_2} e^{-\lambda_1 t} g(t), \\ c_2' &= -\frac{1}{\lambda_1 - \lambda_2} e^{-\lambda_2 t} g(t) \end{aligned}$$

Zu diesen c_1', c_2' bestimme man (irgendwelche!) Stammfunktionen $c_1, c_2 \colon I \to \mathbb{R}$. Diese existieren, da g stetig ist. Dann ist $u_0 = c_1 u_1 + c_2 u_2$ eine partikuläre Lösung der inhomogenen Differentialgleichung, und in der Gestalt

$$u = (c_1 + \alpha_1) u_1 + (c_2 + \alpha_2) u_2 \qquad (\alpha_1, \alpha_2 \in \mathbb{R})$$

erhalten wir alle Lösungen der inhomogenen Differentialgleichung.
D Das Beispiel klappt unverändert über $\mathbb{C}$; so lange $\lambda_1 \neq \lambda_2$, d. h. $a \neq b$ gilt.

Manchmal kann man viel einfacher eine spezielle Lösung finden, wenn die Funktion g auf der rechten Seite der inhomogenen Differentialgleichung von speziellem Typ ist. Folgender Satz gibt ein Beispiel:

1.5.14 Satz *Es seien $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{K}$, und gegeben sei die Differentialgleichung*

$$y^{(n)} + \alpha_{n-1} y^{(n-1)} + \ldots + \alpha_0 y = g(x)\, e^{\lambda x} \qquad (x \in \mathbb{R}) \tag{1.59}$$

mit einem Polynom $g \in \mathbb{K}[X]$ vom Grad m, $\lambda \in \mathbb{K}$. Ferner sei λ Nullstelle der Vielfachheit $k \geqslant 0$ des charakteristischen Polynoms

$$p(X) = X^n + \alpha_{n-1} X^{n-1} + \ldots + \alpha_0.$$

(Dabei heißt $k = 0$: $p(\lambda) \neq 0$.) Dann hat (1.59) *eine spezielle Lösung u der Form $u(x) = x^k \varphi(x)\, e^{\lambda x}$ mit $\varphi \in \mathbb{K}[X]$, mit $\operatorname{grad} \varphi \leqslant \operatorname{grad} g$.*

Beweis per Induktion nach $m := \operatorname{grad} g$. Wir zerlegen $p(X) = q(X)(X - \lambda)^k$, mit $q \in \mathbb{K}[X]$, $q(\lambda) \neq 0$.

$m = 0$: Mit $g(x) = a$ leistet $\varphi(x) := \frac{a}{k!q(\lambda)}$ das Verlangte, denn mit $D := \frac{d}{dx}$ gilt:

$$\begin{aligned} p(D)\, x^k \varphi(x)\, e^{\lambda x} &= q(D)(D - \lambda)^k \frac{a}{k!q(\lambda)} x^k e^{\lambda x} \\ &= q(D) \frac{a}{q(\lambda)} e^{\lambda x} = a e^{\lambda x} = g(x)\, e^{\lambda x}. \end{aligned}$$

$m - 1 \mapsto m$:

$$\begin{aligned} p(D)\left(x^{m+k} e^{\lambda x}\right) &= q(D)(D - \lambda)^k x^{m+k} e^{\lambda x} \\ &= (m + k)(m + k - 1) \cdot \ldots \cdot (m + 1)\, q(D)\, x^m e^{\lambda x} \\ &= h(x)\, e^{\lambda x} \end{aligned}$$

mit geeignetem $h \in \mathbb{K}[X]$ vom Grad m. Sei nun $\operatorname{grad} g = m$, h wie oben und die Behauptung richtig für einen Polynomgrad $\leqslant m - 1$. Dann existiert ein $\alpha \in \mathbb{K}$, so dass für $g_1 := g - \alpha h$ gilt: $\operatorname{grad} g_1 \leqslant m - 1$. Die Induktionsvoraussetzung liefert jetzt die Existenz eines $\varphi_1 \in \mathbb{K}[X]$ mit $\operatorname{grad} \varphi_1 \leqslant \operatorname{grad} g_1 \leqslant m - 1$ und

$$p(D)\, x^k \varphi_1(x)\, e^{\lambda x} = g_1(x)\, e^{\lambda x}.$$

Für $\varphi := \varphi_1 + \alpha x^m$ gilt nun:

$$\begin{aligned} p(D)\, x^k \varphi(x)\, e^{\lambda x} &= p(D)\, x^k \varphi_1(x)\, e^{\lambda x} + \alpha\, p(D)\, x^{m+k} e^{\lambda x} \\ &= \big(g_1(x) + \alpha h(x)\big) e^{\lambda x} = g(x)\, e^{\lambda x}. \end{aligned}$$ □

$$\lim_{x \to \infty} f(x) = 3 \tag{1.60}$$

2 Existenz-, Eindeutigkeitssatz und Abhängigkeitssätze für Systeme von Differentialgleichungen erster Ordnung

Inhaltsangabe

Vorbemerkung: Wir werden im folgenden Differentialgleichungen des Typs

$$y' = f\left(x, y\right) \tag{2.1}$$

betrachten, wobei $f\colon G \to \mathbb{R}^n$, $G \subset \mathbb{R}^{n+1}$; gesucht sind Lösungen $y\colon I \to \mathbb{R}^n$ mit Graph $y \subset G$, so dass (2.1) gilt. In der Regel interessiert man sich für Lösungen y, die der Anfangsbedingung $y\left(x_0\right) = y_0$ genügen.
Viele Differentialgleichungen haben a priori nicht die Gestalt (2.1), lassen sich aber auf die Form (2.1) bringen;

2.0.1 Beispiel Vorgelegt sei die Differentialgleichung

$$y^{(n)} = F\left(x, y, y', \dots, y^{(n-1)}\right) \tag{2.2}$$

mit $F\colon G \to \mathbb{R}$, $G \subset \mathbb{R}^{n+1}$ offen; gesucht ist eine n-mal differenzierbare Funktion $y\colon I \to \mathbb{R}$ mit $\left(x, y\left(x\right), \dots, y^{(n-1)}\left(x\right)\right) \in G$ für alle $x \in I$, so dass (2.2) gilt. Jeder solchen Lösung ordnen wir den Vektor

$$z := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix} : I \to \mathbb{R}^n$$

zu und haben: z ist differenzierbar mit

$$z' = \begin{pmatrix} y' \\ y'' \\ \vdots \\ y^{(n)} \end{pmatrix} = \begin{pmatrix} y' \\ \vdots \\ y^{(n-1)} \\ F\left(x, y, \ldots, y^{(n-1)}\right) \end{pmatrix} = f(x, z) \tag{2.3}$$

mit geeignetem $f\colon G \to \mathbb{R}^n$; wobei G der Definitionsbereich von F ist. Das ist eine Differentialgleichung vom Typ (2.1) für z. – Ist umgekehrt $z\colon I \to \mathbb{R}^n$ eine differenzierbare Lösung von (2.3), d. h. von

$$z' = \begin{pmatrix} z_2 \\ \vdots \\ z_n \\ F(x, z_1, \ldots, z_n) \end{pmatrix}, \qquad \left(z = (z_1, \ldots, z_n)^\top\right),$$

so gilt für $y := z_1$: y ist n-mal differenzierbar mit $y^{(i)} = z_i' = z_{i+1}$ für $i = 1, \ldots, n-1$ und $y^{(n)} = z_n' = F(x, z_1, \ldots, z_n) = F\left(x, y, \ldots, y^{(n-1)}\right)$. Auch die Anfangsbedingungen übertragen sich problemlos von y auf z und umgekehrt: Genügt y der Differentialgleichung (2.2) und der Anfangsbedingung $y(x_0) = y_0$, $y'(x_0) = y_1$, ..., $y^{(n-1)}(x_0) = y_{n-1}$, so genügt

$$z := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix} : I \to \mathbb{R}^n$$

der Anfangsbedingung $z(x_0) = \left(y_0, y_1, \ldots, y_{n-1}\right)^\top =: z_0$. Umgekehrt: Genügt z der Differentialgleichung (2.3) und der Anfangsbedingung $z(x_0) = z_0 \in \mathbb{R}^n$, so genügt $y := z_1$ der Differentialgleichung (2.2) und der Anfangsbedingung

$$\begin{pmatrix} y(x_0) \\ y'(x_0) \\ \vdots \\ y^{(n-1)}(x_0) \end{pmatrix} = z_0 =: \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix}.$$

Also: Die Zuordnung $y \mapsto z$ liefert unter Mitnahme der Anfangsbedingung eine Bijektion zwischen den Lösungen von (2.2) und (2.3). Da (2.1) formal noch etwas allgemeiner aussieht als (2.3), werden wir im folgenden wesentlich mit (2.1) arbeiten. Daher sind (2.2) und (2.3) im Sinne dieser Korrespondenz äquivalent. – Diese Korrespondenz lässt sich noch verallgemeinern, indem man System von Differentialgleichungen des Typs

$$y_j^{(n_j)} = f_j\left(x, y_1, \ldots, y_1^{(n_1-1)}, \ldots, y_m, \ldots, y_m^{(n_m-1)}\right) \qquad \left(j = 1, \ldots, m\right) \tag{2.4}$$

auf die Gestalt (2.1) bringt. Wir werden uns also auf Systeme des Typs (2.1) beschränken und für diese Existenz- und Eindeutigkeitssätze und Sätze über die stetige Abhängigkeit der Lösungen von den Anfangswerten herleiten.

2.1 Der Satz von Arzelà-Ascoli und der Existenzsatz von Peano

benannt nach den italienischen Mathematikern Cesare Arzelà (1547–1912), Giolio Ascoli (1843–1896) und Guiseppe Peano (1858–1932).
Die Differentialgleichung (2.1) hat unter der bloßen Annahme der Stetigkeit von f bereits eine Lösung mit beliebig vorgegebenem Anfangswert, aber die Lösung muss nicht eindeutig sein.
Zur Vorbereitung des Satzes von Peano einige Tatsachen über Funktionenfolgen:

2.1.1 Definition (Gleichgradige Stetigkeit) Sei $D \subset \mathbb{R}^n$ (oder in einem anderen metrischen Raum) und M eine Menge von Funktionen $f\colon D \to \mathbb{R}^m$ (oder einen anderen metrischen Raum).

a) M heißt *gleichgradig stetig* in $a \in D$, wenn zu jedem $\varepsilon > 0$ ein $\delta_a > 0$ existiert, so dass $\left\| f(x) - f(a) \right\|_2 < \varepsilon$ für alle $x \in D$ mit $\|x - a\| < \delta_a$ und alle $f \in M$.

b) M heißt *gleichgradig stetig* auf D, wenn M in jedem $a \in D$ gleichgradig stetig ist.

c) M heißt *gleichstetig* (oder auch gleichmäßig gleichgradig stetig), wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $\left\| f(x) - f\left(y\right) \right\|_2 < \varepsilon$ für alle $x, y \in D$ mit $\left\| x - y \right\| < \delta$ und alle $f \in M$.
Vorsicht: In der englischsprachigen Literatur sind diese Begriffe nicht ganz einheitlich: z. B. verwendet Stromberg den Begriff „equicontinuous“ im Sinne von b).

d) M heißt *punktweise beschränkt*, wenn für alle $x \in D$ die Menge $\left\{ f(x) : f \in M \right\} \subset \mathbb{R}^m$ beschränkt ist, d. h. wenn es zu jedem $x \in D$ ein $C_x \in \mathbb{R}$ gibt, so dass $\left\| f(x) \right\|_2 \leqslant C_x$ für alle $f \in M$, $x \in D$.

e) M heißt *gleichmäßig beschränkt*, wenn es ein $C > 0$ gibt, so dass $\left\| f(x) \right\|_2 \leqslant C$ für alle $f \in M$, $x \in D$.

2.1.2 Beispiele a) Für $a, b \in \mathbb{R}$ sei $\varphi_{a,b}\colon \mathbb{R} \to \mathbb{R}$, $\varphi_{a,b}(x) := ax + b$; ferner sei $C > 0$ und $M := \left\{ \varphi_{a,b}\colon a, b \in \mathbb{R}, |a| \leqslant C \right\}$. Dann ist M gleichstetig, aber nicht punktweise beschränkt. Für $C > 0$ ist $\tilde{M} := \left\{ \varphi_{a,b}\colon |a| \leqslant C, |b| \leqslant C \right\}$ punktweise beschränkt, aber nicht gleichmäßig beschränkt.

b) Ist $K \subset \mathbb{R}^n$ kompakt und $\left(f_k\right)_{k \geqslant 1}$ eine gleichmäßig konvergente Folge stetiger Funktionen $f_k\colon K \to \mathbb{R}^m$, $f_k \xrightarrow{k \to \infty} f$, so ist $M := \{f\} \cup \left\{ f_k\colon k \geqslant 1 \right\}$ gleichstetig und gleichmäßig beschränkt.

Beweis Sei $\varepsilon > 0$. Nach dem Cauchy-Kriterium gibt es zu $\varepsilon > 0$ ein N, so dass für alle $m, n \geqslant N$, $x \in K$ gilt: $\left\| f_m(x) - f_n(x) \right\| < \frac{\varepsilon}{3}$. K ist kompakt und f_N stetig, also ist f_N gleichmäßig stetig, es gibt also ein $\delta > 0$, so dass $\left\| f_N(x) - f_N\left(y\right) \right\| < \frac{\varepsilon}{3}$ für alle $x, y \in K$ mit $\left\| x - y \right\| < \delta$. $\Rightarrow$ Für alle $n \geqslant N$, $x, y \in K$, $\left\| x - y \right\| < \delta$ ist

$$\left\| f_n(x) - f_n\left(y\right) \right\| \leqslant \left\| f_n(x) - f_N(x) \right\| + \left\| f_N(x) - f_N\left(y\right) \right\| + \left\| f_N\left(y\right) - f_n\left(y\right) \right\| < \varepsilon,$$

und für $n \to \infty$ folgt auch $\|f(x) - f(y)\| < \varepsilon$. Da $f_1, \dots, f_{N-1}$ gleichmäßig stetig sind auf K, folgt nach eventueller geeigneter Verkleinerung von δ (zum Minimum der δ_n der gleichmäßigen Stetigkeit von $f_1, \dots, f_{N-1}$):

$$\forall\, x, y \in K,\ \|x - y\| < \delta,\ g \operatorname{im} M \qquad \|g(x) - g(y)\| < \varepsilon.$$

M ist also gleichstetig. Ferner: $f_n \xrightarrow{\text{gleichmäßig}} f$, also

$$\forall\, x \in K,\ n \geqslant n_0(1) \qquad \|f(x) - f_n(x)\| < 1.$$

f ist stetig, $f|_K$ also beschränkt: $\|f(x)\| \leqslant C$ für alle $x \in K$. Also ist $\|f_n(x)\| \leqslant C + 1$ für alle $n \geqslant n_0(1)$ und alle $x \in K$. Somit ist M gleichmäßig beschränkt. □

2.1.3 Satz *Es seien $X \subset \mathbb{R}^n$ und (f_k) eine gleichgradig stetige Folge von Funktionen $f_k\colon X \to \mathbb{R}^m$, die auf einer dichten Teilmenge $D \subset X$ konvergiere. Dann gibt es eine stetige Funktion $f\colon X \to \mathbb{R}^m$, so dass $(f_k)_{k\geqslant 1}$ auf jedem Kompaktum $K \subset X$ gleichmäßig gegen $f|_K$ konvergiert. Für X lässt sich auch ein beliebiger topologischer Raum, für $\mathbb{R}^m$ ein beliebiger vollständiger metrischer Raum wählen, siehe Stromberg (1996), S. 165.*

Beweis Seien $x \in X$, $\varepsilon > 0$. $\Rightarrow \exists \delta_x > 0$, so dass

$$\|f_k(t) - f_k(x)\| < \frac{\varepsilon}{3} \tag{2.5}$$

für alle $k \geqslant 1$ und $t \in X$, $\|t - x\| < \delta_x$. $\overline{D} \supset X \Rightarrow \exists\, t \in D$, $\|t - x\| < \delta_x$, und da $(f_k(t))_{k\geqslant 1}$ konvergiert, folgt: Es gibt ein $N \in \mathbb{N}$ mit

$$\|f_k(t) - f_\ell(t)\| < \frac{\varepsilon}{3} \tag{2.6}$$

für alle $k, \ell \geqslant N$. (2.5) und (2.6) liefern zusammen:

$$\|f_k(x) - f_\ell(x)\| \leqslant \|f_k(x) - f_k(t)\| + \|f_k(t) - f_\ell(t)\| + \|f_\ell(t) - f_\ell(x)\| < \varepsilon, \quad \text{falls } k, \ell \geqslant N.$$

$\Rightarrow (f_k(x))_{k\geqslant 1}$ ist eine Cauchyfolge in $\mathbb{R}^m$, also existiert ein $f\colon X \to \mathbb{R}^m$ mit $f_k \xrightarrow{k\to\infty} f$ (punktweise). Wir zeigen: f ist stetig: Sei $x \in X$, $\varepsilon > 0$, δ_x wie oben, $t \in X$, $\|t - x\| < \delta_x$ und ℓ so groß, dass

$$\|f(t) - f_\ell(t)\| < \frac{\varepsilon}{3} \tag{2.7}$$

$$\|f(x) - f_\ell(x)\| < \frac{\varepsilon}{3} \tag{2.8}$$

$$\overset{(2.5),(2.6),(2.8)}{\Longrightarrow} \|f(t) - f(x)\| \leqslant \|f(t) - f_\ell(t)\| + \|f_\ell(t) - f_\ell(x)\| + \|f_\ell(x) - f(x)\| < \varepsilon. \tag{2.9}$$

$\Rightarrow f$ ist stetig in x. – Sei $K \subset X$ kompakt. Wir zeigen: $f_k|_K \to f|_K$ gleichmäßig. Dazu sei $\varepsilon > 0$. Zu $x \in K$ wählen wir $\delta_x > 0$ gemäß (2.5), und dann gilt mit demselben δ_x auch (2.9) (s.o.). K kompakt liefert: $\exists\, x_1, \dots, x_p \in k$, so dass

$$K \subset \bigcup_{j=1}^{p} \underbrace{K_{\delta_{x_j}}(x_j)}_{=:U_j}.$$

Sei nun N so groß, dass

$$\left\| f\left(x_j\right) - f_k\left(x_j\right) \right\| < \frac{\varepsilon}{2} \quad \text{für alle } j \geqslant N,\ j = 1, \ldots, p. \tag{2.10}$$

Sei $t \in K \Rightarrow \exists\, j$, so dass $t \in U_j$ und damit für alle $k \geqslant N$:

$$\left\| f(t) - f_k(t) \right\| \leqslant \underbrace{\left\| f(t) - f\left(x_j\right) \right\|}_{<\varepsilon \text{ nach (2.9)}} + \underbrace{\left\| f\left(x_j\right) - f_k\left(x_j\right) \right\|}_{<\frac{\varepsilon}{2} \text{ nach (2.10)}} + \underbrace{\left\| f_k\left(x_j\right) - f_k(t) \right\|}_{<\frac{\varepsilon}{3} \text{ nach (2.5)}} < 2\varepsilon$$

$\Rightarrow f_k|_K \xrightarrow{\text{gleichmäßig}} f|_K$. □

2.1.4 Lemma Es seien $D \subset \mathbb{R}^n$ abzählbar und $\left(f_k\right)_{k \geqslant 1}$ eine punktweise beschränkte Folge von Funktionen $f_k \colon D \to \mathbb{R}^m$. Dann hat $\left(f_k\right)$ eine punktweise konvergente Teilfolge $\left(f_{k_j}\right)_{j \geqslant 1}$. Dies ist eine Verallgemeinerung des Satzes von Bolzano und Weierstraß.

Beweis Sei $D = \left\{x_j \colon j \geqslant 1\right\}$. Die Folge $\left(f_k\left(x_1\right)\right)$ ist beschränkt, hat also eine konvergente Teilfolge $\left(f_k^{(1)}\left(x_1\right)\right)_{k \geqslant 1}$. Die Folge $\left(f_k^{(1)}\left(x_2\right)\right)_{k \geqslant 1}$ ist beschränkt, hat also auch eine konvergente Teilfolge $\left(f_k^{(2)}\left(x_2\right)\right)_{k \geqslant 2}$ usw. Es gibt eine Teilfolge $\left(f_k^{(j)}\right)_{k \geqslant 1}$ von $\left(f_k\right)$, so dass $\left(f_k^{(j)}\left(x_\nu\right)\right)_{k \geqslant 1}$ konvergiert für $\nu = 1, \ldots, j$. Die Diagonalfolge $\left(f_k^{(k)}\right)_{k \geqslant 1}$ ist für $k \geqslant j$ eine Teilfolge von $\left(f_k^{(j)}\right)_{k \geqslant 1}$.
$\Rightarrow \left(f_k^{(k)}\left(x_j\right)\right)_{k \geqslant 1}$ konvergiert also für alle $j \geqslant 1$. □

2.1.5 Satz (Arzelà-Ascoli) *Es seien $K \subset \mathbb{R}^n$ kompakt und M eine Menge von stetigen Funktionen $f \colon K \to \mathbb{R}^m$. Dann sind äquivalent:*

a) M ist gleichgradig stetig und punktweise beschränkt.

b) Jede Folge $\left(f_k\right)_{k \geqslant 1}$ von Funktionen aus M hat eine gleichmäßig konvergente Teilfolge $\left(f_{k_j}\right)_{j \geqslant 1}$, die gegen eine stetige Funktion $f \colon K \to \mathbb{R}^m$ konvergiert.

c) M ist gleichstetig und gleichmäßig beschränkt.

Allgemeiner steht a)⇒b) in Stromberg (1996), S. 167. Vgl. Heuser (1990), S. 563. Die Äquivalenz „a)⇔b)“ wird von Heuser als „Satz von Arzelà-Ascoli“ bezeichnet; c) liefert aber eine weitere wesentliche Vertiefung.

Beweis **„a)⇒b)“** Das ist die wesentliche Aussage, die nach Stromberg (1996) noch viel allgemeiner gilt! $\mathbb{R}^n$ ist separabel, also ist auch K separabel, es gibt also eine abzählbare dichte Teilmenge $D \subset K$. Sei $\left(f_k\right)_{k \geqslant 1}$ eine Folge aus M. Dann liefert Lemma 2.1.4 die Existenz einer Teilfolge $\left(f_{k_j}\right)_{j \geqslant 1}$, die auf D konvergiert. Mit Satz 2.1.3 folgt jetzt Aussage b).

„b)⇒c)" Wäre M nicht gleichstetig, so gäbe es ein $\varepsilon_0 > 0$ mit folgender Eigenschaft: Zu jedem $\delta > 0$ existiert ein $f \in M$ und $x, y \in K$ mit $\|x - y\| < \delta$ und $\|f(x) - f(y)\| \geqslant \varepsilon_0$. Zu $\delta := \frac{1}{k}$ wähle man ein solches $f = f_k$ und dazu $x_k, y_k \in K$ mit $\|x_k - y_k\| < \frac{1}{k}$ und $\|f_k(x_k) - f_k(y_k)\| > \varepsilon_0$. Da K kompakt ist, hat $(x_k)_{k\geqslant 1}$ eine konvergente Teilfolge. Wir können somit oBdA gleich annehmen, dass $x_k \to x \in K$, also auch $y_k \to x \in K$. Nach Voraussetzung hat $(f_k)_{k\geqslant 1}$ eine gleichmäßig konvergente Teilfolge $f_{k_j} \xrightarrow{j\to\infty} f$, und dann gilt wegen

$$\left\|f_{k_j}\left(x_{k_j}\right) - f(x)\right\| \leqslant \left\|f_{k_j}\left(x_{k_j}\right) - f\left(x_{k_j}\right)\right\| + \left\|f\left(x_{k_j}\right) - f(x)\right\| \xrightarrow{j\to\infty} 0,$$

und dem Analogon für y_{k_j}:

$$f_{k_j}\left(x_{k_j}\right) - f_{k_j}\left(y_{k_j}\right) \xrightarrow{j\to\infty} f(x) - f(x) = 0.$$

Das ist aber ein Widerspruch, denn die Terme auf der linken Seite sind alle in der Norm $\geqslant \varepsilon_0$. Folglich ist M gleichstetig.
Wäre M nicht gleichmäßig beschränkt, so gäbe es zu jedem $k \in \mathbb{N}$ ein $f_k \in M$ und ein $x_k \in K$ mit $\|f_k(x_k)\| \geqslant k$. $(f_k)_{k\geqslant 1}$ hat eine gleichmäßig konvergente Teilfolge $\left(f_{k_j}\right)_{j\geqslant 1}$, etwa $f_{k_j} \xrightarrow[\text{gleichmäßig}]{j\to\infty} f$ auf K. Aber f ist beschränkt, während für alle j gilt: $\left\|f_{k_j}\left(x_{k_j}\right)\right\| \geqslant k_j > j$. Das ist ein Widerspruch, demnach ist M gleichmäßig beschränkt.

„c)⇒a)" trivial □

Kehren wir nun zurück zu den Differentialgleichungen mit

2.1.6 Satz *Es seien $I := [x_0, x_0 + \ell] \subset \mathbb{R}$, $f\colon I \times \mathbb{R}^n \to \mathbb{R}^n$ stetig und beschränkt. Dann existiert zu jedem $y_0 \in \mathbb{R}^n$ eine Lösung $y\colon I \to \mathbb{R}^n$ der Differentialgleichung $y' = f(x, y)$ mit $y(x_0) = y_0$.*

Beweis Wir approximieren die gesuchte Lösung durch passende Streckenzüge im Richtungsfeld: Teilung von I durch fortgesetzte Halbierung der Intervalle: Teilung $\mathfrak{T}_k$: $x_{k,\nu} := x_0 + \nu\frac{\ell}{2^k}$ für $\nu = 0, 1, \ldots, 2^k$, $k \geqslant 0$; Streckenzug: $y_k\colon I \to \mathbb{R}^n$ mit

$$y_k(x) := \begin{cases} y_0 & \text{für } x = x_0 \text{ (vorg. AW für alle } k \geqslant 0) \\ y_k\left(x_{k,\nu}\right) + \left(x - x_{k,\nu}\right) f\left(x_{k,\nu}, y_k\left(x_{k,\nu}\right)\right) & \text{für } x_{k,\nu} < x \leqslant x_{k,\nu+1},\ \nu = 0, \ldots, 2^k - 1. \end{cases}$$

$\Rightarrow$ $y_k\colon I \to \mathbb{R}^n$ ist stetig (aber nicht notwendig differenzierbar). Wir wollen y_k als Integral einer Funktion z_k schreiben und setzen für $k \geqslant 0$:

$$z_k\colon I \to \mathbb{R}^n, \quad z_k(x) := \begin{cases} f\left(x_0, y_0\right) & \text{für } x = x_0, \\ f\left(x_{k,\nu}, y_k\left(x_{k,\nu}\right)\right), & \text{falls } x_{k,\nu} < x \leqslant x_{k,\nu+1},\ 0 \leqslant \nu \leqslant 2^k - 1, \end{cases}$$

und für diese x ist

$$\begin{aligned} y_k(x) &= y_k\left(x_{k,\nu}\right) + \int_{x_{k,\nu}}^{x} z_k(t)\,dt \\ &= y_k\left(x_{k,0}\right) + \sum_{\mu=1}^{\nu}\left(y_k\left(x_{k,\mu}\right) - y_k\left(x_{k,\mu-1}\right)\right) + \int_{x_{k,\nu}}^{x} z_k(t)\,dt \\ &= y_0 + \sum_{\mu=1}^{\nu}\int_{x_{k,\mu-1}}^{x_{k,\mu}} z_k(t)\,dt + \int_{x_{k,\nu}}^{x} z_k(t)\,dt \\ &= y_0 + \int_{x_0}^{x} z_k(t)\,dt, \end{aligned}$$

also für $x_0 \leqslant x \leqslant x_0 + \ell$:

$$y_k(x) = y_0 + \int_{x_0}^{x} z_k(t)\,dt. \tag{2.11}$$

Auf die Folge $\left(y_k\right)$ wollen wir Satz 2.1.5 anwenden:

(i) Die Folge $\left(y_k\right)$ ist gleichmäßig beschränkt, denn: Sei z. B. $\|f\|_2 \leqslant C$ auf $I \times \mathbb{R}^n$. Dann ist für $x \in I$

$$\left\|y_k(x)\right\|_2 \leqslant \left\|y_0\right\|_2 + \left\|\int_{x_0}^{x} \underbrace{z_k(t)}_{\|\cdot\|_2 \leqslant C}\,dt\right\|_2 \leqslant \left\|y_0\right\|_2 + \ell C. \tag{2.12}$$

(ii) $\left(y_k\right)$ ist gleichgradig stetig, sogar gleichstetig, denn für alle $t, x \in I$ ist

$$\left\|y_k(t) - y_k(x)\right\|_2 = \left\|\int_{x}^{t} z_k(t)\,dt\right\|_2 \leqslant C\,|x-t|\,. \tag{2.13}$$

Satz 2.1.5 liefert nun: Es gibt eine Teilfolge $\left(y_{k_j}\right)_{j\geqslant 1}$ und eine stetige Funktion $y\colon I \to \mathbb{R}^n$ mit $y_{k_j} \xrightarrow[\text{glm.}]{k\to\infty} y$ auf I. Wir zeigen: Mit diesen k_j gilt: $z_{k_j}(x) \xrightarrow[\text{glm.}]{j\to\infty} f\left(x, y(x)\right)$ auf I.
Begründung: Für $x = x_0$ ist die Konvergenz klar. Sei $x_0 < x \leqslant x_0 + \ell$, $j \geqslant 0$. Dann existiert ein $\nu_j = \nu_j(x) \in \left\{0, \ldots, 2^{k_j} - 1\right\}$, so dass $x_{k_j,\nu_j} < x \leqslant x_{k_j,\nu_j+1}$. Dann ist

$$\left|x_{k_j,\nu_j} - x\right| \leqslant 2^{-k_j}\ell \tag{2.14}$$

und damit

$$\left\|y_{k_j}\left(x_{k_j,\nu_j}\right) - y(x)\right\|_2 \leqslant \underbrace{\left\|y_{k_j}\left(x_{k_j,\nu_j}\right) - y_{k_j}(x)\right\|_2}_{\leqslant C2^{-k_j}\ell \text{ nach (2.13) wg. (2.14)}} + \underbrace{\left\|y_{k_j}(x) - y(x)\right\|_2}_{\xrightarrow[\text{glm.}]{j\to\infty} 0}$$

Also $y_{k_j}\left(x_{k_j,\nu_j}\right) \xrightarrow[\text{glm.}]{j\to\infty} y(x)$ auf I und ebenso $x_{k_j,\nu_j} \xrightarrow[\text{glm.}]{j\to\infty} x$ auf I. Weiter ist nach (2.12) $\left\|y_{k_j}\left(x_{k_j,\nu_j}\right)\right\|_2 \leqslant \|y_0\|_2 + \ell C =: R$, und f ist auf dem Kompaktum $K := I \times [-R,R]^n$ gleichmäßig stetig. Daher folgt:

$$\left\|z_{k_j}(x) - f(x, y(x))\right\|_2 = \left\|f\left(x_{k_j,\nu_j}, y_{k_j}\left(x_{k_j,\nu_j}\right)\right) - f(x, y(x))\right\|_2 < \varepsilon$$

für alle $j \geqslant j_0$ und alle $x \in I$. Also: $z_{k_j}(x) \xrightarrow[\text{glm.}]{j\to\infty} f(x, y(x))$. Damit ist die Zwischenbehauptung bewiesen.
Grenzübergang $j \to \infty$ in (2.11) liefert:

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, dt \qquad (x \in I)$$

f ist stetig, nach dem Hauptsatz also y differenzierbar mit $y' = f(x, y)$, und nach Konstruktion ist $y(x_0) = y_0$. □

2.1.7 Satz (Existenzsatz von Peano) *Es seien $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$, $\alpha, \beta > 0$,*

$$K := \left\{\begin{pmatrix} x \\ y \end{pmatrix} : |x - x_0| \leqslant \alpha,\ \|y - y_0\|_2 \leqslant \beta\right\} \subset \mathbb{R}^{n+1},$$

und $f\colon K \to \mathbb{R}^n$ sei stetig (also auch beschränkt),

$$M := \max\left\{\|f(x,y)\|_2 : (x,y) \in K\right\},$$

$$\delta := \begin{cases} \min\left(\alpha, \frac{\beta}{M}\right), & \textit{falls } M > 0, \\ \alpha, & \textit{falls } M = 0. \end{cases}$$

Dann gibt es mindestens eine Lösung $y\colon [x_0 - \delta, x_0 + \delta] \to \mathbb{R}^n$ der Differentialgleichung (2.1) *mit $y(x_0) = y_0$.*

Kurzfassung: Bei stetigem f ist das Anfangswertproblem $y' = f(x, y)$, $y(x_0) = y_0$ lokal lösbar, und über das Existenzintervall der Lösung gibt Satz 2.1.7 Auskunft.

Beweis Wir setzen $J := [x_0 - \alpha, x_0 + \alpha]$,

$$g\colon J \times \mathbb{R}^n \to \mathbb{R}^n, \qquad g(x,y) := \begin{cases} f(x,y) & \text{für } \|y - y_0\|_2 \leqslant \beta, \\ f\left(x, y_0 + \frac{\beta}{\|y - y_0\|_2}(y - y_0)\right) & \text{für } \|y - y_0\|_2 > \beta \end{cases}$$

Dann ist g stetig und beschränkt:

$$\|g(x,y))\|_2 \leqslant M := \max\left\{\|f(x,y)\|_2 : \begin{pmatrix} x \\ y \end{pmatrix} \in K\right\}.$$

Wir betrachten die Differentialgleichungen

$$\begin{aligned} y' &= g(x,y) \qquad \text{auf } [x_0, x_0+\alpha] \\ y' &= -g(-x,y) \qquad \text{auf } [-x_0, -x_0+\alpha] \end{aligned}$$

und Satz 2.1.6 liefert die Existenz von Lösungen y_1, y_2 dieser Differentialgleichungen: $y_1\colon [x_0, x_0+\alpha] \to \mathbb{R}^n$ und $y_2\colon [-x_0, -x_0+\alpha] \to \mathbb{R}^n$ mit $y_1(x_0) = y_0$, $y_2(-x_0) = y_0$. Wir setzen

$$y\colon J \to \mathbb{R}^n, \qquad y(x) := \begin{cases} y_1(x), & \text{falls } x \in [x_0, x_0+\alpha], \\ y_2(-x), & \text{falls } x \in [x_0-\alpha, x_0]. \end{cases}$$

Dann ist y stetig und für $x \neq x_0$ differenzierbar mit $y' = g(x,y)$. Wir prüfen die Differenzierbarkeit im Punkte x_0: Rechtsseitig gilt: $y'_+(x_0) = g(x_0, y_0)$, linksseitig: $y'_-(x_0) = -y'_2(-x_0) = g(x_0, y_2)$. Also ist y in ganz J differenzierbar mit $y' = g(x,y)$ für alle $x \in J$ und $y(x_0) = y_0$. Daher ist für alle $x \in [x_0-\alpha, x_0+\alpha]$ schon

$$y(x) = y_0 + \int_{x_0}^{x} g(t, y(t))\, dt$$

also $\|y(x) - y_0\|_2 \leqslant M|x - x_0|$. Für $|x - x_0| < \delta$ aus Satz 2.1.7 ist $\|y(x) - y_0\|_2 \leqslant M\delta \leqslant \beta$, $[x_0-\delta, x_0+\delta] \subset J$ und $g|_{[x_0-\delta,x_0+\delta]\times\overline{K_\beta(y_0)}} = f|_{[x_0-\delta,x_0+\delta]\times\overline{K_\beta(y_0)}}$. Also leistet $y|_{[x_0-\delta,x_0+\delta]}$ das Verlangte. □

2.1.8 Bemerkung In Satz 2.1.7 braucht die Lösung nicht eindeutig zu sein: In Beispiel 1.1.3 haben wir gesehen, dass das Anfangswertproblem $y' = \sqrt{|y|}$, $y(0) = 0$ unendlich viele Lösungen hat.

2.2 Der Existenz- und Eindeutigkeitssatz von Picard und Lindelöf

benannt nach Emile Picard (1856–1941), französischer Mathematiker, Doktorvater von Henri Lebesgue, bekannt durch die Picardschen Sätze in der Funktionentheorie, und
Ernst Lindelöf (1870–1946), finnischer Mathematiker, Begründer der sog. „finnischen Schule“ der Funktionentheorie.
Bekanntlich heißt eine Funktion $g\colon I \to \mathbb{R}$ ($I \subset \mathbb{R}$ sei ein Intervall) *Lipschitz-stetig* in $x_0 \in I$, benannt nach Rudolf Lipschitz (1832–1903), deutscher Mathematiker, ab 1864 Professor in Bonn, bekannt durch Arbeiten zur Analysis sowie zu Grundsatzfragen und der Zahlentheorie, falls es eine Umgebung U von x_0 und ein $\ell > 0$ gibt, so dass für alle $x \in U \cap I$ gilt:

$$|g(x) - g(x_0)| \leqslant C|x - x_0|.$$

Wir brauchen einen entsprechenden Begriff für die Funktion $g(x,y)$ auf der rechten Seite der Differentialgleichung (2.1) nur in bezug auf die Variable y:

2.2.1 Definition (Lipschitz-Bedingung) Sei $G \subset \mathbb{R}^{n+1}$, $f\colon G \to \mathbb{R}^m$. Wir schreiben die Punkte aus G in der Form (x,y) mit $x \in \mathbb{R}$ und $y \in \mathbb{R}^n$.

a) f genügt in G lokal einer *Lipschitz-Bedingung* (in bezug auf y), wenn zu jedem Punkt $(a,b) \in G$ eine Umgebung U und eine Konstante $C \geqslant 0$ existieren, so dass

$$\left\| f\left(x,y_1\right) - f\left(x,y_2\right)\right\|_2 \leqslant C \left\| y_1 - y_2 \right\|_2$$

für alle $\left(x,y_1\right), \left(x,y_2\right) \in U \cap G$.

b) f genügt in G einer Lipschitz-Bedingung (in bezug auf y), falls eine Konstante $C > 0$ existiert, so dass

$$\left\| f\left(x,y_1\right) - f\left(x,y_2\right)\right\|_2 \leqslant C \left\| y_1 - y_2 \right\|_2$$

für alle $\left(x,y_1\right), \left(x,y_2\right) \in G$.

2.2.2 Satz *Es seien $G \subset \mathbb{R}^{n+1}$ offen, und $f\colon G \to \mathbb{R}^m$, $\left(x,y\right) \mapsto f\left(x,y\right)$ sei in bezug auf y total differenzierbar[1], und die partiellen Ableitungen von f seien lokal beschränkt. Das ist z. B. dann erfüllt, wenn f bzgl. y stetig partiell differenzierbar ist. Dann genügt f lokal einer Lipschitz-Bedingung bzgl. y.*

Beweis Der Beweis genügt im Fall $m = 1$, denn aus

$$\left| f_\mu\left(x,y_1\right) - f_\mu\left(x,y_2\right)\right| \leqslant C \left\| y_1 - y_2 \right\|_2$$

für $\mu = 1,\ldots,m$ folgt für $f = \left(f_1,\ldots,f_m\right)^\top$:

$$\left\| f\left(x,y_1\right) - f\left(x,y_2\right)\right\|_2 = \sqrt{\sum_{\mu=1}^{m} \left| f_\mu\left(x,y_1\right) - f_\mu\left(x,y_2\right)\right|^2} \leqslant \sqrt{m} C \left\| y_1 - y_2 \right\|_2 .$$

Sei $(a,b) \in G$, und die Umgebung $U \subset G$ von (a,b) sei konvex, und es gelte: $\left| \frac{\partial f}{\partial y_\nu}\left(x,y\right)\right| \leqslant M$ für alle $\left(x,y\right) \in U$ und $\nu = 1,\ldots,n$. Seien $(x,u),(x,v) \in U$. Wir betrachten die Funktion $g(t) := f(x, u + t(v-u))$ (beachte: U ist konvex!) für $t \in \left]-\varepsilon, 1+\varepsilon\right[$ mit hinreichend kleinem $\varepsilon > 0$. Dann ist g laut Kettenregel differenzierbar und der Mittelwertsatz für Funktionen einer Variablen liefert:

$$\begin{aligned}
\left| f(x,v) - f(x,u)\right| &= \left| g(1) - g(0)\right| = \left| g'(\xi)\right| \\
&= \left| \sum_{\nu=1}^{n} (v_\nu - u_\nu) \underbrace{\frac{\partial f}{\partial y_\nu}(x, u + \xi(v-u))}_{|\cdot| \leqslant M} \right| \\
&\leqslant M \sum_{\nu=1}^{n} \underbrace{|v_\nu - u_\nu|}_{\leqslant \|v-u\|_2} \leqslant nM \|u - v\|_2 .
\end{aligned}$$

mit passendem $\xi \in [0,1]$. □

[1] nach Roelcke (1961) reicht sogar partielle Differenzierbarkeit

Der Beweis lässt leicht erkennen, wann f auf ganz $K \subset G$ einer (globalen) Lipschitz-Bedingung genügt.

2.2.3 Satz (Existenz- und Eindeutigkeitssatz von Picard und Linedelöf) *Es seien $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$, $\alpha, \beta > 0$,*

$$K := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+1} : \; |x - x_0| \leqslant \alpha, \; \left\| y - y_0 \right\|_2 \leqslant \beta \right\}$$

und $f \colon K \to \mathbb{R}^n$ sei stetig und genüge auf K bzgl. y einer Lipschitz-Bedingung. Ferner seien

$$M := \max \left\{ \left\| f\left(x, y\right) \right\|_2 : \; \left(x, y\right) \in K \right\},$$
$$\delta := \begin{cases} \min\left(\alpha, \frac{\beta}{M}\right), & \text{falls } M > 0, \\ \alpha, & \text{falls } M = 0. \end{cases}$$

Dann gibt es genau eine differenzierbare Funktion $y \colon [x_0 - \delta, x_0 + \delta] \to \mathbb{R}^n$, die der Differentialgleichung $y' = f\left(x, y\right)$ und der Anfangsbedingung $y(x_0) = y_0$ genügt. Diese Lösung y wird erhalten durch die Methode der sukzessiven Approximation: Man wähle irgendeine stetige Funktion $y_1 \colon [x_0 - \delta, x_0 + \delta] \to \mathbb{R}^n$ mit $y_1(x_0) = y_0$, $\left\| y_1(x) - y_0 \right\|_2 \leqslant \beta$ für alle $x \in [x_0 - \delta, x_0 + \delta]$ (z. B. $y_1(x) := y_0$ für alle x) und setze iterativ für $k \geqslant 1$:

$$y_{k+1}(x) := y_0 + \int_{x_0}^{x} f\left(t, y_k(t)\right) dt \qquad (x \in [x_0 - \delta, x_0 + \delta]). \tag{2.15}$$

Dann ist $\left(y_k\right)$ sinnvoll und konvergiert gegen y gleichmäßig auf $[x_0 - \delta, x_0 + \delta]$.

Bemerkung Eine axiomiatische Fassung der Methode der sukzessiven Approximation findet sich bei Barner und Flohr (2000) auf S. 146 (Banachscher Fixpunktsatz) und die Anwendung auf den Satz von Picard und Lindelöf ebenda auf S. 154f. Dort wird der Fixpunktsatz für vollständige metrische Räume ausgesprochen; bei Heuser (1995a) und Walter (1976a) wird dieser Satz nur für Banach-Räume formuliert.

Beweis **Existenz:** Die y_k sind wohldefiniert. Das zeigt man per Induktion:

$k = 1$: klar

$k \mapsto k+1$: Sei bekannt, dass $\left\| y_k(t) - y_0 \right\|_2 \leqslant \beta$ für alle $t \in [x_0 - \delta, x_0 + \delta]$. Dann ist $f\left(t, y_k(t)\right)$ sinnvoll und

$$\left\| y_{k+1}(x) - y_0 \right\|_2 = \left\| \int_{x_0}^{x} \underbrace{f\left(t, y_k(t)\right)}_{\|\cdot\|_2 \leqslant M} dt \right\|_2 \leqslant M \underbrace{|x - x_0|}_{\leqslant \delta} \leqslant \beta$$

für $|x - x_0| \leqslant \delta$.

Die Definition der y_k ist somit sinnvoll.

Behauptung Es sei (Lipschitz-Bedingung)

$$\left\| f(x,u) - f(x,v) \right\|_2 \leqslant C \left\| u - v \right\|$$

für alle $(x,u), (x,v) \in K$. Dann gilt für $|x - x_0| < \delta$:

$$\left\| y_{k+1}(x) - y_k(x) \right\|_2 \leqslant 2\beta \frac{(C|x-x_0|)^{k-1}}{(k-1)!} \qquad (k \geqslant 1).$$

BEGRÜNDUNG $k = 1$:

$$\left\| y_2(x) - y_1(x) \right\|_2 \leqslant \underbrace{\left\| y_2(x) - y_0 \right\|_2}_{\leqslant \beta \text{ s.o.}} + \underbrace{\left\| y_1(x) - y_0 \right\|_2}_{\leqslant \beta}.$$

Bemerkung Wählt man gleich $y_1 := y_0$, so ist

$$\left\| y_2(x) - y_1(x) \right\|_2 = \left\| \int_{x_0}^{x} \underbrace{f\left(t, y_0\right)}_{\|\cdot\|_2 \leqslant M} dt \right\| \leqslant M|x - x_0| = \frac{M}{C}(C|x-x_0|) \leqslant M\delta \leqslant \beta$$

Steigt man damit in die weitere Induktion ein mit

$$\left\| y_{k+1}(x) - x_k(x) \right\|_2 \leqslant \frac{M}{C} \frac{(C|x-x_0|)^k}{k!},$$

so folgt dort

$$\left\| y_{k+2}(x) - y_{k+1}(x) \right\|_2 \leqslant C \left| \int_{x_0}^{x} \frac{M}{C} \frac{|C(t-x_0)|^k}{k!} dt \right| = \frac{M}{C} \frac{(C|x-x_0|)^{k+1}}{(k+1)!}.$$

Das führt zu etwas schärferen Abschätzungen.

$k \mapsto k+1$:

$$\begin{aligned}
\left\| y_{k+2}(x) - y_{k+1}(x) \right\|_2 &= \left\| \int_{x_0}^{x} \left(f\left(t, y_{k+1}(t)\right) - f\left(t, y_k(t)\right) \right) dt \right\|_2 \\
&\leqslant C \left| \int_{x_0}^{x} \underbrace{\left\| y_{k+1}(t) - y_k(t) \right\|_2}_{\leqslant 2\beta \frac{(C|t-x_0|)^{k-1}}{(k-1)!}} dt \right| \\
&\leqslant \frac{2\beta C^k}{(k-1)!} \left| \int_{x_0}^{x} |t - x_0|^{k-1} \, dt \right| \\
&= 2\beta \frac{(C|x-x_0|)^k}{k!}
\end{aligned}$$

✓

Da die Reihe $\sum_{k=1}^{\infty}\frac{(C|x-x_0|)^{k-1}}{(k-1)!}$ gleichmäßig für $|x-x_0|\leqslant\delta$ konvergiert (Exponentialreihe), folgt: Die Reihe

$$y_1(x)+\sum_{k=1}^{\infty}\left(y_{k+1}(x)-y_k(x)\right)$$

konvergiert gleichmäßig auf $[x_0-\delta, x_0+\delta]$, d. h. $y(x) := \lim_{k\to\infty} y_k(x)$ ist stetig auf $[x_0-\delta, x_0+\delta]$.

Behauptung y ist differenzierbar, $y' = f\left(x,y\right)$ auf $[x_0-\delta, x_0+\delta]$. und $y(x_0)=y_0$.

Begründung Für $|x-x_0|\leqslant\delta$ ist nach dem obigen

$$\left\|y(x)-y_0\right\|_2 = \lim_{k\to\infty}\left\|y_k(x)-y_0\right\|_2 \leqslant \beta.$$

Die Abbildung $x\mapsto f\left(x,y(x)\right)$ ist also sinnvoll, und nach der Lipschitz-Bedingung ist

$$\left\|f\left(x,y_k(x)\right)-f\left(x,y(x)\right)\right\|_2 \leqslant C\left\|y_k(x)-y(x)\right\|_2 \xrightarrow[\text{glm.}]{k\to\infty} 0.$$

$\Rightarrow f\left(x,y_k(x)\right)\xrightarrow[\text{glm.}]{k\to\infty} f\left(x,y(x)\right)$ bzgl. $x\in[x_0-\delta,x_0+\delta]$. Der Grenzübergang $k\to\infty$ in (2.15) ist wegen dieser gleichmäßigen Konvergenz legitim und liefert:

$$y(x) = y_0+\int_{x_0}^{x} f\left(t,y(t)\right)dt \qquad (|x-x_0|\leqslant\delta) \qquad \checkmark$$

Damit ist der Existenzbeweis abgeschlossen. Wir zeigen als

Zusatz Es gilt die Fehlerabschätzung:

$$\left\|y(x)-y_{k+1}(x)\right\|_2 \leqslant 2\beta\sum_{j=k}^{\infty}\frac{(C|x-x_0|)^j}{j!} \leqslant 2\beta\frac{(C|x-x_0|)^k}{k!}e^{C|x-x_0|}$$

für alle $k\geqslant 0$, $|x-x_0|\leqslant\delta$.

Begründung Es ist

$$y(x) = y_{k+1}(x)+\sum_{j=k+1}^{\infty}\left(y_{j+1}(x)-y_j(x)\right) \qquad (|x-x_0|\leqslant\delta).$$

und damit

$$\begin{aligned}\left\|y(x)-y_{k+1}(x)\right\|_2 &\leqslant \sum_{j=k+1}^{\infty}\left\|y_{j+1}(x)-y_j(x)\right\|_2\\ &\leqslant \sum_{j=k+1}^{\infty} 2\beta\frac{(C|x-x_0|)^{j-1}}{\left(j-1\right)!}\\ &= 2\beta\sum_{j=k}^{\infty}\frac{(C|x-x_0|)^j}{j!}\\ &= 2\beta\left(e^{C|x-x_0|}-\sum_{j=0}^{k-1}\frac{(C|x-x_0|)^j}{j!}\right).\end{aligned}$$

Nun gilt nach der Taylor-Formel: $e^x = \sum_{j=0}^{k-1} \frac{x^j}{j!} + R_k(x)$ mit dem Restglied von Lagrange $R_k(x) = \frac{\varphi^{(k)}(\xi)}{k!} x^k$, wobei $\varphi(x) := e^x$, mit passendem ξ zwischen 0 und x, also $R_k(x) = e^\xi \frac{x^k}{k!}$. Das liefert

$$\left| e^x - \sum_{j=0}^{k-1} \frac{x^j}{j!} \right| \leqslant \frac{|x|^k}{k!} e^{|x|}.$$

und damit

$$\left\| y(x) - y_{k+1}(x) \right\|_2 \leqslant 2\beta \frac{(C|x-x_0|)^k}{k!} e^{C|x-x_0|} \qquad (|x-x_0| \leqslant \delta). \qquad \checkmark$$

Eindeutigkeit: Sei $z\colon [x_0-\delta, x_0+\delta] \to \mathbb{R}^n$ eine zweite Lösung mit $z(x_0) = y_0$, $z' = f(x,z)$. Dann ist

$$z(x) = y_0 + \int_{x_0}^x f(t, z(t))\, dt \qquad (|x-x_0| \leqslant \delta)$$

und ebenso mit y.

$$\begin{aligned}
\Rightarrow \left\| z(x) - y(x) \right\|_2 &= \left\| \int_{x_0}^x \big(f(t,z(t)) - f\big(t, y(t)\big) \big)\, dt \right\|_2 \\
&\overset{\text{Lip.-Bed.}}{\leqslant} C \left| \int_{x_0}^x \underbrace{\left\| z(t) - y(t) \right\|_2}_{\leqslant \|z(t)-y_0\| + \|y_0 - y(t)\| \leqslant 2\beta} dt \right| \\
&\leqslant 2\beta C|x-x_0|.
\end{aligned}$$

$$\Rightarrow \left\| z(x) - y(x) \right\|_2 \leqslant C \left| \int_{x_0}^x \underbrace{\left\| z(t) - y(t) \right\|_2}_{\leqslant 2\beta C|t-x_0|} dt \right| \leqslant 2\beta C^2 \frac{|x-x_0|^2}{2!}$$

und mit Induktion folgt weiter:

$$\left\| z(x) - y(x) \right\|_2 \leqslant 2\beta C^k \frac{|x-x_0|^k}{k!}$$

für alle $k \geqslant 1$, $|x-x_0| \leqslant \delta$. Grenzübergang $k \to \infty$ liefert: $z(x) = y(x)$ für $|x-x_0| \leqslant \delta$. $\square$

Zusatz Satz 2.2.2 gilt entsprechend für

$$K = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x_0 \leqslant x \leqslant x_0 + \alpha,\ \left\| y - y_0 \right\|_2 \leqslant \beta \right\}$$

mit $y\colon [x_0, x_0+\delta] \to \mathbb{R}^n$. Der Beweis verläuft wie gehabt.

Manchmal kann man mit der Methode der sukzessiven Approximation wirklich die Lösung der Differentialgleichung bestimmen:

2.2.4 Beispiel $A \in M(n \times n, \mathbb{R})$, $y_0 \in \mathbb{R}^n$, $x_0 \in \mathbb{R}$, und gesucht sei die Lösung der Differentialgleichung $y' = Ay$, $y(x_0) = y_0$. Man sieht leicht, dass eine Lipschitz-Bedingung erfüllt ist. Sukzessive Approximation:

$$\begin{aligned} y_1(x) &:= y_0, \\ y_{k+1}(x) &:= y_0 + \int_{x_0}^{x} A y_k(t)\, dt \qquad (x \in \mathbb{R},\ k \geqslant 1) \end{aligned}$$

Behauptung

$$y_{k+1}(x) = \sum_{j=0}^{k} \frac{(x-x_0)^j}{j!} A^j y_0 \qquad (k \geqslant 0) \tag{2.16}$$

BEGRÜNDUNG Induktion.

$k = 0$: ist klar

$k \mapsto k+1$: Gelte (2.16) mit k an Stelle von $k+1$. Dann folgt

$$y_{k+1}(x) = y_0 + \int_{x_0}^{x} \sum_{j=0}^{k-1} \frac{(t-x_0)^j}{j!} A^{j+1} y_0\, dt = \sum_{j=0}^{k} \frac{(x-x_0)^j}{j!} A^j y_0. \qquad \checkmark$$

Für $k \to \infty$ gilt daher:

$$y(x) = \lim_{k\to\infty} y_{k+1}(x) = \sum_{j=0}^{\infty} \frac{(x-x_0)^j}{j!} A^j y_0 = e^{(x-x_0)A} y_0,$$

und wir wissen schon von früher, dass das die Lösung des Problems ist.

2.2.5 Beispiel $y' = 2xy$ auf $\mathbb{R}^2$, $y(0) = y_0$. Iteration: $y_1 := y_0$,

$$y_{k+1}(x) := y_0 + 2\int_0^x t y_k(t)\, dt.$$

Induktion liefert:

$$y_{k+1}(x) = y_0 \left(1 + x^2 + \frac{x^4}{2!} + \ldots + \frac{x^{2k}}{k!}\right).$$

Also $y_k(x) \xrightarrow{k\to\infty} y_0 e^{x^2}$, und man überzeugt sich sofort, dass dies die Lösung ist.

Die Eindeutigkeitsaussage von Satz 2.2.3 gilt allgemeiner („global“) und nicht nur „lokal“ auf hinreichend kleinem Intervall $[x_0 - \delta, x_0 + \delta]$:

2.2.6 Satz (Eindeutigkeitssatz) *Es seien $G \subset \mathbb{R}^{n+1}$ und $f\colon G \to \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung (bzgl. y) genüge. Ferner seien $I \subset \mathbb{R}$ ein Intervall, $y_1, y_2\colon I \to \mathbb{R}^n$ zwei Funktionen mit* Graph $y_j \subset G$ *für $j = 1,2$, und y_1, y_2 seien Lösungen der Differentialgleichung $y' = f\left(x, y\right)$ auf I. Gilt dann $y_1\left(x_0\right) = y_2\left(x_0\right)$ für ein $x_0 \in I$, so ist $y_1 = y_2$.*

Beweis **Behauptung** Gilt $y_1(a) = y_2(a)$ für ein $a \in I$, so gibt es ein $\delta > 0$, so dass $y_1(x) = y_2(x)$ für alle $x \in I$ mit $|x - a| \leqslant \delta$.

Begründung $y_j' = f\left(x, y_j\right)$ für alle $x \in I$, $j = 1,2$.

$$y_j(x) = y_j(a) + \int_a^x f\left(t, y_j(t)\right) dt \qquad \left(x \in I,\ j = 1,2\right)$$

$$\Rightarrow y_1(x) - y_2(x) = \int_a^x \left(f\left(t, y_1(t)\right) - f\left(t, y_2(t)\right)\right) dt \qquad (x \in I)$$

f ist lokal Lipschitz-stetig, zum Punkt $\left(a, y_1(a)\right) = \left(a, y_2(a)\right) \in G$ gibt es also eine Umgebung U und ein $C > 0$, so dass

$$\left\| f(t,u) - f(t,v) \right\|_2 \leqslant C \left\| u - v \right\|_2$$

für alle $(t,u), (t,v) \in U \cap G$. – Wir nehmen an, a sei kein rechter Eckpunkt von I und zeigen die Behauptung für ein Intervall $[a, a+\delta]$. (Ist a auch kein linker Eckpunkt von I, so zeigt man ebenso die Existenz eines δ, so dass die Behauptung auf $[a-\delta, a]$ gilt.)
y_1, y_2 sind stetig, es gibt also ein $\delta_1 > 0$, so dass $[a, a+\delta_1] \subset I$ und $\left(t, y_j(t)\right) \in U \cap G$ für $j = 1,2$. $\Rightarrow$ Für $a \leqslant x \leqslant a + \delta_1$ gilt:

$$\left\| y_1(x) - y_2(x) \right\|_2 \leqslant C \left| \int_a^x \left\| y_1(t) - y_2(t) \right\|_2 dt \right| .$$

Wir setzen für $a \leqslant x \leqslant a + \delta_1$:

$$M(x) := \max\left\{ \left\| y_1(t) - y_2(t) \right\|_2 :\ t \in [a,x] \right\}$$

und haben $\left\| y_1(x) - y_2(x) \right\|_2 \leqslant C|x-a|\, M(x)$, also für $a \leqslant t \leqslant x \leqslant a + \delta_1$:

$$\left\| y_1(t) - y_2(t) \right\|_2 \leqslant C|t-a|\, M(t) \leqslant C|x-a|\, M(x),$$

und das Supremum auf der linken Seite gibt: $M(x) \leqslant C|x-a|\, M(x)$. Mit $\delta := \min\left(\delta_1, \frac{1}{2(C+1)}\right)$ gilt dann für $a \leqslant x \leqslant a + \delta$: $M(x) \leqslant \frac{1}{2} M(x)$, d. h. $M(x) = 0$. $\Rightarrow y_1(x) = y_2(x)$ für $a \leqslant x \leqslant a + \delta$. Gleicher Schluss „nach links“ liefert die Behauptung. ✓

Behauptung $y_1(x) = y_2(x)$ für alle $x \in I$, $x \geqslant x_0$.

Begründung Sei $a := \sup\left\{x \in I\colon\ y_1(t) = y_2(t) \text{ für } x_0 \leqslant t \leqslant x\right\}$. Ist $a = \infty$ oder a rechter Eckpunkt von I, so ist die Behauptung klar wegen der Stetigkeit von y_1, y_2. Sonst gibt es ein $\delta_1 > 0$, so dass $[a, a+\delta-1] \subset I$, und wegen der Stetigkeit von y_1, y_2 ist $y_1(a) = y_2(a)$. Die erste Behauptung liefert die Existenz eines $\delta > 0$ mit $[a, a+\delta] \subset I$, so dass

$$y_1|_{[a,a+\delta]} = y_2|_{[a,a+\delta]} .$$

Das aber ist ein Widerspruch zur Maximalität von a, und so folgt die Behauptung. ✓

Gleicher Schluss „nach links“ liefert: $y_1(x) = y_2(x)$ für alle $x \in I$, $x \leqslant x_0$. Damit folgt die Behauptung. □

2.2.7 Satz (Existenz- und Eindeutigkeitssatz) *Es seien $G \subset \mathbb{R}^{n+1}$ offen, und $f\colon G \to \mathbb{R}^n$ sei stetig und genüge lokal einer Lipschitz-Bedingung bzgl. y. Dann gilt: Für jedes $(x_0, y_0) \in G$ hat das Anfangswertproblem $y' = f(x,y)$, $y(x_0) = y_0$ genau eine Lösung $y\colon J \to \mathbb{R}^n$, Graph $y \subset G$, die nicht echt fortsetzbar ist. Diese Lösung kommt nach rechts und nach links dem Rand von G beliebig nahe in folgendem Sinne:*

$$\overline{\{(x, y(x)) : x \in J, x \geqslant x_0\}} =: \overline{\Gamma^+}, \text{ und}$$
$$\overline{\{(x, y(x)) : x \in J, x \leqslant x_0\}} =: \overline{\Gamma^-}$$

sind keine kompakten Teilmengen von G.

Bemerkung $\overline{\Gamma^+}$ keine kompakte Teilmenge von G ist gleichbedeutend mit dem Vorliegen eines der folgenden Fälle (b rechter Endpunkt von J):

(i) $b = \infty$ oder

(ii) $b \in \mathbb{R}$ und $\overline{\lim}_{x \to b-0} \|y(x)\|_2 = \infty$ oder

(iii) $b \in \mathbb{R}$ und $\overline{\lim}_{x \to b-0} \|y(x)\|_2 < \infty$, aber $\overline{\Gamma^+} \cap \operatorname{Rd} G \neq \emptyset$.

Entsprechend mit $\overline{\Gamma^-}$.

Beweis Seien $\alpha > 0, \beta > 0$ so gewählt, dass

$$K = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : |x - x_0| \leqslant \alpha,\ \|y - y_0\|_2 \leqslant \beta \right\} \subset G.$$

Nach Aufgabe 44 ist $f|_K$ Lipschitz-stetig bzgl. y; ferner ist $f|_K$ stetig. Satz 2.2.3 liefert damit die Existenz eines $\delta > 0$ und einer Lösung $y\colon [x_0 - \delta, x_0 + \delta] \to \mathbb{R}^n$ mit $y' = f(x,y)$, $y(x_0) = y_0$, und dieses y ist eindeutig bestimmt. Gleicher Schluss mit $x_0 + \delta$ und Anfangswert $y(x_0 + \delta)$ und der Zusatz zum Beweis von Satz 2.2.2 zeigen: y kann über $x_0 + \delta$ hinaus nach rechts (und analog über $x_0 - \delta$ hinaus nach links) fortgesetzt werden, und zwar eindeutig (nach Satz 2.2.6). Sei

$$\mathfrak{I} := \{I \subset \mathbb{R} : I \text{ Intervall},\ x_0 \in I,\ \exists y_I\colon I \to \mathbb{R}^n,\ y_I' = f(x, y_I),\ y_I(x_0) = y_0\}$$

Dann ist $J := \bigcup_{I \in \mathfrak{I}} I$ ein Intervall. Wir definieren für $x \in J$: $y(x) := y_I(x)$, falls $x \in I$. Diese Definition ist sinnvoll, denn ist auch $x \in \tilde{I} \in \mathfrak{I}$, so ist $y_I(x) = y_{\tilde{I}}(x)$, da $y_I|_{I \cap \tilde{I}} = y_{\tilde{I}}|_{I \cap \tilde{I}}$ nach Satz 2.2.6. Also ist $y\colon J \to \mathbb{R}^n$ die Lösung des Anfangswertproblems $y' = f(x,y)$, $y(x_0) = y_0$ mit maximalem Definitionsbereich, d. h. y ist nicht mehr echt fortsetzebar zu einer Lösung der Differentialgleichung.

Behauptung $\overline{\Gamma^+}$ ist keine kompakte Teilmenge von G.

BEGRÜNDUNG Angenommen, b sei das rechte Intervallende von J und $\overline{\Gamma^+} \subset G$ sei kompakt. Dann ist $b \in \mathbb{R}$. Wir zeigen zunächst: $b \in J$: $\overline{Gamma^+}$ ist kompakt und f stetig, also ist $\left\| f|_{\overline{\Gamma^+}} \right\|_2 \leqslant C$ für geeignetes $C > 0$. Nach der Differentialgleichung ist

$$y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\, dt \qquad (x_0 \leqslant x \leqslant b)$$

Also ist für $x_0 \leqslant x, x' \leqslant b$:

$$\left\| y(x) - y\left(x'\right) \right\|_2 = \left\| \int_{x'}^{x} \underbrace{f\big(t, y(t)\big)}_{\|\cdot\|_2 \leqslant C} dt \right\|_2 \leqslant C\left|x - x'\right|$$

In dieser Situation liefert das Cauchy-Kriterium: $\lim_{x \to b-0} y(x) =: y^*$ existiert; offenbar ist $\left(b, y^*\right) \in \overline{\Gamma^+} \subset G$. Angenommen, es sei $b \notin J$. Dann setzen wir

$$z(x) := \begin{cases} y(x) & \text{für } x \in J, \\ y* & \text{für } x = b \end{cases}$$

und zeigen: z ist Lösung des Anfangswertproblems. Offenbar ist $z|_J = y$ Lösung des Anfangswertproblems, und z ist stetig. Weiter ist $\operatorname{Graph} z \subset \overline{\Gamma^+} \subset G$ und

$$\begin{aligned} z(b) = y^* &= \lim_{x \to b-0} y(x) \\ &= \lim_{x \to b-0} \left(y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\, dt \right) \\ &= y_0 + \int_{x_0}^{b} f(t, z(t))\, dt, \end{aligned}$$

denn f und z sind stetig. Für $x \in J$ ist ohnehin

$$z(x) = y_0 + \int_{x_0}^{x} f(t, z(t))\, dt,$$

d. h. diese Gleichung gilt für alle $x \in J \cup \{b\}$. Also ist z auf ganz $J \cup \{b\}$ differenzierbar mit $z' = f(x, z)$, ferner ist $z(x_0) = y_0$, z also Lösung des Anfangswertproblems und echte Fortsetzung von y. Das aber ist ein Widerspruch, also ist $b \in J$.
$\Rightarrow \big(b, y(b)\big) \in G$, und die Lösung y lässt sich nach Satz 2.2.3 über b hinaus fortsetzen. Das aber ist ein Widerspruch. Folglich ist $\overline{\Gamma^+}$ keine kompakte Teilmenge von G. ✓

Analog zeigt man: $\overline{\Gamma^-}$ ist keine kompakte Teilmenge von G. □

2.3 Systeme linearer Differentialgleichungen

2.3.1 Definition (System linearer Differentialgleichungen) Seien $I \subset \mathbb{R}$ ein Intervall und $A\colon I \to M(n \times n, \mathbb{R})$, $b\colon I \to \mathbb{R}^n$ stetig. Dann heißt $y' = A(x)\, y + b(x)$ für $x \in I$ ein *System linearer Differentialgleichungen*. Analog über $\mathbb{C}$.

2.3.2 Satz *Sind A, b, I wie in Definition 2.3.1, so gibt es zu jedem $(x_0, y_0) \in I \times \mathbb{R}^n$ genau eine auf ganz I definierte Lösung $y\colon I \to \mathbb{R}^n$ des Anfangswertproblems*

$$\begin{aligned} y' &= A(x)\,y + b(x) \qquad (x \in I) \\ y(x_0) &= y_0. \end{aligned} \tag{2.17}$$

Ebenso für komplexwertige Lösungen.

Beweis Unsere Differentialgleichung hat die Form $y' = f(x,y)$ mit $f(x,y) := A(x)\,y + b(x)$ für $(x,y) \in I \times \mathbb{R}^n$. Sei $L \subset I$ kompakt. Dann ist für alle $x \in L$ und $y, y' \in \mathbb{R}^n$

$$\left\| f(x,y) - f(x,y') \right\|_2 = \left\| A(x)\,(y - y') \right\|_2 \leqslant C \left\| y - y' \right\|_2$$

mit geeignetem $C > 0$, denn: Ist $A = (\alpha_{j,k})_{j,k}$, $y = (y_1, \ldots, y_n)^\top$, so ist

$$\|Ay\|_2 = \left(\sum_{j=1}^n \left(\sum_{k=1}^n \alpha_{j,k} y_k \right)^2 \right)^{\frac{1}{2}} \overset{\text{CSU}}{\leqslant} \left(\sum_{j=1}^n \left(\sum_{k=1}^n \alpha_{j,k}^2 \underbrace{\sum_{\ell=1}^n y_\ell^2}_{=\|y\|_2^2} \right) \right)^{\frac{1}{2}} = \left(\sum_{j,k=1}^n \alpha_{j,k}^2 \right)^{\frac{1}{2}} \|y\|,$$

und hier sind die $\alpha_{j,k}$ auf I stetig, also auf L beschränkt. Satz 2.2.7 liefert jetzt die Existenz einer Lösung $y\colon J \to \mathbb{R}^n$ mit maximalem Definitionsbereich $J \subset I$.

Behauptung $J = I$

Begründung Seien c rechter Eckpunkt von J, c_1 rechter Eckpunkt von I und angenommen, dass $c < c_1$. Für $x_0 \leqslant x < c$ ist

$$y(x) = y_0 + \int_{x_0}^x f\bigl(t, y(t)\bigr)\,dt,$$

und wie im Beweis von Satz 2.2.7 folgt: $y^* := \lim_{x \to c-0} y(x)$ existiert, ist Lösung der Differentialgleichung, und Satz 2.2.3 mit Anfangswert (c, y^*) liefert: y ist über c hinaus nach rechts fortsetzbar. Das aber ist ein Widerspruch zur Maximalität von J. $\Rightarrow c = c_1$.
Es könnte immer noch gelten $c = c_1$, aber $c \notin J$, $c \in I$. Um diesen Fall zu widerlegen, verwenden wir nochmals das obige Argument: $\lim_{x \to c-0} y(x) =: y(c)$ existiert, und der Grenzübergang liefert:

$$y(x) = y_0 + \int_{x_0}^x f\bigl(t, y(t)\bigr)\,dt,$$

gilt auch für $x_0 \leqslant x \leqslant c$. Also kann c zum Definitionsbereich von y hinzugenommen werden. J ist aber maximal, mit $c \in I$ folgt also $c \in J$. Der gleiche Schluss nach links liefert die Behauptung $I = J$. ✓

2.3.3 Satz *Sind A, b, I wie in Definition 2.3.1, so gilt:*

a) $V := \{y\colon I \to \mathbb{R}^n\colon y' = A(x)\,y\}$ *ist ein n-dimensionaler $\mathbb{R}$-Vektorraum.*

b) Man erhält genau alle Lösungen der inhomogenen Differentialgleichung $y' = A(x)\,y + b(x)$, indem man zu einer speziellen („partikulären") Lösung der inhomogenen Differentialgleichung eine geeignete Lösung der homoegenen Differentialgleichung $y' = A(x)\,y$ addiert.

Beweis a) Sei $x_0 \in I$ fest und $\varphi\colon V \to \mathbb{R}^n$, $\varphi(y) := y(x_0)$ für $y \in V$. Dann ist φ ein Vektorraum-Isomorphismus nach Satz 2.3.2.

b) klar □

2.3.4 Folgerung Sei $A\colon I \to M(n \times n, \mathbb{R})$ stetig, und vorgelegt sei die homogene Differentialgleichung

$$y' = Ay \qquad (\text{auf } I)\,. \tag{2.18}$$

a) Es gibt n linear unabhängige Lösungen von (2.18) $y_1, \ldots, y_n\colon I \to \mathbb{R}^n$; je n linear unabhängige Lösungen bilden eine Basis des Lösungsraums V. Jedes solche System $y_1, \ldots, y_n$ heißt ein *Hauptsystem* oder ein *Fundamentalsystem* von (2.18).

b) Ist $y_1, \ldots, y_n$ irgendein System von Lösungen von (2.18), so genügt die zugehörige „Lösungsmatrix" $Y := (y_1, \ldots, y_n) : I \to M(n \times n, \mathbb{R})$ der Differentialgleichung $Y' = AY$.

c) Sei $y_1, \ldots, y_n$ ein Hauptsystem. Ist $z\colon I \to \mathbb{R}^n$ eine Lösung von $y' = A(x)\,y$, so existieren $\gamma_1, \ldots, \gamma_n \in \mathbb{R}$ mit

$$z = \gamma_1 y_1 + \ldots + \gamma_n y_n = Yc, \qquad c = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Das bedeutet:

(1) $Z = (z_1, \ldots, z_n) : I \to M(n \times n, \mathbb{R})$ ist Lösungsmatrix genau dann, wenn es ein $C \in M(n \times n, \mathbb{R})$ gibt mit $Z = YC$.

(2) $Z = (z_1, \ldots, z_n)$ ist Hauptsystem genau dann, wenn es ein $C \in \mathrm{GL}_n(\mathbb{R})$ gibt mit $Z = YC$.

d) Ist $x_0 \in I$, $Y_0 \in M(n \times n, \mathbb{R})$, so existiert genau eine Lösungsmatrix Y mit $Y(x_0) = Y_0$.

e) Seien $y_1, \ldots, y_n$ Lösungen von (2.18), $Y := (y_1, \ldots, y_n)$. Nach dem Beweis von Satz 2.3.3 a) ist für jedes $x_0 \in I$ die Abbildung

$$\varphi_{x_0}\colon V \to \mathbb{R}^n, \quad \varphi_{x_0}(y) := y(x_0) \in \mathbb{R}^n$$

ein Isomorphismus. Daher gilt:

$$\begin{aligned}&Y \text{ ist Hauptsystem}\\ \Leftrightarrow&\forall\, x_0 \in I \quad Y(x_0) \in \mathrm{GL}_n(\mathbb{R})\\ \Leftrightarrow&\exists\, x_0 \in I \quad Y(x_0) \in \mathrm{GL}_n(\mathbb{R})\end{aligned}$$

(Wenn ein solches $x_0 \in I$ existiert, so sind $y_1, \ldots, y_n$ linear unabhängig, da $\varphi_{x_0}(y_1), \ldots, \varphi_{x_0}(y_n)$ linear unabhängig sind)

$$\begin{aligned}\Leftrightarrow& Y\colon \mathbb{R}^n \to V, \quad c \mapsto Yc \text{ ist surjektiv}\\ \Leftrightarrow& Y\colon \mathbb{R}^n \to V, \quad c \mapsto Yc \text{ ist injektiv}\end{aligned}$$

f) Sei $x_0 \in I$, und das Hauptsystem $\Phi = (\varphi_1, \ldots, \varphi_n)\colon I \to M(n \times n, \mathbb{K})$ sei definiert durch $\Phi(x_0) = E_n$, d. h. $\varphi_j(x_0) = e_j := j$-ter Einheitsvektor im $\mathbb{R}^n$. Ist dann $y_0 \in \mathbb{R}^n$ beliebig vorgegeben, so ist $y = \Phi y_0$ die Lösung des Anfangswertproblems $y' = A(x)\,y$, $y(x_0) = y_0$. Also: Sind $y_1, \ldots, y_n$ Lösungen von (2.18) und $Y = (y_1, \ldots, y_n)$, so gilt: $Y(x) = \Phi(x)\,Y(x_0)$.

2.3.5 Satz (Wronskische Determinante) *Sind $A\colon I \to M(n \times n, \mathbb{R})$ stetig und $y_1, \ldots, y_n\colon I \to \mathbb{R}^n$ Lösungen der homogenen Differentialgleichung $y' = Ay$ und $W(x) := \det\big(y_1(x), \ldots, y_n(x)\big)$ für $x \in I$, so gilt: $W' = \operatorname{Spur} A \cdot W$, also*

$$W(x) = W(x_0)\exp\left(\int_{x_0}^{x} \big(\operatorname{Spur} A(t)\big)\,dt\right) \qquad (x, x_0 \in I).$$

$W(x)$ *heißt die* Wronskische Determinante*, benannt nach dem polnischen Mathematiker Graf Hoëné Wronski, 1778–1853, der diese Determinante 1821 einführte.*

Beweis Sei $y_k := (y_{1,k}, \ldots, y_{n,k})^\top$ der k-te Spaltenvektor von W, $z_j := (y_{j,1}, \ldots, y_{j,n})$ der j-te Zeilenvektor von W. Dann ist nach dem in Lemma 1.5.11 Gesagten:

$$W' = \begin{vmatrix} z_1' \\ z_2 \\ \vdots \\ z_n \end{vmatrix} + \ldots + \begin{vmatrix} z_1 \\ \vdots \\ z_{n-1} \\ z_n' \end{vmatrix} = \sum_{j=1}^{n} \begin{vmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j' \\ z_{j+1} \\ \vdots \\ z_n \end{vmatrix}.$$

Hier ist

$$\begin{aligned} z_j' &= \left(y_{j,1}', \ldots, y_{j,n}'\right)\\ &= \left((Ay_1)_j, \ldots, (Ay_n)_j\right) = \left(\sum_{k=1}^{n} a_{j,k} y_{k,1}, \ldots, \sum_{k=1}^{n} a_{j,k} y_{k,n}\right)\\ &= \sum_{k=1}^{n} a_{j,k}\,(y_{k,1}, \ldots, y_{k,n}) = \sum_{k=1}^{n} a_{j,k} z_k. \end{aligned}$$

$$\Rightarrow \begin{vmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j' \\ z_{j+1} \\ \vdots \\ z_n \end{vmatrix} = a_{j,j} \begin{vmatrix} z_1 \\ \vdots \\ z_{j-1} \\ z_j \\ z_{j+1} \\ \vdots \\ z_n \end{vmatrix} = a_{j,j} W.$$

$\Rightarrow W' = \left(\operatorname{Spur} A\right) W.$

$$\Rightarrow W(x) = W(x_0) \exp\left(\int_{x_0}^{x} \left(\operatorname{Spur} A(t)\right) dt\right). \qquad \square$$

2.3.6 Korollar In Satz 2.3.5 gilt: Die Wronski-Determinante $W(x) = \det\left(y_1(x), \dots, y_n(x)\right)$ ist entweder konstant $= 0$ oder nullstellenfrei auf I. Es gilt: $y_1, \dots, y_n$ bilden ein Hauptsystem genau dann, wenn für alle $x \in I$ gilt: $W(x) \neq 0$, was genau dann der Fall ist, wenn für irgendein $x_0 \in I$ gilt: $W(x_0) \neq 0$.

Bemerkung Eine kunstvolle andere Art der Berechnung von W findet man bei Walter (1976a) auf S. 112; dabei wird Folgerung 2.3.4 wesentlich benutzt. – Roelcke (1961) schließt auf S. 51 reichlich kompliziert mit dem Entwicklungssatz. In beiden Fällen werden die Spalten von W differenziert, was wegen der Differentialgleichung $y_k' = Ay_k$ nahe liegt, aber die weitere Rechnung nicht so transparent macht wie der obige Beweis mit der Differentiation der Zeilen.

Das Problem der Bestimmung der Lösungen von (2.17) zerfällt in zwei Teile:

(A) Bestimmung einer partikulären Lösung;

(B) Bestimmung eines Hauptsystems von (2.18).

(A) können wir lösen, wenn ein Hauptsystem bekannt ist:

2.3.7 Satz *Seien I, A, b wie in Definition 2.3.1 und $y_1, \dots, y_n$ ein Hauptsystem von* (2.18). *Dann gibt es zu jeder Lösung y von* (2.17) *stetig differenzierbare Funktionen $c_1, \dots, c_n \colon I \to \mathbb{R}$, so dass $y = \sum_{j=1}^n c_j y_j$. Notwendig und hinreichend dafür, dass $\sum_{j=1}^n c_j y_j$ eine Lösung von* (2.17) *ist, ist die Differentialgleichung*

$$\sum_{j=1}^{n} c_j' y_j = b \qquad (\textit{auf}\, I) \tag{2.19}$$

für $c_1, \dots, c_n$ („Variation der Konstanten“). Ist $Y := \left(y_1, \dots, y_n\right)$, $c_0 \in \mathbb{R}^n$ beliebig, $x_0 \in I$, so liefert

$$c(x) = \begin{pmatrix} c_1(x) \\ \vdots \\ c_n(x) \end{pmatrix} := c_0 + \int_{x_0}^{x} Y^{-1}(t)\, b(t)\, dt \qquad (x \in I)$$

eine Lösung der Differentialgleichung (2.19) *und damit eine partikuläre Lösung von* (2.17)*, und zwar die Lösung* $y = \sum_{j=1}^{n} c_j y_j = Y c$.

Beweis $y = \sum_{j=1}^{n} c_j y_j$ ist Lösung von (2.17) genau dann, wenn $y = Y c$ Lösung von (2.17) ist, also wenn

$$y' = Y' c + Y c' = A y + b = A Y c + b,$$

und das ist genau dann der Fall, wenn $Y c' = b$ ist (das ist (2.19)), also $c' = Y^{-1} b$, das aber heißt, dass

$$c(x) = c_0 + \int_{x_0}^{x} Y^{-1}(t)\, b(t)\, dt \qquad (x, x_0 \in I)$$

mit irgendeinem $c_0 \in \mathbb{R}^n$. Beachte: Y^{-1} ist stetig (sogar stetig differenzierbar) nach der Cramerschen Regel, d. h. $t \mapsto Y^{-1}(t)\, b(t)$ ist stetig. □

Wir haben also das Problem (A) auf das Problem (B) zurückgeführt. Dieses Problem lässt sich allgemein nicht ohne weiteres lösen (wohl aber im Spezialfall, dass A konstante Koeffizienten hat). Aber: Wenn man eine nicht-triviale Lösung (z. B. durch Raten) gefunden hat, so lässt sich das Problem reduzieren auf ein System von $n-1$ Differentialgleichungen mit Satz 2.3.8.

Bekannt seien $p \geqslant 1$ linear unabhängige Lösungen $y_1, \dots, y_p$ für $1 \leqslant p < n$ der homogenen Differentialgleichung (2.18); wir bilden die Matrix

$$Y = \underbrace{(y_1, \dots, y_p)}_{\in M(n \times p, \mathbb{R})} = \begin{pmatrix} U \\ V \end{pmatrix},$$

wobei $U \in M(p \times p, \mathbb{R})$ und $V \in M((n-p) \times p, \mathbb{R})$ sei, und setzen voraus: U ist invertierbar auf I. Ggf. muss man dazu die Koordinaten von $y_1, \dots, y_p$ vorher geeignet nummerieren, damit diese Voraussetzung (wenigstens auf einem Teilintervall von I) erfüllt ist. Die Matrix $\begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix}$ ist invertierbar, und $\begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix}^{-1} = \begin{pmatrix} U^{-1} & 0 \\ -VU^{-1} & E_q \end{pmatrix}$ ist stetig differenzierbar. Daher können wir die weiteren Lösungen von (2.18) ansetzen in der Form

$$y = \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} = Y w + \begin{pmatrix} 0 \\ z \end{pmatrix} = \begin{pmatrix} U w \\ V w + z \end{pmatrix}$$

mit $w\colon I \to \mathbb{R}^p$ und $z\colon I \to \mathbb{R}^q$ stetig differenzierbar. Wir formulieren die Differentialgleichung $y' = A y$ um zu Bedingungen an w und z. Bei diesem Ansatz gilt:

$$y' = \left(Y w + \begin{pmatrix} 0 \\ z \end{pmatrix} \right)' = Y' w + Y w' + \begin{pmatrix} 0 \\ z' \end{pmatrix} = A Y w + \begin{pmatrix} U w' \\ V w' + z' \end{pmatrix}, \tag{2.20}$$

und zerlegen wir $A = \begin{pmatrix} * & B \\ * & C \end{pmatrix}$ mit $B \in M(p \times q, \mathbb{R})$ und $C \in M(q \times q, \mathbb{R})$, so folgt:

$$A y = A \left(Y w + \begin{pmatrix} 0 \\ z \end{pmatrix} \right) = A Y w + \begin{pmatrix} * & B \\ * & C \end{pmatrix} \begin{pmatrix} 0 \\ z \end{pmatrix} = A Y w + \begin{pmatrix} B z \\ C z \end{pmatrix}. \tag{2.21}$$

Also gilt nach (2.20) und (2.21):

$$y' = Ay \Longleftrightarrow \begin{cases} Uw' = Bz, \\ Vw' + z' = Cz \end{cases} \qquad \text{d. h.}$$

$$y' = Ay \Longleftrightarrow \begin{cases} w' = U^{-1}Bz, \\ z' = \left(C - VU^{-1}B\right)z \end{cases}$$

Das ist ein homogenes System von nur noch q Gleichungen für z, und wenn man die Lösungen dieses Systems bestimmt hat, findet man w aus $w' = U^{-1}Bz$ durch einfache Quadratur. Diese Überlegungen fassen wir zusammen zu

2.3.8 Satz (Reduktionsverfahren von d'Alembert) *benannt nach Jean-Baptist le Rond d'Alembert, 1717–1783, französischer Mathematiker, Physiker und Philosoph.*
Gegeben sei die stetige Funktion $A\colon I \to M(n \times n, \mathbb{R})$, und bekannt seien p $(1 \leqslant p < n)$ linear unabhängige Lösungen $y_1, \dots, y_n$ der Differentialgleichung (2.18) *mit der Eigenschaft, dass in*

$$\left(y_1, \dots, y_p\right) = Y = \begin{pmatrix} U \\ V \end{pmatrix}$$

mit $U \in M\left(p \times p, \mathbb{R}\right)$ und $V \in M\left(q \times p, \mathbb{R}\right)$ $(q = n - p)$ gilt: $U(t) \in \mathrm{GL}_p(\mathbb{R})$ für alle $t \in I$. Ferner sei $A = \left(\begin{smallmatrix} * & B \\ * & C \end{smallmatrix}\right)$ mit $B \in M\left(p \times q, \mathbb{R}\right)$ und $C \in M\left(q \times q, \mathbb{R}\right)$. Dann erhält man genau alle Lösungen y von (2.18) *in der Form*

$$y = \begin{pmatrix} Uw \\ Vw + z \end{pmatrix}$$

mit $w\colon I \to \mathbb{R}^p$ und $z\colon I \to \mathbb{R}^q$ stetig differenzierbar, wenn z, w die Lösungen des Systems

$$\begin{cases} z' = \left(C - VU^{-1}B\right)z \\ w' = U^{-1}Bz \end{cases}$$

durchlaufen. ($z = 0$ und $w = e_j$ liefert die bekannte Lösung y_j.)

Ein Beispiel dazu findet sich in Aufgabe 41.

Zusatz (zu Satz 2.3.8) Es seien in der Situation von Satz 2.3.8 $z_1, \dots, z_q\colon I \to \mathbb{R}^q$ ein Hauptsystem der Differentialgleichung

$$z' = \underbrace{\left(C - VU^{-1}B\right)}_{\in M(q \times q, \mathbb{R})} z$$

und $w_1, \dots, w_q\colon I \to \mathbb{R}^p$ ein zugehöriges System von Stammfunktionen gemäß $w_k' = U^{-1}Bz_k$ für $k = 1, \dots, q$ und

$$y_{p+k} := \begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix} \qquad \left(k = 1, \dots, q\right).$$

Dann ist $y_1, \dots, y_p, y_{p+1}, \dots, y_{p+q} = y_n$ ein Hauptsystem von (2.18).

Beweis Angenommen, es seien $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ und $\sum_{j=1}^n \lambda_j y_j = 0$, d. h.

$$\sum_{j=1}^p \lambda_j y_j + \sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix} = Y \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} + \sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix}$$
$$= \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix} = 0.$$

Weiter ist

$$\begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix} = \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \begin{pmatrix} w_k \\ z_k \end{pmatrix},$$

also

$$\sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} Uw_k \\ Vw_k + z_k \end{pmatrix} = \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} w_k \\ z_k \end{pmatrix},$$

und wegen der Invertierbarkeit von $\left(\begin{smallmatrix} U & 0 \\ V & E_q \end{smallmatrix}\right)$ haben wir

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{k=1}^q \lambda_{p+k} \begin{pmatrix} w_k \\ z_k \end{pmatrix} = 0. \tag{2.22}$$

$\Rightarrow \sum_{k=1} q\lambda_{p+k} z_k = 0$, also wegen der linearen Unabhängigkeit der z_k schon $\lambda_{p+1}, \dots, \lambda_{p+q} = \lambda_n = 0$. Dann sagt aber (2.22): $\lambda_1 = \ldots = \lambda_p = 0$, also verschwinden alle λ_j. □

Beweis (alternativ) mit Hilfe der Wronski-Determinante: Wir schreiben:

$$\begin{aligned}
(y_1, \dots, y_p, y_{p+1}, \dots, y_{p+q}) &= \left(Y, \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \begin{pmatrix} w_1 & \dots & w_q \\ z_1 & \dots & z_q \end{pmatrix} \right) \\
&= \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \left(e_1, \dots, e_p, \begin{pmatrix} w_1 \\ z_1 \end{pmatrix}, \dots, \begin{pmatrix} w_q \\ z_q \end{pmatrix} \right) \\
&= \begin{pmatrix} U & 0 \\ V & E_q \end{pmatrix} \begin{pmatrix} E_p & W \\ 0 & Z \end{pmatrix}
\end{aligned}$$

mit $Y = \binom{U}{V} \in M\left(n \times p, \mathbb{R}\right)$ ($U \in M\left(p \times p, \mathbb{R}\right)$, $V \in M\left(q \times p, \mathbb{R}\right)$) und $W = \left(w_1, \ldots, w_q\right) \in M\left(p \times q, \mathbb{R}\right)$ und $Z = \left(z_1, \ldots, z_q\right) \in M\left(q \times q, \mathbb{R}\right)$. Daher ist

$$\det\left(y_1, \ldots, y_p, y_{p+1}, \ldots, y_{p+q}\right) = \det U \det Z \neq 0,$$

denn U ist nach Voraussetzung invertierbar, und $z_1, \ldots, z_q$ ist ein Hauptsystem. □

2.4 Lineare Differentialgleichungen n-ter Ordnung

2.4.1 Definition (Lineare Differentialgleichung n-ter Ordnung) Eine lineare Differentialgleichung n-ter Ordnung in Normalform ist eine Differentialgleichung der Form

$$y^{(n)} + a_{n-1}(x)\, y^{(n-1)} + \ldots + a_1(x)\, y' + a_0(x)\, y = b(x) \qquad (x \in I). \tag{2.23}$$

mit stetigen Funktionen $a_0, \ldots, a_{n-1}, b\colon I \to \mathbb{R}$. Ist $b = 0$, so heißt die Differentialgleichung homogen, ansonsten inhomogen. Analog für $\mathbb{C}$.

Setzt man $z := \left(y, y', \ldots, y^{(n-1)}\right)^\top = (z_1, \ldots, z_n)^\top$, so ist (2.23) äquivalent zum System

$$z' = \begin{pmatrix} z_1' \\ \vdots \\ z_n' \end{pmatrix} = \begin{pmatrix} y' \\ \vdots \\ y^{(n-1)} \\ -\left(a_0 y + \ldots + a_{n-1} y^{(n-1)}\right) + b \end{pmatrix} = A(x)\, z + b(x)\, e_n$$

mit

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

Da A und b stetig sind auf I, haben wir sofort

2.4.2 Satz *Seien $a_0, \ldots, a_{n-1}, b\colon I \to \mathbb{R}$ stetig.*

a) *Zu allen $x_0 \in I$, $y_0, \ldots, y_{n-1} \in \mathbb{R}$ gibt es genau eine (auf ganz I definierte) Lösung $y\colon I \to \mathbb{R}$ der Differentialgleichung* (2.23) *mit $y(x_0) = y_0$, $y'(x_0) = y_1$, ..., $y^{(n-1)}(x_0) = y_{n-1}$.*

b) *Die Menge V der Lösungen der homogenen Differentialgleichung*

$$y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_0 y = 0 \qquad \left(\text{auf } I\right) \tag{2.24}$$

bildet einen n-dimensionalen Vektorraum über $\mathbb{R}$. Eine Basis von V nennt man ein Hauptsystem *oder* Fundamentalsystem *der Differentialgleichung* (2.24).

c) *Man erhält genau alle Lösungen der inhomogenen Differentialgleichung* (2.23)*, indem man zu einer speziellen (sog. „partikulären") Lösung der inhomogenen Differentialgleichung eine geeignete Lösung der homogenen Differentialgleichung addiert.*

Beweis a) klar nach der Vorbemerkung und Satz 2.3.2.

b) Sei $\mathscr{L}$ der Lösungsraum des Systems $z' = A(x)\,z$ (A wie oben). Dann ist $\varphi\colon V \to \mathscr{L}$, $y \mapsto z = \left(y, \ldots, y^{(n-1)}\right)^{\top}$ ein Isomorphismus. Daher ist die Behauptung klar nach Satz 2.3.3

c) klar □

2.4.3 Satz (Wronskische Determinante) *Es seien $a_0, \ldots, a_{n-1}\colon I \to \mathbb{R}$ stetig und $y_1, \ldots, y_n$ Lösungen der homogenen Differentialgleichung* (2.24)*,*

$$W := \begin{vmatrix} y_1 & \ldots & y_n \\ y_1' & \ldots & y_n' \\ \vdots & & \vdots \\ y_1^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix}$$

die Wronskische Determinante. Dann gilt:

a) $W'(x) = -a_{n-1}(x)\,W(x)$*, also*

$$W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} a_{n-1}(t)\,dt\right) \qquad (x, x_0 \in I)$$

b) $y_1, \ldots, y_n$ *sind linear unabhängig genau dann, wenn es ein* $x_0 \in I$ *gibt mit* $W(x_0) \neq 0$*, und das ist genau dann der Fall, wenn für alle* $x \in I$ *gilt:* $W(x) \neq 0$.

Beweis a) Man betrachte die Bijektion $y \mapsto z := \left(y, y', \ldots, y^{(n-1)}\right)^{\top}$ zwischen den Lösungen von (2.24) und

$$z' = A(x)\,z, \qquad (A(x) \text{ von Seite 87}). \tag{2.25}$$

Dann liefert Folgerung 2.3.4: $W' = \left(\operatorname{Spur} A(x)\right) W = -a_{n-1}(x)\,W$, also

$$W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} a_{n-1}(t)\,dt\right) \qquad (x, x_0 \in I).$$

b) $y_1, \ldots, y_n$ sind nach Satz 2.4.2 a) und b) linear unabhängig genau dann, wenn es ein $x_0 \in I$ gibt, so dass $W(x_0) \neq 0$, da das dortige φ ein Isomorphismus ist, und das ist nach der Formel aus a) genau dann der Fall, wenn für alle $x \in I$ schon $W(x) \neq 0$ gilt.□

Das Problem der Bestimmung der Lösungen der Differentialgleichung (2.23) mit stetigen Koeffizienten $a_0, \ldots, a_{n-1}, b$ zerfällt in zwei Teile:

(A) Bestimmung einer partikulären Lösung der Differentialgleichung (2.23).

(B) Bestimmung eines Hauptsystems der homogenen Differentialgleichung (2.24).

Die Lösung von (A) bei bekannter Lösung von (B) ist Thema in

2.4.4 Satz (Variation der Konstanten) *Es seien $a_0,\dots,a_{n-1},b\colon I\to\mathbb{R}$ stetig und $y_1,\dots,y_n$ ein Hauptsystem der homogenen Differentialgleichung* (2.24). *Dann gibt es zu jeder Lösung y der inhomogenen Differentialgleichung* (2.23) *stetig differenzierbare Funktionen $c_1,\dots,c_n\colon I\to\mathbb{R}$, so dass*

$$y=\sum_{j=1}^{n} c_j y_j. \tag{2.26}$$

Beachte: Gleicher Ansatz wie im Falle konstanter Koeffizienten $a_0,\dots,a_{n-1}$: siehe Konstruktion 1.5.12! Gleiche Bedingung (2.27) *wie in Konstruktion 1.5.12!*
Notwendig und hinreichend dafür, dass (2.26) *eine Lösung von* (2.23) *ist, ist das Bestehen der Differentialgleichungen*

$$\sum_{j=1}^{n} c_j' y_j^{(\nu)}=\delta_{\nu,n-1}b \qquad (\nu=0,\dots,n-1) \tag{2.27}$$

für $c_1,\dots,c_n$. Die Funktionen $c_1,\dots,c_n$ genügen genau dann den Gleichungen (2.27)*, wenn $c_1,\dots,c_n$ die Form haben*

$$c_j(x)=c_{j,0}+(-1)^{j+n}\int_{x_0}^{x}\frac{W_j(t)}{W(t)}b(t)\,dt \qquad (x\in I)$$

mit beliebigem festen $x_0\in I$ und beliebigen Konstanten $c_{1,0},\dots,c_{n,0}$, wobei

$$W:=\begin{vmatrix} y_1 & \dots & y_n\\ y_1' & \dots & y_n'\\ \vdots & & \vdots\\ y_1^{(n-1)} & \dots & y_n^{(n-1)}\end{vmatrix}$$

$$W_j:=\begin{vmatrix} y_1 & \dots & y_{j-1} & y_{j+1} & \dots & y_n\\ y_1' & \dots & y_{j-1}' & y_{j+1}' & \dots & y_n'\\ \vdots & & \vdots & \vdots & & \vdots\\ y_1^{(n-2)} & \dots & y_{j-1}^{(n-2)} & y_{j+1}^{(n-2)} & \dots & y_n^{(n-2)}\end{vmatrix}$$

die Wronski-Determinante von $y_1,\dots,y_n$ bzw. die Wronski-Determinante der Ordnung $n-1$ von $y_1,\dots,y_{j-1},y_{j+1},\dots,y_n$ seien.

Beweis Wir betrachten die Zuordnung

$$y\mapsto z:=\begin{pmatrix} y\\ y'\\ \vdots\\ y^{(n-1)}\end{pmatrix}, \qquad y_j\mapsto z_j:=\begin{pmatrix} y_j\\ y_j'\\ \vdots\\ y_j^{(n-1)}\end{pmatrix}$$

wie oben. $y_1,\dots,y_n$ ist ein Hauptsystem für (2.24), also $z_1,\dots,z_n$ ein Hauptsystem für

$$z' = A(x)\,z + be_n, \qquad (A(x) \text{ von Seite 87}) \tag{2.28}$$

Nun haben wir: y ist eine Lösung von (2.23) genau dann, wenn z Lösung von (2.28) ist. Das ist nach Satz 2.3.7 genau dann der Fall, wenn es stetig differenzierbare Funktionen $c_1,\dots,c_n\colon I \to \mathbb{R}$ gibt mit $z = \sum_{j=1}^n c_j z_j$ und $\sum_{j=1}^n c_j' z_j = be_n$. Unter Verwendung obiger Zuordnung heißt das: Es gibt stetig differenzierbare Funktionen $c_1,\dots,c_n\colon I \to \mathbb{R}$ mit $y = \sum_{j=1}^n c_j y_j$ und $\sum_{j=1}^n c_j' y_j^{(\nu)} = \delta_{\nu,n-1} b$ für $\nu = 0,\dots,n-1$. Das hier auftretende lineare Gleichungssystem für die c_j' hat die Determinante $W(x)$, und da $y_1,\dots,y_n$ linear unabhängig sind, ist $W(x)$ nullstellenfrei, d. h. man kann $c_1',\dots,c_n'$ mit der Cramerschen Regel ausrechnen. Sei

$$C_j := \begin{vmatrix} y_1 & \cdots & y_{j-1} & 0 & y_{j+1} & \cdots & y_n \\ y_1' & \cdots & y_{j-1}' & 0 & y_{j+1}' & \cdots & y_n' \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ y_1^{(n-2)} & \cdots & y_{j-1}^{(n-2)} & b & y_{j+1}^{(n-2)} & \cdots & y_n^{(n-2)} \end{vmatrix};$$

Dann ist $c_j'(x) = \frac{C_j(x)}{W(x)}$ für $x \in I$ und $j = 1,\dots,n$. C_j lässt sich bequem berechnen durch Entwicklung nach der j-ten Spalte:

$$C_j(x) = (-1)^{j+n} \begin{vmatrix} y_1 & \cdots & y_{j-1} & y_{j+1} & \cdots & y_n \\ y_1' & \cdots & y_{j-1}' & y_{j+1}' & \cdots & y_n' \\ \vdots & & \vdots & \vdots & & \vdots \\ y_1^{(n-2)} & \cdots & y_{j-1}^{(n-2)} & y_{j+1}^{(n-2)} & \cdots & y_n^{(n-2)} \end{vmatrix} b(x) = (-1)^{j+n} W_j(x)\, b(x).$$

Ergebnis: y ist Lösung von (2.23) genau dann, wenn $y = \sum_{j=1}^n c_j y_j$, wobei $c_1,\dots,c_n\colon I \to \mathbb{R}$ die Form

$$c_j(x) = c_{j,0} + (-1)^{j+n} \int_{x_0}^x \frac{W_j(t)}{W(t)} b(t)\, dt \qquad (x \in I,\ x_0 \in I \text{ fest})$$

haben mit $c_{j,0} \in \mathbb{R}$ für $j = 1,\dots,n$. □

Zusatz Für jedes $x_0 \in I$ ist

$$y(x) = \sum_{j=1}^n \left((-1)^{j+n} \int_{x_0}^x \frac{W_j(t)}{W(t)} b(t)\, dt \right) y_j(x) \qquad (x \in I)$$

eine partikuläre Lösung von (2.23).

2.4.5 Beispiel Sei $n = 2$, y_1, y_2 ein Hauptsystem der Differentialgleichung

$$y'' + a_1 y' + a_0 y = b \tag{2.29}$$

mit stetigen $a_0, a_1, b\colon I \to \mathbb{R}$. Dann ist $W_1(x) = y_2(x)$ und $W_2(x) = y_1(x)$, also ist

$$\begin{aligned} y(x) &:= \left(-\int_{x_0}^{x} \frac{y_2(t)}{W(t)} b(t)\,dt\right) y_1(x) + \left(\int_{x_0}^{x} \frac{y_1(t)}{W(t)} b(t)\,dt\right) y_2(x) \\ &= \int_{x_0}^{x} \frac{\begin{vmatrix} y_1(t) & y_2(t) \\ y_1(x) & y_2(x) \end{vmatrix}}{W(t)} b(t)\,dt \end{aligned}$$

eine spezielle Lösung von (2.29). Hier ist noch bekannt aus Satz 2.4.3:

$$W(t) = W(t_0) \exp\left(-\int_{t_0}^{t} a_1(s)\,ds\right),$$

so dass alles elementar berechenbar wird: Für beliebige $x_0, t_0 \in I$ ist

$$y(x) = \frac{1}{W(t_0)} \int_{x_0}^{x} \big(y_1(t)\,y_2(x) - y_2(t)\,y_1(x)\big) \exp\left(\int_{t_0}^{t} a_1(s)\,ds\right) b(t)\,dt$$

eine partikuläre Lösung von (2.29).

Nun wollen wir mit der Bestimmung eines Hauptsystems für (2.23) auch Problem (B) angehen. Wie bei Systemen lässt sich Problem (B) nicht allgemein lösen (wohl aber z. B. im Falle konstanter Koeffizienten!). Daher bietet sich das Reduktionsverfahren von d'Alembert aus Satz 2.3.8 an. Wendet man aber dieses Verfahren an auf das System (2.28), das zur Differentialgleichung (2.23) in Normalform (d. h. Koeffizient 1 bei $y^{(n)}$) gehört, so wird man auf Systeme geführt, die nicht mehr die spezielle Gestalt (2.28) haben, d. h. die nicht zu linearen Differentialgleichungen der Ordnung q in Normalform gehören. Daher bietet sich folgende Variante des früheren Reduktionsverfahrens von d'Alembert an:

2.4.6 Satz (Reduktionsverfahren von d'Alembert) *Es seien $a_0, \dots, a_{n-1}\colon I \to \mathbb{R}$ stetig, und $y_1\colon I \to \mathbb{R}$ sei eine nullstellenfreie Lösung der Differentialgleichung* (2.24). *Vermöge $y = y_1 v$ entsprechen die Lösungen y von* (2.24) *bijektiv den Lösungen v von*

$$v^{(n)} + b_{n-1} v^{(n-1)} + \ldots + b_1 v' = 0, \tag{2.30}$$

wobei

$$b_j(x) = \frac{1}{y_1(x)} \sum_{i=j}^{n} \binom{i}{j} a_i(x)\, y_1^{(i-j)}(x) \qquad \big(j = 1, \dots, n-1,\ a_n := 1\big). \tag{2.31}$$

Ist $w_1, \dots, w_{n-1}$ ein Hauptsystem der Differentialgleichung

$$w^{(n-1)} + b_{n-1} w^{(n-2)} + \ldots + b_2 w' + b_1 w = 0, \tag{2.32}$$

und sind $v_1, \dots, v_{n-1}$ Stammfunktionen von $w_1, \dots, w_{n-1}$, so ist $y_1, y_2 := v_1 y_1, \dots, y_n := v_{n-1} y_1$ ein Hauptsystem von (2.24).

Beweis Wir setzen $y = y_1 v$ und haben mit $a_n := 1$:

$$\begin{aligned}\sum_{i=0}^{n} a_i y^{(i)} &= \sum_{i=0}^{n} a_i \sum_{j=0}^{i} \binom{i}{j} y_1^{(i-j)} v^{(j)} \\ &= \sum_{j=0}^{n} v^{(j)} \sum_{i=j}^{n} \binom{i}{j} a_i y_1^{(i-j)}\end{aligned}$$

Der Summand für $j = 0$ ist

$$v \underbrace{\sum_{i=0}^{n} a_i y_1^{(i)}}_{=0} = 0,$$

da y_1 Lösung von (2.24) ist, also beginnt obige Summe erst mit $j = 1$. Weiter ist der Beitrag für $j = n$ gleich

$$v^{(n)} \binom{n}{n} a_n y_1^{(n-n)} = v^{(n)} y_1,$$

so dass wir wegen der Nullstellenfreiheit von y_1 mit $b_n := 1$ haben:

$$\sum_{i=0}^{n} a_i y^{(i)} = y_1 \sum_{j=1}^{n} b_j v^{(j)}.$$

Damit ist die Bijektion zwischen den Lösungen von (2.24) und (2.30) klar. Seien nun $w_1, \ldots, w_{n-1}$ ein Hauptsystem von (2.32), $v_1, \ldots, v_{n-1}$ Stammfunktionen von $w_1, \ldots, w_{n-1}$. Dann sind $y_1, y_2 := v_1 y_1, \ldots, y_n := v_{n-1} y_1$ Lösungen von (2.24). Angenommen, $\sum_{j=1} n\lambda_j y_j = 0$ mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Abspalten des Faktors y_1 und Division liefert:

$$\lambda_1 + \sum_{j=2}^{n} \lambda_j v_{j-1} = 0. \tag{2.33}$$

Wir differenzieren und erhalten $\sum_{j=2}^{n} \lambda_j w_{j-1} = 0$. Aber $w_1, \ldots, w_{n-1}$ bilden ein Hauptsystem, folglich sind $\lambda_2 = \ldots, \lambda_n = 0$ und wegen (2.33) auch $\lambda_1 = 0$. Also sind $y_1, \ldots, y_n$ linear unabhängig und somit ein Hauptsystem. □

Damit ist Problem (B) auf das Auffinden einer Lösung der homogenen Differentialgleichung reduziert. Dafür gibt es kein patentrezept, aber oft führt geschicktes Raten und Probieren zum Ziel... Beispiele dazu finden sich bei Heuser (1995a), S. 255ff

2.4.7 Beispiel Sei $n = 2$. Ist y_1 eine Lösung von

$$y'' + a_1 y' + a_0 y = 0, \tag{2.34}$$

und ist y_1 nullstellenfrei auf I, so ist im Sinne von (2.31)

$$\begin{aligned} b_1(x) &= \frac{1}{y_1(x)} \sum_{i=1}^{2} i\, a_i(x)\, y_1^{(i-1)}(x) \qquad (a_2 := 1!) \\ &= \frac{1}{y_1(x)} \left(a_1(x)\, y_1(x) + 2 y_1'(x) \right) \\ &= a_1(x) + 2 \frac{y_1'(x)}{y_1(x)} = a_1(x) + \left(\log y_1^2(x) \right)'. \end{aligned}$$

Beachte: Hier müssen wir $\log y_1^2$ nehmen und nicht $2 \log y_1$, da y_1 negativ sein kann. Die Differentialgleichung (2.32) lautet nun $w' + b_1(x)\, w = 0$, und diese hat die Lösung

$$\begin{aligned} w_1(x) &= \exp\left(c_0 - \int_{x_0}^{x} b_1(t)\, dt \right) \qquad (c_0 \text{ konstant, } x \in I,\ x_0 \in I \text{ fest}) \\ &\overset{(2.31)}{=} C_0 y_1^{-2}(x) \exp\left(- \int_{x_0}^{x} a_1(t)\, dt \right) \end{aligned}$$

mit einer Konstanten $C_0 \neq 0$. w_1 hat die Stammfunktion

$$v_1(x) = C_1 + \int_{x_0}^{x} w_1(t)\, dt$$

mit einer Konstanten C_1 und x_0, x wie oben, und wir erhalten das Hauptsystem $y_1, y_2 := v_1 y_1$. Auf die Werte der Konstanten $C_0 \neq 0$ und $c_1 \in \mathbb{R}$ kommt es dabei überhaupt nicht an. Probe der Differentialgleichung:

$$\begin{aligned} y_2 &= v_1 y_1 \\ y_2' &= v_1' y_1 + v_1 y_1' = w_1 y_1 + v_1 y_1' \\ y_2'' &= w_1' y_1 + w_1 y_1' + v_1' y_1' + v_1 y_1'' = w_1' y_1 + 2 w_1 y_1' + v_1 y_1'' \end{aligned}$$

Also ist

$$\begin{aligned} y_2'' + a_1 y_2' + a_0 y_2 &= v_1 \underbrace{\left(y_1'' + a_1 y_1' + a_0 y_1 \right)}_{=0!} + 2 w_1 y_1' + w_1' y_1 + a_1 w_1 y_1 \\ &= y_1 \left(w_1' + \underbrace{\left(a_1 + \frac{2 y_1'}{y_1} \right)}_{=b_1} w_1 \right) \\ &= 0 \end{aligned}$$

2.4.8 Beispiele a) $y'' + y = 0$. Wir kennen die Lösung $y_1(x) = \cos x$. Wir betrachten diese Lösung im Intervall $I = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$; dort ist y_1 nullstellenfrei. Nach Beispiel 2.4.7 finden wir eine linear unabhängige Lösung y_2 in der Form $y_2 = v_1 y_1$, wobei v_1 Stammfunktion von w_1 ist und w_1 Lösung von $w' + b_1(x)\, w = 0$. Hier ist $a_1(x) = 0$, also

$$\begin{aligned}
b_1(x) &= a_1(x) + 2\frac{y_1'(x)}{y_1(x)} = \left(\log\cos^2(x)\right)', \\
w_1(x) &= C_= y_1^{-2}(x) = C_0 \frac{1}{\cos^2 x}, \\
v_1(x) &= \int_{x_0}^{x} C_0 \frac{1}{\cos^2 x}\, dx = C_0 \tan x + C_1 .
\end{aligned}$$

Wir wählen $C_0 = 1$, $C_1 = 0$; dann ist $v_1 = \tan x$, $y_2 = v_1 y_1 = \sin x$, und wir haben das wohlbekannte Hauptsystem (zunächst nur auf I, aber die Funktionen sind auf ganz $\mathbb{R}$ ein Hauptsystem.)

b) Wir betrachten die Differentialgleichung $y'' - y'\cos x + y \sin x = 0$ mit $I = \mathbb{R}$. Eine Lösung ist $y_1(x) = e^{\sin x}$ für $x \in \mathbb{R}$. Dazu gehört

$$b_1(x) = a_1(x) + 2\frac{y_1'(x)}{y_1(x)} = -\cos x + 2\cos x = \cos x,$$

und ebenso für w: $w' + \cos x\, w = 0$, spezielle nicht-triviale Lösung dazu: $w(x) = e^{-\sin x}$ mit Stammfunktion $v(x) = \int_{x_0}^{x} e^{-\sin t}\, dt$. Das gibt das Hauptsystem

$$\begin{aligned}
y_1(x) &= e^{\sin x}, \\
y_2(x) &= y_1(x)\, v(x) = \int_{x_0}^{x} e^{\sin x - \sin t}\, dt .
\end{aligned}$$

2.5 Stetige und differenzierbare Abhängigkeit der Lösungen

2.5.1 Satz *Es seien $G \subset \mathbb{R}^{n+1}$ ein Gebiet, $f\colon G \to \mathbb{R}^n$ stetig und beschränkt: $\left|f\right| \leqslant K$ auf G, ferner gelte die Lipschitz-Bedingung*

$$\left\| f(x,u) - f(x,v) \right\|_2 \leqslant C \left\| u - v \right\|_2$$

für alle $(x,u), (x,v) \in G$ mit einem $C > 0$. Die Funktionen $y, \hat{y}\colon I \to \mathbb{R}$ seien stetig differenzierbar mit Graph $y \subset G$, Graph $\hat{y} \subset G$, *$y(x_0) = y_0$, $\hat{y}(\hat{x}_0) = \hat{y}_0$ mit $x_0, \hat{x}_0 \in I$, $y_0, \hat{y}_0 \in \mathbb{R}^n$, und es gelte mit gewissen $\varepsilon, \hat{\varepsilon} \geqslant 0$:*

$$\begin{aligned}
&\left\| y'(x) - f\left(x, y(x)\right) \right\|_2 \leqslant \varepsilon, \\
&\left\| \hat{y}'(x) - f\left(x, \hat{y}(x)\right) \right\|_2 \leqslant \hat{\varepsilon} \qquad (x \in I) .
\end{aligned}$$

Dann gilt:

$$\left\| y(x) - \hat{y}(x) \right\|_2 \leqslant \frac{\varepsilon + \hat{\varepsilon}}{C} \left(e^{C|x-x_0|} - 1\right) + \left((K + \hat{\varepsilon})\, |x_0 - \hat{x}_0| + \left\| y_0 - \hat{y}_0 \right\|_2\right) e^{C|x-x_0|}$$

Spezialfälle:

a) $y' = f(x,y)$, $x_0 = \hat{x}_0$, $y_0 = \hat{y}_0$. Dann ist

$$\left\| y(x) - \hat{y}(x) \right\|_2 \leqslant \frac{\hat{\varepsilon}}{C} \left(e^{C|x-x_0|} - 1 \right) \qquad (x \in I)$$

b) $y' = f(x,y)$, $\hat{y}' = f(x,\hat{y})$. Dann ist

$$\left\| y(x) - \hat{y}(x) \right\|_2 \leqslant \left(K\,|x_0 - \hat{x}_0| + \left\| y_0 - \hat{y}_0 \right\|_2 \right) e^{C|x-x_0|} \qquad (x \in I)$$

(Lipschitz-stetige Abhängigkeit von den Anfangswerten)

Beweis Setze $z := y - \hat{y}$. Dann gilt:

$$\begin{aligned}
\|z(x_0)\|_2 &\leqslant \left\| y(x_0) - \hat{y}(\hat{x}_0) \right\|_2 + \left\| \hat{y}(\hat{x}_0) - \hat{y}(x_0) \right\|_2 \\
&= \left\| y_0 - \hat{y}_0 \right\|_2 + \left\| \int_{x_0}^{\hat{x}_0} \hat{y}'(t)\,dt \right\|_2 \\
&\leqslant \left\| y_0 - \hat{y}_0 \right\|_2 + \left\| \int_{x_0}^{\hat{x}_0} \underbrace{\left(\hat{y}'(t) - f\left(t, \hat{y}(t)\right) \right)}_{\|\cdot\|_2 \leqslant \hat{\varepsilon}} dt \right\|_2 + \left\| \int_{x_0}^{\hat{x}_0} \underbrace{f\left(t, \hat{y}(t)\right)}_{\|\cdot\|_2 \leqslant K} dt \right\|_2 \\
&\leqslant \left\| y_0 - \hat{y}_0 \right\|_2 + (K + \hat{\varepsilon})\,|x_0 - \hat{x}_0|\,.
\end{aligned}$$

Das ist die Behauptung für $x = x_0$.
Sei nun $x \in I$, $x > x_0$. Ist $z(x) = 0$, so ist die Behauptung klar. Sei also $z(x) \neq 0$,

$$c := \min\{t \in [x_0, x] : z(s) \neq 0 \quad \forall\, s \in\,]t, x]\}\,.$$

Ist $c > x_0$, so ist $z(c) = 0$. Also gilt stets: $\|z(c)\|_2 \leqslant \|z(x_0)\|_2$. $\|z\|_2\,|_{]c,x]}$ ist stetig differenzierbar mit $\|z\|_2' = \frac{\langle z, z' \rangle}{\|z\|_2}$, und die Cauchy-Schwarzsche Ungleichung liefert:

$$\left| \|z\|_2' \right| \leqslant \left\| z' \right\|_2 \qquad \text{auf }]c, x]\,.$$

Nun gilt für $t \in\,]c, x]$:

$$\begin{aligned}
\left| \|z\|_2'(t) \right| &\leqslant \left\| z'(t) \right\|_2 = \left\| y'(t) - \hat{y}'(t) \right\|_2 \\
&\leqslant \varepsilon + \hat{\varepsilon} + \left\| f\left(t, y(t)\right) - f\left(t, \hat{y}(t)\right) \right\|_2 \\
&\leqslant \varepsilon + \hat{\varepsilon} + C \left\| y(t) - \hat{y}(t) \right\|_2 \\
&= \varepsilon + \hat{\varepsilon} + C\,\|z(t)\|_2\,,
\end{aligned}$$

also gilt für $c < d < x$ wegen $z(t) \neq 0$ für $d \leqslant t \leqslant x$:

$$\left| \int_d^x \frac{\|z\|_2'(t)}{\varepsilon + \hat{\varepsilon} + C\,\|z(t)\|_2}\,dt \right| \leqslant \int_d^x \frac{\left| \|z\|_2'(t) \right|}{\varepsilon + \hat{\varepsilon} + C\,\|z(t)\|_2}\,dt \leqslant \int_d^x 1\,dt = x - d,$$

und somit

$$\frac{1}{C}\left|\log\frac{\varepsilon+\hat{\varepsilon}+C\,\|z(x)\|_2}{\varepsilon+\hat{\varepsilon}+C\,\|z(d)\|_2}\right|\leqslant x-d,$$

d. h.

$$\varepsilon+\hat{\varepsilon}+C\,\|z(x)\|_2\leqslant(\varepsilon+\hat{\varepsilon}+C\,\|z(d)\|_2)\,e^{C(x-d)},$$

d. h. für $d\downarrow x$:

$$\varepsilon+\hat{\varepsilon}+C\,\|z(x)\|_2\leqslant(\varepsilon+\hat{\varepsilon}+C\,\|z(c)\|_2)\,e^{C(x-c)}\leqslant\left(\varepsilon+\hat{\varepsilon}+C\,\|z(x_0)\|_2\right)e^{C(x-x_0)}$$

Das aber bedeutet:

$$\begin{aligned}\|z(x)\|_2&\leqslant\frac{\varepsilon+\hat{\varepsilon}}{C}\left(e^{C(x-x_0)}-1\right)+\|z(x_0)\|_2\,e^{C(x-x_0)}\\&\leqslant\frac{\varepsilon+\hat{\varepsilon}}{C}\left(e^{C(x-x_0)}-1\right)+\left((K+\hat{\varepsilon})\,|x-\hat{x}_0|+\left\|y_0-\hat{y}_0\right\|_2\right)e^{C(x-x_0)}\end{aligned}$$

Das gilt also für alle $x\in I$, $x>x_0$. Analog für $x<x_0$. □

2.5.2 Beispiel (Differentialgleichung des mathematischen Pendels) In Beispiel 0.2.7 hatten wir die Differentialgleichung

$$\ddot{\varphi}=-\frac{g}{\ell}\sin\varphi=-\omega^2\sin\varphi\qquad\left(\omega:=\sqrt{\frac{g}{\ell}}\right)$$

kennen gelernt. Für kleine φ approximiert man $\sin\varphi\approx\varphi$ und löst statt der wahren Differentialgleichung die Näherungsgleichung $\ddot{\psi}=-\omega^2\psi$, die eine harmonische Schwingung beschreibt. Wir groß ist der Fehler in den Lösungen? Wir nehmen als Anfangsbedingung $t_0=0$, $\varphi(0)=\psi(0)=0$, $\dot{\varphi}(0)=\dot{\psi}(0)=\eta_0$ und schreiben um auf Systeme:

$$\begin{aligned}y&=\begin{pmatrix}\varphi\\\dot{\varphi}\end{pmatrix}=\begin{pmatrix}y_1\\y_2\end{pmatrix},\qquad\hat{y}=\begin{pmatrix}\psi\\\dot{\psi}\end{pmatrix}=\begin{pmatrix}\hat{y}_1\\\hat{y}_2\end{pmatrix},\\y'&=\begin{pmatrix}\dot{\varphi}\\\ddot{\varphi}\end{pmatrix}=\begin{pmatrix}y_2\\-\omega^2\sin y_1\end{pmatrix}=f\left(t,y\right),\qquad\varepsilon=0,\\\hat{y}'&=\begin{pmatrix}\dot{\psi}\\\ddot{\psi}\end{pmatrix}=\begin{pmatrix}\hat{y}_2\\-\omega^2\hat{y}_1\end{pmatrix},\\\left\|\hat{y}'-f\left(t,\hat{y}\right)\right\|_2&=\left|\left(-\omega^2\hat{y}_1+\omega^2\sin\hat{y}_1\right)^2\right|^{\frac{1}{2}}=\omega^2\left|\hat{y}_1-\sin\hat{y}_1\right|.\end{aligned}$$

Wir kennen aber $\hat{y}$: $\psi(t)=\frac{\eta_0}{\omega}\sin\omega t$ für $t\in\mathbb{R}$, d. h. $\hat{y}_1(t)=\frac{\eta_0}{\omega}\sin\omega t$,

$$\begin{aligned}\left\|\hat{y}'(t)-f\left(t,\hat{y}(t)\right)\right\|_2&=\omega^2\left|\sin\left(\frac{\eta_0}{\omega}\sin\omega t\right)-\frac{\eta_0}{\omega}\sin\omega t\right|\\&\leqslant\omega^2\frac{1}{3!}\left|\frac{\eta_0}{\omega}\sin\omega t\right|^3\leqslant\frac{1}{6}\frac{\left|\eta_0\right|^3}{\omega}=:\hat{\varepsilon},\end{aligned}$$

falls $\left|\frac{\eta_0}{\omega}\right| \leqslant 1,\ t \in \mathbb{R}$, denn eine alternierende Reihe mit monoton abnehmenden Gliedern ist betragsmäßig höchstens so groß wie der Betrag des ersten Terms.
Lipschitz-Bedingung:

$$\begin{aligned}
\left\| f(t,u) - f(t,v) \right\|_2 &= \left\| \begin{pmatrix} u_2 \\ -\omega^2 \sin u_1 \end{pmatrix} - \begin{pmatrix} v_2 \\ -\omega^2 \sin v_1 \end{pmatrix} \right\|_2 \\
&= \left(\omega^4 \underbrace{(\sin u_1 - \sin v_1)}_{=(u_1 - v_1)\cos\xi}{}^2 + (u_2 - v_2)^2 \right)^{\frac{1}{2}} \\
&\leqslant \left(\omega^4 (u_1 - v_1)^2 + (u_2 - v_2)^2 \right)^{\frac{1}{2}} \\
&\leqslant C \left\| u - v \right\|_2
\end{aligned}$$

mit $C := \max\left(1, \omega^2\right)$. Wir befinden uns im Spezialfall Satz 2.5.1 a), so dass gilt:

$$\left\| y(t) - \hat{y}(t) \right\|_2 \leqslant \frac{1}{6} \frac{\left|\eta_0\right|^3}{\omega} \frac{1}{C} \left(e^{C|t|} - 1 \right) \leqslant \frac{1}{6} \frac{\left|\eta_0\right|^3}{\omega} |t|\, e^{C|t|},$$

denn nach dem Mittelwertsatz ist $e^x - 1 \leqslant x e^x$ für $x \geqslant 0$.
Ergebnis: Für $\left|\frac{\eta_0}{\omega}\right| \leqslant 1$ gilt mit $C := \max\left(1, \omega^2\right)$:

$$\left| \varphi(t) - \frac{\eta_0}{\omega} \sin \omega t \right| \leqslant \frac{1}{6} \frac{\left|\eta_0\right|^3}{\omega} |t|\, e^{C|t|} \qquad (t \in \mathbb{R}).$$

Die Bewegung des Pendels ist also (wenigstens für kleine Zeiten bzw. kleine Werte von η_0, also kleine Ausschläge, näherungsweise harmonisch.

Nicht nur in Abhängigkeit von den Anfangswerten sind die Lösungen des Anfangswertproblems $y' = f\left(x, y\right)$ mit $y(x_0) = y_0$ stetig. Man kann sich vorstellen, dass auch f noch stetig von weiteren Parametern abhängt. Beschreibt die Gleichung z. B. die Bewegung eines Systems von Massenpunkten, so werden die Massen solche Parameter sein. In der Praxis sind sie nur bis auf unvermeidliche Messfehler genau bekannt, und für den Mathematiker ergibt sich die nahe liegende Aufgabe zu zeigen, dass kleine Änderungen in den Parametern nur zu kleinen Änderungen in der Lösung des Anfangswertproblems führen. Das ist in der Tat der Fall:

2.5.3 Satz *Sei in der Situation von Satz 2.5.1 zusätzlich die Funktion f stetig abhängig vom Parameter λ für $\|\lambda - \lambda_0\|_2 < c$ $(c > 0)$, und gelte die Lipschitz-Bedingung bzgl. y gleichmäßig in bezug auf λ, d. h.*

$$\left\| f(x,u,\lambda) - f(x,v,\lambda) \right\|_2 \leqslant C \left\| u - v \right\|_2$$

für alle $(x,u), (x,v) \in G$, $\|\lambda - \lambda_0\| < c$. Dann hängt die Lösung $y_\lambda\left(\cdot; x_0, y_0\right)$ des Anfangswertproblems $y' = f\left(x, y, \lambda\right)$, $y(x_0) = y_0$ stetig ab von $\left(x_0, y_0, \lambda\right)$.

Der Beweis verläuft ähnlich wie bei Satz 2.5.1 und kann nachgelesen werden bei Coddington und Levinson (1998) in Theorem 7.4 auf S. 29. Dort wird sogar eine präzisere und schärfere Aussage bewiesen als Satz 2.5.3. Der Beweis stützt sich aber nicht auf die Methode von Satz 2.5.1 sondern benutzt die Methode der sukzessiven Approximation. Noch etwas allgemeiner steht die Sache bei Walter (1976a) auf S. 93 als „Satz über stetige Abhängigkeit“, und ein ähnlicher Satz steht bei Heuser (1995a) auf S. 144 als Satz 13.1. Bei Knobloch und Kappel (1974a) steht im wesentlichen unser Spezialfall b) von Satz 2.5.1, allerdings in einem Punkt wesentlich allgemeiner und daher sehr technisch.

Neben der stetigen Abhängigkeit von den Anfangswerten interessiert auch die Frage nach der Differenzierbarkeit der Lösungen nach den Anfangswerten und entsprechend die Frage nach der differenzierbaren Abhängigkeit von eventuellen weiteren Parametern. Dabei wird man natürlich annehmen, dass f nicht nur stetig ist und einer Lipschitz-Bedingung genügt, sondern darüber hinaus auch hinreichend oft differenzierbar ist.

2.5.4 Satz *Es seien $G \subset \mathbb{R}^{n+1}$ ein Gebiet und $f\colon G \to \mathbb{R}^n$, $(x, y) \mapsto f(x, y)$ sei stetig und bzgl. y stetig (partiell) differenzierbar. Dann genügt f lokal einer Lipschitz-Bedingung bzgl. y nach Satz 2.2.2, also gilt Satz 2.2.7. Für $(x_0, y_0) \in G$ sei $y(\cdot; x_0, y_0)$ die eindeutig bestimmte Lösung des Anfangswertproblems $y' = f(x, y)$, $y(x_0) = y_0$ mit maximalem Definitions-Intervall $I(x_0, y_0)$ und*

$$D := \left\{ \begin{pmatrix} x \\ x_0 \\ y_0 \end{pmatrix} : x \in I(x_0, y_0) \right\} \subset \mathbb{R}^{n+2}.$$

Dann ist $D \subset \mathbb{R}^{n+2}$ offen, und die Funktion

$$D \ni \begin{pmatrix} x \\ x_0 \\ y_0 \end{pmatrix} \mapsto y(x, x_0, y_0)$$

ist auf D stetig partiell differenzierbar.

Beweis Dies und noch mehr wird bewiesen bei Knobloch und Kappel (1974a), und zwar die Offenheit von D als Satz 3.1 auf S. 122, die Differenzierbarkeit und der Rest als Satz 3.2 auf S. 125. Vgl. auch Walter (1976d) und Arnol'd (1979). Satz 2.5.4 steht auch im wesentlichen bei Stepanov (1976) auf S. 291 bzw. S. 296 (höhere Ableitungen). Ein allgemeinerer Satz, der zusätzlich die Differenzierbarkeit nach Parametern beinhaltet, steht als Theorem 7.5 auf S. 30 bei Coddington und Levinson (1998). Ein einfacher Satz über Differenzierbarkeit nach einem Paramter steht bei Heuser (1995a), S. 145 und Stepanov (1976), S. 294. Ein allgemeiner Satz über differenzierbare Abhängigkeit von Parametern steht bei Walter (1976c). Eine sehr gute Darstellung des Themas gibt Knobloch und Kappel (1974b). □

2.6 Trennungs-, Vergleichs-, Oszillations- und Amplitudensatz

Ziel wird die genauere Untersuchung der Eigenschaften der Lösungen linearer Differentialgleichungen 2. Ordnung sein.

2.6.1 Beispiel Die Differentialgleichung $y'' + y = 0$ auf $\mathbb{R}$ hat das Hauptsystem $y_1(x) = \cos x$, $y_2(x) = \sin x$ für $x \in \mathbb{R}$. Diese Funktionen haben bekanntlich folgende Eigenschaften:

a) y_1, y_2 haben höchstens abzählbar viele Nullstellen. Jede Nullstelle ξ ist einfach (d. h. y_j und y_j' haben keine gemeinsame Nullstelle), und die Nullstellen häufen sich nicht in $\mathbb{R}$.

b) Die Nullstellen von y_1, y_2 „trennen sich", d. h.: zwischen je zwei aufeinanderfolgenden Nullstellen von y_1 liegt genau eine Nullstelle von y_2 und umgekehrt.

Dies ist ein Spezialfall eines allgemeinen Sachverhalts:

2.6.2 Satz (Sturmscher Trennungssatz) *benannt nach Charles Sturm, 1803–1855, schweizer Mathematiker.*
Es seien $a_0, a_1\colon I \to \mathbb{R}$ stetig. Dann gelten für die Lösungen der homogenen Differentialgleichung

$$y'' + a_1(x)\, y' + a_0(x)\, y = 0 \qquad \big(auf\, I\big) \tag{2.35}$$

folgende Aussagen:

a) Ist $y \neq 0$ eine Lösung von (2.35), so hat y höchstens abzählbar viele Nullstellen. Die Nullstellen sind sämtlich einfach und haben keinen Häufungspunkt in I.

b) Die Nullstellen zweier linear unabhängiger Lösungen y_1, y_2 von (2.35) trennen sich.

BEWEIS siehe Heuser (1995a), S. 329.

a) Sei $\xi \in I$ Nullstelle der Lösung $y\colon I \to \mathbb{R}$, $y \neq 0$. Dann ist $y'(\xi) \neq 0$, denn sonst wäre ja y Lösung des Anfangswertproblems bestehend aus der Differentialgleichung (2.35) mit $y(\xi) = y'(\xi) = 0$. Dieses Anfangswertproblem hat aber nur die triviale Lösung $y = 0$ nach Satz 2.4.2 a). Also ist $y'(\xi) \neq 0$, ξ eine einfache Nullstelle.
Angenommen, die Nullstellen häufen sich in $x_0 \in I$: Dann gibt es eine Folge von Nullstellen $\xi_n \in I$ mit $\xi_n \xrightarrow{n\to\infty} x_0$, $\xi_n \neq x_0$ für alle $n \in \mathbb{N}$. Dann ist $y(x_0) = \lim_{n\to\infty} y(\xi_n) = 0$, $y'(x_0) = \lim_{n\to\infty} \frac{y(\xi_n) - y(x_0)}{\xi_n - x_0} = 0$. Das aber ist ein Widerspruch, denn wegen $y \neq 0$ und $y(x_0) = 0$ ist $y'(x_0) \neq 0$.
Die Menge der Nullstellen ist abzählbar (ggf. endlich oder leer): Sei $K \subset I$ ein kompaktes Teilintervall. Da sich die Nullstellen nirgends in I häufen, hat jeder Punkt $a \in K$ eine offene Umgebung U_a, in der keine Nullstelle $\xi \neq a$ liegt. K ist kompakt, folglich gibt es $a_1, \ldots, a_m \in K$, so dass $K \subset \bigcup_{\nu=1}^{m} U_{a_\nu}$. K enthält daher höchstens endlich viele Nullstellen, denn ist $\xi \in K$ eine Nullstelle, so gilt: $\xi \in \{a_1, \ldots, a_m\}$. Damit enthält I höchstens abzählbar viele Nullstellen von y.

b) Sei nun y_1, y_2 ein Hauptsystem von (2.35), und seien ξ_1, ξ_2 zwei aufeinander folgende Nullstellen von y_2, $\xi_1 < \xi - 2$. Wir wissen:

$$W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix} \qquad (x \in I)$$

ist nullstellenfrei auf I. OBdA gelte $W(x) > 0$ für alle $x \in I$ (anderenfalls ersetze man $y_1 \mapsto -y_1$). Dann ist $W\left(\xi_j\right) = y_1\left(\xi_j\right) y_2'\left(\xi_j\right) > 0$ für $j = 1,2$. OBdA dürfen wir annehmen: $y_2'(\xi_1) > 0$ (sonst: $y_1 \mapsto -y_1$, $y_2 \mapsto -y_2$). y_2 ist stetig differenzierbar, es gibt also ein $\delta > 0$, so dass y_2 auf $[\xi_1 - \delta, \xi_1 + \delta] \cap I$ streng monoton wachsend ist, also $y_2(t) > 0$ für $t \in]\xi_1, \xi_1 + \delta] \cap I \neq \emptyset$. Wäre nun auch $y_2'(\xi_2) > 0$, so hätten wir die Existenz eines $\delta_1 > 0$, so dass y_2 auf $[\xi_2 - \delta_1, \xi_2 + \delta_1] \cap I$ streng monoton wachsend wäre, insbesondere also $y_2(s) < 0$ für $s \in [\xi_2 - \delta_1, \xi - 2[\cap I$. Wählen wir solche s, t, so haben wir $\xi_1 < t < s < \xi_2$, und nach dem Zwischenwertsatz liegt eine weitere Nullstelle von y_2 zwischen t und s: Das aber ist ein Widerspruch zur Voraussetzung, dass ξ_1 und ξ_2 aufeinander folgende Nullstellen seien. Demnach ist $y_2'(\xi_2) < 0$ (der Fall $y_2'(\xi_2) = 0$ scheidet nach a) aus). Nun haben wir $y_1(\xi_1)\, y_2'(\xi_1) > 0$, $y_2'(\xi_1) > 0$, also $y_1(\xi_1) > 0$ und ebenso $y_1(\xi - 2)\, y_2'(\xi_2) > 0$ und $y_2'(\xi_2) < 0$, also $y_1(\xi_2) < 0$. Nach dem Zwischenwertsatz liegt zwischen ξ_1 und ξ_2 mindestens eine Nullstelle x_1 von y_1. Gäbe es eine zweite Nullstelle $x_2 \neq x_1$ von y_1, $\xi_1 < x_2, x_1 < \xi_2$, so gäbe es (gleicher Schluss wie oben mit x_1, x_2 und y_1 an Stelle von ξ_1, ξ_2 und y_2) zwischen x_1 und x_2 eine Nullstelle von y_2, was zu einem Widerspruch führte. □

Im folgenden legen wir lineare Differentialgleilchungen zweiter Ordnung nicht in Normalform

$$y'' + a_1 y' + a_0 y = b \tag{2.36}$$

mit stetigen $a_0, a_1, b\colon I \to \mathbb{R}$ zugrunde sondern in etwas anderer Form:

2.6.3 Lemma (Äquivalente Umformulierung von (2.36)**)** Jede lineare Differentialgleichung zweiter Ordnung in der Normalform (2.36) ist äquivalent zu einer Differentialgleichung des Typs

$$\left(py'\right)' + qy = r \tag{2.37}$$

mit stetigen $p, q, r\colon I \to \mathbb{R}$, $p > 0$ stetig differenzierbar.

Beweis (2.37)→(2.36) (2.37) besagt: $py'' + p'y' + qy = r$, also wegen $p > 0$:

$$y'' + \frac{p'}{p} y' + \frac{q}{p} y = \frac{r}{p},$$

und hier sind alle Koeffizienten-Funktionen stetig, da p stetig differenzierbar ist.

(2.36)→(2.37) Gegeben sei (2.36), F eine Stammfunktion von a_1 (diese existiert, da a_1 stetig ist). Mit $p := e^F$, $q := a_0 e^F$, $r := be^F$ sind p, q, r stetig, $p > 0$ stetig differenzierbar und

$$\left(py'\right)' + qy - r = e^F y'' + e^F a_1 y' + e^F a_0 y - e^F b = e^F \left(y'' + a_1 y' + a_0 y - b\right),$$

d. h. wenn (2.36) gilt, so gilt (2.37) mit den angegebenen p, q, r. □

2.6.4 Satz (Lagrangesche Identität) *Es seien $p\colon I \to \mathbb{R}$ stetig differenzierbar, $p > 0$, $q\colon I \to \mathbb{R}$ stetig, $Lu := \left(pu'\right)' + qu$, falls $u\colon I \to \mathbb{R}$ zweimal differenzierbar ist. Sind $u, v\colon I \to \mathbb{R}$ zweimal differenzierbar, so gilt:*

$$uLv - vLu = \frac{d}{dx}\left(p\left(uv' - vu'\right)\right) = \frac{d}{dx}\left(p \cdot W(u,v)\right),$$

wobei $W(u,v) = \left|\begin{smallmatrix} u & v \\ u' & v' \end{smallmatrix}\right|$ die Wronskische Determinante von u und v bezeichne.

Beweis Nachrechnen:

$$\begin{aligned} uLv - vLu &= u\left(pv'' + p'v' + qv\right) - v\left(pu'' + p'u' + qu\right) \\ &= p\left(uv'' - vu''\right) + p'\left(uv' - vu'\right) \\ &= \frac{d}{dx}\left(p\left(uv' - vu'\right)\right). \end{aligned}$$

□

2.6.5 Satz (Sturmscher Vergleichssatz) *Es seien $p\colon I \to \mathbb{R}$ stetig differenzierbar, $p > 0$, $q_1, q_2\colon I \to \mathbb{R}$ stetig und $q_2 < q_1$ auf I. Ferner seien $u \neq 0$ eine Lösung von*

$$\left(pu'\right)' + q_1 u = 0 \qquad (\mathit{auf}\, I) \tag{2.38}$$

und $v \neq 0$ eine Lösung von

$$\left(pv'\right)' + q_2 v = 0 \qquad (\mathit{auf}\, I). \tag{2.39}$$

Dann liegt zwischen je zwei aufeinander folgenden Nullstellen von v (siehe Satz 2.6.2 a)) mindestens eine Nullstelle von u. Achtung: Es wird nicht behauptet, dass v überhaupt Nullstellen hat!

Beweis Es seien $x_1 < x_2$ zwei aufeinanderfolgende Nullstellen von v. OBdA dürfen wir annehmen: $v(x) > 0$ für $x_1 < x < x_2$. Angenommen, die Behauptung ist falsch. Dann ist $u(x) \neq 0$ für $x_1 < x < x_2$, also darf oBdA gleich angenommen werden, dass $u(x) > 0$ für $x_1 < x < x_2$. Nun multiplizieren wir (2.38) mit v und (2.39) mit u, setzen $Lw := \left(pw'\right)' + q_2 w$, und erhalten durch Subtraktion der ersten Gleichung von der zweiten:

$$\begin{aligned} \left(u\left(pv'\right)' + q_2 uv\right) - \left(v\left(pu'\right)' + q_1 uv\right) &= u\left(pv'\right)' - v\left(pu'\right)' + \left(q_2 - q_1\right) uv \\ &= uLv - vLu + \left(q_2 - q_1\right) uv \\ &\overset{\text{Satz 2.6.4}}{=} \frac{d}{dx}\left(p\left(uv' - vu'\right)\right) + \left(q_2 - q_1\right) uv = 0 \end{aligned}$$

Integration über $[x_1, x_2]$ liefert:

$$\int_{x_1}^{x_2} \underbrace{\left(q_1 - q_2\right)}_{>0 \text{ auf }]x_1, x_2[} uv\,dx = \left[p\left(uv' - vu'\right)\right]_{x_1}^{x_2} = p(x_2)\,u(x_2)\,v'(x_2) - p(x_1)\,u(x_1)\,v'(x_1),$$

da $v(x_1) = v(x_2) = 0$. Nach unseren Normierungen ist das Integral positiv, also:

$$p(x_2)\,u(x_2)\,v'(x_2) - p(x_1)\,u(x_1)\,v'(x_1) > 0. \tag{2.40}$$

Hier ist $v'(x_1) \neq 0 \neq v'(x_2)$, da nach Satz 2.6.2 a) die Nullstellen einfach sind, und wegen $v(x_1) = v(x_2) = 0$ und $v|_{]x_1,x_2[} > 0$ folgt: $v'(x_1) > 0$ und $v'(x_2) < 0$. Weiter ist $p(x_1) > 0$, $p(x_2) > 0$ und $u(x_1) \geqslant 0$, $u(x_2) \geqslant 0$. Daher ist

$$\underbrace{p(x_2)}_{>0}\,\underbrace{u(x_2)}_{\geqslant 0}\,\underbrace{v'(x_2)}_{<0} - \underbrace{p(x_1)}_{>0}\,\underbrace{u(x_1)}_{\geqslant 0}\,\underbrace{v'(x_1)}_{>0} \leqslant 0,$$

und das steht im Widerspruch zu (2.40). Folglich hat u eine Nullstelle in $]x_1, x_2[$. □

2.6.6 Beispiel Die Differentialgleichung $y'' + y = 0$ ($q_1 = 1$, $I = \mathbb{R}$) hat die nicht-trivialen Lösungen $u(x) = A\cos(x+\alpha)$ mit $A, \alpha \in \mathbb{R}$, $A \neq 0$, und diese Lösungen haben Nullstellen im Abstand π.
Für $0 < \omega < 1$ hat die Differentialgleichung $y'' + \omega^2 y = 0$ ($q_2 = \omega^2 < q_1$, $I = \mathbb{R}$) die nicht-trivialen Lösungen $v(x) = B\cos(\omega x + \beta)$ mit $B, \beta \in \mathbb{R}$, $B \neq 0$. Diese Lösungen haben Nullstellen im Abstand $\frac{\pi}{\omega} > \pi$. Also liegt zwischen je zwei aufeinander folgenden Nullstellen von v mindestens eine Nullstelle von u (evtl. liegen auch mehrere Nullstellen von u zwischen zwei aufeinander folgenden Nullstellen von v).
Für $\omega = 0$ hat die Differentialgleichung $y'' = 0$ ($q_2 = 0 < q_1$, $I = \mathbb{R}$) die nicht-trivialen Lösungen $v(x) = ax + b$ mit $a, b \in \mathbb{R}$, $(a,b) \neq (0,0)$, und keine dieser Lösungen hat zwei Nullstellen. Dann liefert Satz 2.6.5 keine Information. Ähnlich ist die Lage für $y'' - \omega^2 y = 0$ ($\omega > 0$, $q_2 = -\omega^2 < q_1$, $I = \mathbb{R}$). Hauptsystem ist $e^{\omega x}$, $e^{-\omega x}$, und eine nicht-triviale Lösung $v(x) = ae^{\omega x} + be^{-\omega x}$ mit $a, b \in \mathbb{R}$, $(a,b) \neq (0,0)$ hat für $a = 0$ oder $b = 0$ gar keine Nullstelle. Für $a \neq 0 \neq b$ kann $a > 0$ angenommen werden, und für $b < 0$ ist dann v wachsend und hat genau eine Nullstelle, d. h. auch hier liefert Satz 2.6.5 keine Information.

2.6.7 Beispiel (Die Nullstellen der Besselschen Funktionen) Die Differentialgleichung

$$x^2 y'' + xy' + \left(x^2 - \nu^2\right) y = 0 \tag{2.41}$$

heißt die *Besselsche Differentialgleichung*, benannt nach dem Astronomen, Mathematiker und Geodäten Friedrich Wilhelm Bessel, 1784–1846, der 1838 die Parallaxe von 61 Cygni bestimmt hat und damit einen fundamentalen Beweis der Richtigkeit des kopernikanischen Systems erbrachte. Normalerweise betrachtet man die Besselsche Differentialgleichung auf $\mathbb{C}$ und lässt auch komplexe ν zu. Die Lösungen sind genau bekannt; für sie gelten zahlreiche Formeln, die man bei Bedarf (neben vielen anderen) nachsehen kann bei Magnus u. a. (1966). Wir diskutieren (2.41) nur in $\mathbb{R}$ und setzen voraus:

$$\nu \geqslant 0, \quad x > 0. \tag{2.42}$$

Substituiert man

$$u := \sqrt{x} \cdot y, \tag{2.43}$$

so ist (2.41) äquivalent zu

$$u'' + \left(1 + \frac{1-4\nu^2}{4x^2}\right) u = 0. \tag{2.44}$$

Probe:

$$\begin{aligned} u &= x^{\frac{1}{2}} y, \\ u' &= \frac{1}{2} x^{-\frac{1}{2}} y + x^{\frac{1}{2}} y', \\ u'' &= -\frac{1}{4} x^{-\frac{3}{2}} y + x^{-\frac{1}{2}} y' + x^{\frac{1}{2}} y'', \end{aligned}$$

also

$$\begin{aligned} u'' + \left(1 + \frac{1-4\nu^2}{4x^2}\right) u &= x^{\frac{1}{2}} y'' + x^{-\frac{1}{2}} y' - \frac{1}{4} x^{-\frac{3}{2}} y + \frac{x^2-\nu^2}{x^2} x^{\frac{1}{2}} y + \frac{1}{4} x^{-\frac{3}{2}} y \\ &= x^{-\frac{3}{2}} \left(x^2 y'' + x y' + \left(x^2 - \nu^2\right) y\right). \end{aligned}$$

Da u und y dieselben Nullstellen haben, betrachten wir die Differentialgleichung (2.44), denn sie hat unsere Standardform (2.37), die wir in Satz Satz 2.6.5 zugrunde legen. Wir unterscheiden die drei Fälle: $0 \leqslant \nu < \frac{1}{2}$, $\nu = \frac{1}{2}$ und $\nu > \frac{1}{2}$.

$0 \leqslant \nu < \frac{1}{2}$: Es gilt: $q_2 := 1 < 1 + \frac{1-4\nu^2}{x^2} =: q_1$ für alle $x > 0$. Daher liegt der Vergleich mit der Differentialgleichung

$$v'' + v = 0 \qquad \left(q_2 = 1\right) \tag{2.45}$$

nahe: Nicht-triviale Lösungen von (2.45) sind die Funktionen $v(x) = \sin\left(x - \varphi\right)$ für $x \in \mathbb{R}$, wobei $\varphi \in \mathbb{R}$ eine beliebige Konstante ist, und diese haben die Nullstellen $\varphi + k\pi$ mit $k \in \mathbb{Z}$. Satz 2.6.5 liefert: Für $0 \leqslant \nu < \frac{1}{2}$ hat jede Lösung $y \neq 0$ der Differentialgleichung (2.41) in jedem Intervall der Form $0 < \varphi + k\pi < x < \varphi + (k+1)\pi$ für beliebiges $\varphi \in \mathbb{R}$ und $k \in \mathbb{Z}$ mindestens eine Nullstelle. D. h.: Jede Lösung $y \neq 0$ von (2.41) hat im Falle $0 \leqslant \nu < \frac{1}{2}$ unendlich viele Nullstellen, und aufeinander folgende Nullstellen haben immer einen Abstand $< \pi$. Heuser (1995a) gibt auf S. 332 an, dass der Abstand aufeinander folgender Nullstellen für $n \to \infty$ gegen π konvergiert.

$\nu = \frac{1}{2}$: Dann lautet (2.44) einfach: $u'' + u = 0$, d. h. die Differentialgleichung ist elementar und hat die nicht-trivialen Lösungen $u = A \sin\left(x - \varphi\right)$ mit $A \neq 0$ und $\varphi \in \mathbb{R}$. Im Falle $\nu = \frac{1}{2}$ haben aufeinander folgende Nullstellen jeder nicht-trivialen Lösung y von (2.41) den konstanten Abstand π.

$\nu > \frac{1}{2}$: Nun ist $q_2 = 1 + \frac{1-4\nu^2}{4x^2} < 1 = q_1$, und wir vergleichen (2.44) mit $v'' + v = 0$: Gegenüber dem ersten Fall sind jetzt aber die Rollen vertauscht, da $q_2 = 1 + \frac{1-4\nu^2}{4x^2}$, $q_1 = 1$. Satz 2.6.5 liefert: Zwischen je zwei aufeinander folgenden Nullstellen jeder nicht-trivialen Lösung y von (2.41) liegt mindestens eine Nullstelle von $v(x) = \sin\left(x - \varphi\right)$, also mindestens eine der Zahlen $\varphi + k\pi$ mit beliebigem $\varphi \in \mathbb{R}$ und $k \in \mathbb{Z}$. Sei nun $\varphi_0 \in \mathbb{R}$. In

$]\varphi_0, \varphi_0+\pi[$ hat dann $v(x) = \sin\bigl(x-\varphi_0\bigr)$ keine Nullstelle. Also kann y in $]\varphi_0, \varphi_0+\pi[$ keine zwei Nullstellen haben. Im Falle $\nu > \frac{1}{2}$ haben aufeinander folgende positive Nullstellen einer nicht-trivialen Lösung von (2.41) also stets einen Abstand $\geqslant \pi$. Die Existenz von Nullstellen im Fall $\nu > \frac{1}{2}$ wissen wir bisher nicht; sie folgt so: Sei $0 < \alpha < 1$ fest; man denke sich α dicht bei 1. Wegen $1 + \frac{1-4\nu^2}{4x^2} \xrightarrow{x\to\infty} 1$ gibt es zu α ein $x_0 > 0$, so dass $q_2 = \alpha^2 < 1 + \frac{1-4\nu^2}{4x^2} = q_1$ für alle $x \geqslant x_0$. Die Differentialgleichung

$$v'' + \alpha^2 v = 0 \tag{2.46}$$

bietet sich jetzt zum Vergleich an: Zwischen je zwei Nullstellen der nicht-trivialen Lösung $v(x) = \sin\bigl(\alpha x - \varphi\bigr)$ von (2.46) liegt mindestens eine Nullstelle jeder nicht-trivialen Lösung y von (2.41). Die Nullstellen sind hier die Zahlen $\frac{1}{\alpha}\bigl(\varphi + k\pi\bigr)$ für $k \in \mathbb{Z}$, diese haben den Abstand $\frac{\pi}{\alpha}$. Insbesondere existieren also für jedes solche y unendlich viele Nullstellen $0 < \xi_1 < \xi_2 < \ldots < \xi_n < \ldots$, wobei $\xi_n \xrightarrow{n\to\infty} \infty$, und zusammen mit dem obigen Resultat folgt für alle hinreichend großen n:

$$\pi \leqslant \xi_{n+1} - \xi_n \leqslant \frac{\pi}{\alpha}.$$

Da hier α beliebig dicht bei 1 liegen darf, folgt: Für $\nu > \frac{1}{2}$ hat jede nicht-triviale Lösung y von (2.41) unendlich viele Nullstellen $0 < \xi_1 < \xi_2 < \ldots < \xi_n < \ldots$, $\xi_n \xrightarrow{n\to\infty} \infty$, und es gilt: $\xi_{n+1} - \xi_n \xrightarrow{n\to\infty} \pi$.

In Satz 2.6.5 braucht die Lösung v gar keine Nullstellen zu haben, und dann liefert der Satz keine Aussage. Der folgende Satz 2.6.8 gibt ein hinreichendes Kriterium für die Existenz unendlich vieler Nullstellen:

2.6.8 Satz (Oszillationssatz) *Es seien $I = [a, \infty[$ und p, q wie in Satz 2.6.4, und es gelte: $q > 0$ auf I, $\int_a^\infty \frac{dx}{p(x)}$ und $\int_a^\infty q(x)\,dx$ divergieren. Dann ist jede nicht-triviale Lösung der Differentialgleichung $\bigl(pu'\bigr)' + qu = 0$ (auf I) oszillatorisch, d. h. hat unendlich viele Nullstellen. Nach Satz 2.6.2 bilden also die Nullstellen eine monoton wachsende Folge $a \leqslant \xi_1 < \xi_2 < \ldots < \xi_n < \ldots$, $\xi_n \xrightarrow{n\to\infty} \infty$, alle Nullstellen sind einfach, u oszilliert.*

Beweis Heuser (1995a), S. 334f. Wir setzen $v := pu'$; dann ist $v' = \bigl(pu'\bigr)' = -qu$, d. h.

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} \frac{1}{p}v \\ -qu \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{p} \\ -q & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \tag{2.47}$$

Für jedes $\xi \geqslant a$ ist

$$\begin{pmatrix} u(\xi) \\ v(\xi) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{2.48}$$

denn das lineare System (2.47) hat nur die triviale Lösung mit dem Anfangswert $\left(\begin{smallmatrix} u(\xi) \\ v(\xi) \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, und nach Voraussetzung ist $u \neq 0$. Wegen (2.48) ist

$$\begin{pmatrix} u(\xi) \\ v(\xi) \end{pmatrix} = r(x) \begin{pmatrix} \sin\vartheta(x) \\ \cos\vartheta(x) \end{pmatrix} \qquad (x \geqslant a) \tag{2.49}$$

mit den stetig differenzierbaren Funktionen $r(x) := \left(u^2(x) + v^2(x)\right)^{\frac{1}{2}}$ und $\vartheta(x)$ für $x \geqslant a$.

Behauptung $\vartheta(x) \xrightarrow{x\to\infty} \infty$. Dann hat $\sin\vartheta(x)$ unendlich viele Nullstellen und damit auch u.

Begründung Aus (2.49) folgt:

$$u' = r'\sin\vartheta + r\vartheta'\cos\vartheta = \frac{1}{p}\cos\vartheta,$$
$$v' = r'\cos\vartheta - r\vartheta'\sin\vartheta = -qr\sin\vartheta.$$

Hier wird die erste Gleichung mit $\cos\vartheta$, die zweite mit $-\sin\vartheta$ multipliziert und addiert:

$$\Rightarrow \vartheta' = \frac{1}{p}\cos^2\vartheta + q\sin^2\vartheta > 0. \tag{2.50}$$

$\Rightarrow \vartheta\colon [a,\infty[\to \mathbb{R}$ ist streng monoton wachsend. Die Behauptung wird bewiesen sein, wenn wir zeigen: ϑ ist unbeschränkt. Das zeigen wir indirekt: Wäre ϑ beschränkt, so existierten die Grenzwerte $\alpha := \lim_{x\to\infty}\cos^2\vartheta(x)$ und $\beta := \lim_{x\to\infty}\sin^2\vartheta(x)$, und es wäre $\alpha+\beta = 1$, insbesondere $\alpha > 0$ oder $\beta > 0$. Sei $x_0 > a$ so groß, dass $\cos^2\vartheta(x) \geqslant \frac{\alpha}{2}$, $\sin^2\vartheta(x) \geqslant \frac{\beta}{2}$ für alle $x \geqslant x_0$. Dann liefert (2.50) für alle $x \geqslant x_0$:

$$\begin{aligned}
\vartheta(x) - \vartheta(x_0) &= \int_{x_0}^{x} \vartheta'(x)\,dx \\
&= \int_{x_0}^{x} \left(\frac{1}{p(t)} \underbrace{\cos^2\vartheta(t)}_{\geqslant\frac{\alpha}{2}} + q(t) \underbrace{\sin^2\vartheta(t)}_{\geqslant\frac{\beta}{2}} \right) dt \\
&\geqslant \frac{\alpha}{2}\int_{x_0}^{x} \frac{dt}{p(t)} + \frac{\beta}{2}\int_{x_0}^{x} q(t)\,dt \xrightarrow{x\to\infty} \infty
\end{aligned}$$

nach der Voraussetzung des Satzes. Das aber ist ein Widerspruch, ϑ somit unbeschränkt. ✓

Wenn nun eine Lösung $u \neq 0$ von $(pu')' + qu = 0$ unendlich viele Nullstellen hat, wie hoch/tief sind dann die „Ausschläge" („Amplituden", d. h. die Maxima und Minima) zwischen den aufeinander folgenden Nullstellen? Darüber gibt Satz 2.6.9 Auskunft:

2.6.9 Satz (Amplitudensatz) *Es gelten die Voraussetzungen aus Satz 2.6.4, und zusätzlich seien q stetig differenzierbar und nullstellenfrei auf I, pq monoton (wachsend oder fallend) auf I. Dann gilt: Ist $u \neq 0$ eine Lösung der Differentialgleichung $Lu = (pu')' + qu = 0$, und sind $x_k < x_{k+1}$ zwei aufeinander folgende Extrema von u, so gilt: $|u(x_k)| \leqslant |u(x_{k+1})|$, falls $(pq)' \leqslant 0$, und $|u(x_k)| \geqslant |u(x_{k+1})|$, falls $(pq)' \geqslant 0$. D. h.: Das Wachsen bzw. Fallen der Amplituden von u ist dem von pq entgegengesetzt.*

Beweis siehe Heuser (1995a), S. 335, oder Walter (1976a), S. 193. Für $v := u^2 + \frac{1}{pq}\left(pu'\right)^2$ (Beachte: pu' ist differenzierbar, da $\left(pu'\right)' = -qu$) gilt:

$$v' = 2uu' + \underbrace{\frac{1}{pq}2\left(pu'\right)\underbrace{\left(pu'\right)'}_{=-qu}}_{=-2uu'} - \frac{\left(pq\right)'}{\left(pq\right)^2}\left(pu'\right)^2 = -\left(pq\right)'\left(\frac{u'}{q}\right)^2 \begin{cases} \geqslant 0, & \text{falls } \left(pq\right)' \leqslant 0, \\ \leqslant 0, & \text{falls } \left(pq\right)' \geqslant 0. \end{cases}$$

Wegen $v\left(x_k\right) = u^2\left(x_k\right)$ folgt:

$$\left|u\left(x_k\right)\right| \begin{cases} \leqslant \left|u\left(x_{k+1}\right)\right|, & \text{falls } \left(pq\right)' \leqslant 0, \\ \geqslant \left|u\left(x_{k+1}\right)\right|, & \text{falls } \left(pq\right)' \geqslant 0. \end{cases}$$

□

Zusatz Ist $\left(pq\right)' < 0$ bzw. > 0, so nehmen die Amplituden streng zu bzw. streng ab.

2.6.10 Beispiel Schon in Beispiel 2.6.7 haben wir gezeigt: Jede Lösung $y\colon\ ]0,\infty[\ \to \mathbb{R}$, $y \neq 0$ der Besselschen Differentialgleichung

$$x^2y'' + xy' + \left(x^2 - \nu^2\right)y = 0 \qquad (\nu \geqslant 0 \text{ fest}; x > 0) \tag{2.51}$$

hat unendlich viele Nullstellen. Wir haben sogar gesehen: Die Abstände aufeinander folgender Nullstellen konvergieren gegen π. Wir zeigen noch einmal die Existenz unendlich vieler Nullstellen mit Satz 2.6.8. Die Differentialgleichung (2.51) lautet:

$$\left(xy'\right)' + \left(x - \frac{\nu^2}{x}\right) = 0 \qquad (x > 0)\,,$$

d. h. $p = x$, $q = x - \frac{\nu^2}{x}$. Betrachten wir die Differentialgleichung nur für $x \geqslant \nu+1$, so ist klar: Auf $[\nu+1,\infty[$ ist $p > 0$, $q > 0$ und die Integrale $\int_{\nu+1}^{\infty}\frac{dx}{p(x)}$ und $\int_{\nu+1}^{\infty} q\,dx$ divergieren. Daher hat jede Lösung $y \neq 0$ von (2.51) auf $[\nu+1,\infty[$ unendlich viele Nullstellen. Weiter ist $\left(pq\right)(x) = x^2 - \nu^2$ streng monoton wachsend, q nullstellenfrei für $x > \nu$. Dann liefert Satz 2.6.9: Ist $y \neq 0$ eine Lösung von (2.51), und sind $\nu < x_k < x_{k+1}$ zwei aufeinander folgende Extrema von y, so gilt: $\left|y\left(x_k\right)\right| > \left|y\left(x_{k+1}\right)\right|$.

3 Rand- und Eigenwertprobleme

Inhaltsangabe

3.1 Randwertprobleme

Problemstellung: Vorgelegt sei die lineare Differentialgleichung zweiter Ordnung $y'' + a_1 y' + a_0 y = b$ auf dem kompakten Intervall $I = [a, b]$ mit stetigen $a_0, a_1, b \colon I \to \mathbb{R}$. Bisher haben wir für $x_0 \in I$, $y_0, y_1 \in \mathbb{R}$ das Anfangswertproblem $y(x_0) = y_0$, $y'(x_0) = y_1$ gelöst und gezeigt, dass dieses Anfangswertproblem stets eindeutig lösbar ist. Neben dem Anfangswertproblem sind in Mathematik, Naturwissenschaften und Technik besonders die Randwertprobleme von Bedeutung: Dabei ist die gesuchte Funktion besonderen Bedingungen in den *beiden* Endpunkten des Intervalls $[a, b]$ unterworfen. Ein Beispiel dafür ist eine schwingende Saite, die an beiden Enden aufgehängt ist (siehe dazu Heuser (1995a), S. 372–376).

Solche Randbedingungen sind z. B. die sog. Randbedingungen

erster Art: $y(a) = \eta_1,\ y(b) = \eta_2$;

zweiter Art: $y'(a) = \eta_1,\ y'(b) = \eta_2$;

dritter Art: $\alpha_1 y(a) + \alpha_2 y'(a) = \eta_1,\ \beta_1 y(b) + \beta_2 y'(b) = \eta_2$ mit $(\alpha_1, \alpha_2) \neq (0,0) \neq \left(\beta_1, \beta_2\right)$.

Diese Randbedingungen sind linear in y, aber nicht die allgemeinsten linearen Randbedingungen; man könnte auch Randbedingungen des Typs

$$\alpha_1 y(a) + \alpha_2 y'(a) + \alpha_3 y(b) + \alpha_4 y'(b) = \eta_1 \beta_1 y(a) + \beta_2 y'(a) + \beta_3 y(b) + \beta_4 y'(b) \quad = \eta_2$$

mit linear unabhängigen Vektoren $(\alpha_1, \ldots, \alpha_4)$, $\left(\beta_1, \ldots, \beta_4\right)$ betrachten. Natürlich kann man entsprechende Probleme für lineare Differentialgleichungen n-ter Ordnung betrachten; das geschieht z. B. bei Weise (1966).

Zur Vereinfachung betrachten wir nur den Fall von Differentialgleichungen zweiter Ordnung. Dieser Fall ist durchaus typisch und für Anwendungen besonders wichtig. Wir konzentrieren uns also auf die sog.

3.1.1 Problem (Sturmsche Randwertaufgabe) Gegeben seien im kompakten Intervall $I = [a,b]$ die Funktionen $p,q,g\colon I \to \mathbb{R}$, $p > 0$ stetig differenzierbar, q und g stetig. Gesucht sind die Lösungen y der Differentialgleichung

$$\begin{aligned} &\left(py'\right)' + qy = g \qquad (\text{auf } I), \quad \text{die den Randbedingungen} \\ &\begin{cases} R_a y := \alpha_1 y(a) + \alpha_2 y'(a) = \eta_1 \\ R_b y := \beta_1 y(b) + \beta_2 y'(b) = \eta_2 \end{cases} \end{aligned} \tag{3.1}$$

mit vorgelegten konstanten $(\alpha_1,\alpha_2) \neq (0,0) \neq \left(\beta_1,\beta_2\right)$, $\eta_1,\eta_2 \in \mathbb{R}$ genügen.

Warnung: Schon sehr einfache Randwertprobleme können unlösbar oder nicht eindeutig lösbar sein.

3.1.2 Beispiele a) $y'' + y = 0$, $I = [0,2\pi]$

Randbedingung I: $y(0) = 0$, $y(2\pi) = 1$. Jedes Lösung $y(x) = c_1 \cos x + c_2 \sin x$ der Differentialgleichung ist 2π-periodisch, die Randbedingung ist also niemals erfüllt. – Ändert man diese Randbedingung zu

Randbedingung II: $y(0) = 0$, $y(2\pi) = 0$, so heißt das: $c_1 = 0$. Das Randwertproblem hat also die unendlich vielen Lösungen $y_c(x) = c \sin x$ mit $c \in \mathbb{R}$.

Randbedingung III: $y(0) = 0$, $y'(2\pi) = 1$. Dann besagen diese Bedingungen für y: $c_1 = 0$, $c_2 = 1$, d. h. das Randwertproblem hat die eindeutige Lösung $y(x) = \sin x$ für $x \in I$.

b) $y'' = 0$, $I = [0,1]$; Lösungen $y(x) = \alpha x + \beta$ mit $\alpha,\beta \in \mathbb{R}$.

Randbedingung I: $y(0) = \eta{-}1$, $y(1) = \eta_2$ liefert das Gleichungssystem $\beta = \eta_1$, $\alpha + \beta = \eta_2$, und das ist eindeutig lösbar für alle η_1,η_2.

Randbedingung II: $y'(0) = \eta_1$, $y'(1) = \eta_2$ liefert das Gleichungssystem für α,β: $\alpha = \eta_1$, $\alpha = \eta_2$, und das ist für $\eta_1 \neq \eta_2$ gar nicht, für $\eta_1 = \eta_2$ nicht eindeutig lösbar.

3.1.3 Lemma (Alternative) Folgende Aussagen sind äquivalent:

a) Die Sturmsche Randwertaufgabe aus Problem 3.1.1 hat genau eine Lösung.

b) Die zugehörige homogene Sturmsche Randwertaufgabe

$$\left(py'\right)' + qy = 0, \qquad R_a y = R_b y = 0 \tag{3.2}$$

ist nur trivial lösbar.

c) Für jedes Hauptsystem y_1, y_2 von (3.1.3) gilt: $\begin{vmatrix} R_a y_1 & R_a y_2 \\ R_b y_1 & R_b y_2 \end{vmatrix} \neq 0$.

Beweis Sei y_0 irgendeine Lösung der inhomogenen Gleichung (3.1) und y_1, y_2 ein Hauptsystem von (3.1.3). Dann hat jede Lösung y von (3.1) die Form $y = y_0 + c_1 y_1 + c_2 y_2$. Nun gilt: Aussage a) ist gleichbedeutend mit der eindeutigen Lösbarkeit des linearen Gleichungssystems

$$\begin{aligned} R_a y &= R_a y_0 + c_1 R_a y_1 + c_2 R_a y_2 = \eta_1 \\ R_b y &= R_b y_0 + c_1 R_b y_1 + c_2 R_b y_2 = \eta_2 \end{aligned}$$

nach c_1, c_2, das wiederum besagt: $\begin{vmatrix} R_a y_1 & R_a y_2 \\ R_b y_1 & R_b y_2 \end{vmatrix} \neq 0$. $\Leftrightarrow$ Das homogene Randwertproblem (3.2) ist nur trivial lösbar. □

Frage: Wenn das homogene Randwertproblem (3.2) nur die triviale Lösung hat, wie findet man dann die Lösung des inhomogenen Problems (3.1)? Die Antwort erfolgt in zwei Schritten:

(A) Es genügt die Lösung für $\eta_1 = \eta_2 = 0$: siehe Satz 3.1.4

(B) Bestimmung der Lösung im Falle $\eta_1 = \eta_2 = 0$: siehe Satz 3.1.5

3.1.4 Satz *Unter den Voraussetzungen von Problem 3.1.1 sei das homogene Randwertproblem* (3.2) *nur trivial lösbar. Ferner sei $w\colon [a,b] \to \mathbb{R}$ die (nach Lemma 3.1.3 vorhandene eindeutig bestimmte) Lösung des halbhomogenen Problems*

$$\begin{aligned} &\big(pw'\big)' + qw = g - \Big(\big(pf'\big)' + qf\Big) \qquad \big(\mathit{auf}\, I = [a,b]\big), \\ &R_a w = R_b w = 0, \end{aligned}$$

wobei $f\colon [a,b] \to \mathbb{R}$ irgendeine zweimal stetig differenzierbare Funktion ist mit $R_a f = \eta_1$, $R_b f = \eta_2$ (man kann z. B. stets f als Polynom wählen). Dann ist $y := w + f$ die (nach Lemma 3.1.3 vorhandene eindeutig bestimmte) Lösung des Sturmschen Randwertproblems (3.1).

Beweis

$$\big(py'\big)' + qy = \underbrace{\big(pw'\big)' + qw}_{=g-((pf')'+qf)} + \big(pf'\big)' + qf = g,$$

und $R_a y = R_a w + R_a f = \eta_1$, $R_b y = R_b w + R_b f = \eta_2$. □

3.1.5 Satz *Vorgelegt sei das halbhomogene Sturmsche Randwertproblem*

$$\begin{cases} \big(py'\big)' + qy = g & \mathit{auf}\, I = [a,b] \\ R_a y = R_b y = 0 \end{cases} \tag{3.3}$$

mit den Voraussetzungen aus Problem 3.1.1, und das zugehörige homogene Randwertproblem (3.2) *sei nur trivial lösbar. Dann gibt es ein Hauptsystem y_1, y_2 der Differentialgleichung $\big(py'\big)' + qy = 0$, welches den Randbedingungen $R_a y_1 = 0$, $R_b y_2 = 0$ genügt. Mit y_1, y_2 bilde man die* Greensche Funktion *(benannt nach George Green, 1793–1841):*

$$G(x,t) := \begin{cases} \frac{y_1(x) y_2(t)}{p(a) W(a)} & a \leqslant x \leqslant t \leqslant b \\ \frac{y_1(t) y_2(x)}{p(a) W(a)} & a \leqslant t \leqslant x \leqslant b, \end{cases}$$

wobei $W(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{vmatrix}$ *die Wronskische Determinante sei. Dann ist* $p(x)\,W(x) = p(a)\,W(a)$ *für alle* $x \in [a,b]$, *da* $\left(pW\right)' = p'W + pW' = p'W + p\frac{-p'}{p}W = 0$. *Die eindeutig bestimmte Lösung* y *des Randwertproblems* (3.3) *lässt sich dann angeben in der Form*

$$y(x) = \int_a^b G(x,t)\,g(t)\,dt \qquad (x \in [a,b]).$$

Die Greensche Funktion hat folgende Eigenschaften:

(i) $G\colon [a,b] \times [a,b] \to \mathbb{R}$ *ist stetig und symmetrisch:* $G(x,t) = G(t,x)$ *für alle* $t, x \in [a,b]$.

(ii) Für $x \neq t$ *ist* G *nach* x *differenzierbar und erfüllt die Differentialgleichung*

$$\left(pG_x(\cdot,t)\right)' + qG(\cdot,t) = 0,$$

wobei G_x *die Ableitung nach dem ersten Argument bezeichne.*

(iii) Auf der Diagonalen $t = x$ *existieren noch die einseitigen Ableitungen von* G, *und es gilt:*

$$G_x(x+0,x) = \frac{y_1(x)\,y_2'(x)}{p(a)\,W(a)} \qquad \text{für } a \leqslant x < b,$$
$$G_x(x-0,x) = \frac{y_1'(x)\,y_2(x)}{p(a)\,W(a)} \qquad \text{für } a < x \leqslant b;$$

insbesondere gilt die Sprungrelation

$$G_x(x+0,x) - G_x(x-0,x) = \frac{1}{p(x)} \qquad \text{für } a < x < b.$$

Resultat: Ist G *bekannt (d. h. sind* y_1, y_2 *bekannt), so lässt sich für jedes* g *das Randwertproblem* (3.3) *lösen. Bestimmung von* y_1, y_2*: siehe Beweis.*

Beweis Die homogene Differentialgleichung (3.2) hat zwei eindeutig bestimmte Lösungen y_1, y_2, die den Anfangswertbedingungen $y_1(a) = \alpha_2$, $y_1'(a) = -\alpha_1$ bzw. $y_2(b) = \beta_2$, $y_2'(b) = -\beta_2$ mit den Koeffizienten $\alpha_1, \alpha_2, \beta_1, \beta_2$ aus (3.1) genügen. $\Rightarrow R_a y_1 = 0$, $R_b y_2 = 0$.

Behauptung y_1, y_2 bilden ein Hauptsystem von (3.2)

Begründung Wegen $(\alpha_1,\alpha_2) \neq (0,0) \neq \left(\beta_1,\beta_2\right)$ sind $y_1 \neq 0 \neq y_2$. Wären y_1, y_2 linear abhängig, so wäre $y_1 = \lambda y_2$ mit $\lambda \in \mathbb{R}$, $\lambda \neq 0$, und y_1 wäre Lösung des homogenen Problems (3.2). Das ist ein Widerspruch, denn nach Voraussetzung hat das homogene Problem nur die triviale Lösung, aber $y_1 \neq 0$. ✓

Sei nun $y\colon [a,b] \to \mathbb{R}$ die Lösung des halbhomogenen Problems (3.3). Dann lässt sich (Variation der Konstanten) y schreiben in der Form $y = c_1 y_1 + c_2 y_2$ mit stetig differenzierbaren Funktionen $c_1, c_2\colon [a,b] \to \mathbb{R}$. Wir wollen die Formeln aus Satz 2.4.4 und Beispiel 2.4.5 für c_1, c_2 benutzen und schreiben dazu unsere Differentialgleichung in der Form:

$$y'' + \frac{p'}{p}y' + \frac{q}{p}y = \frac{g}{p}$$

mit $a_0 = \frac{q}{p}$, $a_1 = \frac{p'}{p}$ und $b = \frac{g}{p}$. Dann ist nach Satz 2.4.4 und Beispiel 2.4.5 mit $x_0 = a$:

$$\begin{aligned}
c_1(x) &= c_{1,0} - \int_a^x \frac{y_2(t)}{W(t)} \underbrace{\frac{g(t)}{p(t)}}_{=b(t)} \, dt, \\
c_2(x) &= c_{2,0} + \int_a^x \frac{y_1(t)}{W(t)} \frac{g(t)}{p(t)} \, dt
\end{aligned}$$

mit

$$\begin{aligned}
W(x) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} &= W(x_0) \exp\left(- \int_{x_0}^x \underbrace{\frac{p'(t)}{p(t)}}_{=a_1(t)} \, dt \right) \\
&= W(x_0) \exp\left(- [\log p(t)]_{x_0}^x \right) = \frac{W(x_0)\, p(x_0)}{p(x)},
\end{aligned}$$

d. h. $p(x)\, W(x) = p(a)\, W(a)$ konstant für alle $x \in [a, b]$. Das setzen wir oben ein und erhalten:

$$\begin{aligned}
c_1(x) &= c_{1,0} - \frac{1}{p(a)\, W(a)} \int_a^x y_2(t)\, g(t)\, dt, \\
c_2(x) &= c_{2,0} - \frac{1}{p(a)\, W(a)} \int_a^x y_1(t)\, g(t)\, dt
\end{aligned} \tag{3.4}$$

Die Konstanten $c_{1,0}, c_{2,0}$ sind aus den Randbedingungen zu bestimmen. Wir wissen a priori, dass solche eindeutig bestimmten Konstanten existieren, welche y zur eindeutig bestimmten Lösung des Randwertproblems (3.3) machen. Nach (2.27) mit $b = \frac{g}{p}$ ist

$$\begin{aligned}
c_1' y_1 + c_2' y_2 &= 0 \\
c_1' y_1' + c_2' y_2' &= \frac{g}{p},
\end{aligned}$$

also

$$\begin{aligned}
y(a) &= c_1(a)\, y_1(a) + c_2(a)\, y_2(a) \\
y'(a) &= c_1(a)\, y_1'(a) + c_2(a)\, y_2'(a).
\end{aligned}$$

$$\Rightarrow R_a y = \alpha_1 y(a) + \alpha_2 y'(a) = c_1(a) \underbrace{R_a y_1}_{=0} + c_2(a)\, R_a y_2 = c_2(a)\, R_a y_2.$$

Hier ist $R_a y_2 \neq 0$, denn wäre $R_a y_2 = 0$, so hätten wir

$$\begin{aligned}
\alpha_1 y_1(a) + \alpha_2 y_1'(a) &= 0, \\
\alpha_1 y_2(a) + \alpha_2 y_2'(a) &= 0,
\end{aligned} \tag{3.5}$$

d. h. $(\alpha_1, \alpha_2) \neq (0, 0)$ wäre Lösung des homogenen linearen Gleichungssystems (3.5). Aber (3.5) hat die Determinante $W(a) \neq 0$, das ist ein Widerspruch.

$\Rightarrow R_a y = c_2(a) R_a y_2$ mit $R_a y_2 \neq 0$, und die Randbedingung $R_a y = 0$ ist gleichbedeutend mit $c_2(a) = 0$. Analog bei b: $R_b y = 0 \Leftrightarrow c_1(b) = 0$. Das setzen wir in (3.4) ein und erhalten:

$$c_{1,0} = \frac{1}{p(a)W(a)} \int_a^b y_2(t)\, g(t)\, dt, \quad \text{d. h.} \qquad c_1(x) = \frac{1}{p(a)W(a)} \int_x^b y_2(t)\, g(t)\, dt;$$

$$c_{2,0} = 0 \quad \text{d. h.} \qquad c_2(x) = \frac{1}{p(a)W(a)} \int_a^x y_1(t)\, g(t)\, dt.$$

Die Lösung y von (3.3) ist also gegeben durch

$$y(x) = c_1(x)\, y_1(x) + c_2(x)\, y_2(x) = \int_a^b G(x,t)\, g(t)\, dt$$

mit der Greenschen Funktion

$$G(x,t) := \begin{cases} \frac{y_1(x) y_2(t)}{p(a) W(a)} & a \leqslant x \leqslant t \leqslant b \\ \frac{y_1(t) y_2(x)}{p(a) W(a)} & a \leqslant t \leqslant x \leqslant b. \end{cases}$$

Die Greensche Funktion hat die in Satz 3.1.5 genannten Eigenschaften:

(i) $G(x,t) = G(t,x)$ für alle $t, x \in [a,b]$ und die Stetigkeit von G sind klar.

(ii) Für $x \neq t$ ist $G(\cdot,t)$ nach x differenzierbar und erfüllt die Differentialgleichung

$$\big(pG_x(\cdot,t)\big)' + qG(\cdot,t) = 0.$$

(iii) Auf der Diagonalen $t = x$ existieren noch die einseitigen Ableitungen von G, und es gilt:

$$G_x(x+0,x) = \frac{y_1(x)\, y_2'(x)}{p(a)\, W(a)} \qquad \text{für } a \leqslant x < b,$$

$$G_x(x-0,x) = \frac{y_1'(x)\, y_2(x)}{p(a)\, W(a)} \qquad \text{für } a < x \leqslant b,$$

und es gilt die Sprungrelation

$$G_x(x+0,x) - G_x(x-0,x) = \frac{1}{p(x)} \qquad \text{für } a < x < b,$$

denn

$$y_1(x)\, y_2'(x) - y_1'(x)\, y_2(x) = W(x) = \frac{p(a)\, W(a)}{p(x)}. \qquad \square$$

Zusatz Als Funktion von x genügt G auch den Randbedingungen $R_a G(\cdot,t) = R_b G(\cdot,t) = 0$. Das folgt sofort aus den Randbedingungen $R_a y_1 = 0$ und $R_b y_2 = 0$.

Frage: Wenn das homogene Randwertproblem (3.2) eine Lösung $y \neq 0$ hat, unter welchen Bedingungen ist dann das halbhomogene Problem (3.3) lösbar? Die Antwort gibt der folgende

3.1.6 Satz *Wenn das homogene Randwertproblem* (3.2) *eine Lösung* $y_1 \neq 0$ *hat*[1]*, so ist das halbhomogene Problem* (3.3) *genau dann lösbar, wenn*

$$\int_a^b g(t)\, y_1(t)\, dt = 0. \tag{3.6}$$

Bemerkung Diese Lösung kann wegen (3.2) nicht eindeutig bestimmt sein.

Beweis Wir setzen $Ly := \left(py'\right)' + qy$ und haben die Lagrangesche Identität aus Satz 2.6.4:

$$vLu - uLv = \left(p\left(vu' - uv'\right)\right)'.$$

„⇒" Sei (3.3) lösbar mit Lösung y und $y_1 \neq 0$ eine Lösung von (3.2). Dann ist

$$\int_a^b \underbrace{L(y)}_{=g} y_1 - \underbrace{L(y_1)}_{=0} y\, dt = \int_a^b g(t)\, y_1(t)\, dt$$

$$\overset{\text{Satz 2.6.4}}{=} \left[p(t)\left(y_1(t)\, y'(t) - y_1'(t)\, y(t)\right)\right]_a^b = \left[p(t) \begin{vmatrix} y_1(t) & y(t) \\ y_1'(t) & y'(t) \end{vmatrix}\right]_a^b.$$

Ist nun $\alpha_1 = 0$ in (3.1), so ist $\alpha_2 \neq 0$, d. h. $y'(a) = y_1'(a) = 0$, also

$$\begin{vmatrix} y_1(a) & y(a) \\ y_1'(a) & y'(a) \end{vmatrix} = 0.$$

Im Falle $\alpha_1 \neq 0$ ist

$$\begin{vmatrix} y_1(a) & y(a) \\ y_1'(a) & y'(a) \end{vmatrix} = \frac{1}{\alpha_1} \begin{vmatrix} \alpha_1 y_1(a) + \alpha_2 y_1'(a) & \alpha_1 y(a) + \alpha_2 y'(a) \\ y_1'(a) & y'(a) \end{vmatrix} = \frac{1}{\alpha_1} \begin{vmatrix} R_a y_1 & R_a y \\ y_1'(a) & y'(a) \end{vmatrix} = 0.$$

Gleicher Schluss bei b liefert (3.6).

„⇐" Gelte (3.6), und y_0 sei Lösung von $Ly_0 = g$. Weiter sei y_2 Lösung von (3.2), so dass y_1, y_2 ein Hauptsystem von (3.2) bilden. Nach der Fußnote auf Seite 113 genügt dann y_2 nicht den Randbedingungen in (3.2). Schärfer: Dann ist $R_a y_2 \neq 0$: Wäre nämlich $R_a y_2 = 0$, so wäre

$$\alpha_1 y_1(a) + \alpha_2 y_1'(a) = 0,$$
$$\alpha_1 y_2(a) + \alpha_2 y_2'(a) = 0,$$

[1] Der Lösungsraum von (3.2)ist dann eindimensional, der Lösungsraum von (3.2) ohne Randbedingungen zweidimensional, und es gibt eineLösung mit $R_a y \neq 0$ wegen der Lösbarkeit des Anfangswertproblems.

und dieses lineare Gleichungssystem hätte die Lösung $(\alpha_1, \alpha_2) \neq (0,0)$. Das aber ist ein Widerspruch, denn

$$\begin{vmatrix} y_1(a) & y_2(a) \\ y_1'(a) & y_2'(a) \end{vmatrix} \neq 0.$$

Also: $R_a y_2 \neq 0$. Wir setzen nun eine Lösung von (3.3) an in der Form $y = y_0 + c y_2$, also $Ly = g$, wobei $c \in \mathbb{R}$ so bestimmt sei, dass $R_a y = R_a y_0 + c R_a y_2 = 0$. Dadurch ist c eindeutig bestimmt. Nun haben wir

$$\int_a^b \underbrace{L(y)}_{=g} y_1 - \underbrace{L(y_1)}_{=0} y \, dt = \int_a^b g(t)\, y_1(t)\, dt = 0$$

nach Voraussetzung; andererseits ist nach der Lagrangeschen Identität die linke Seite gleich

$$\left[p(t) \begin{vmatrix} y_1(t) & y(t) \\ y_1'(t) & y'(t) \end{vmatrix} \right]_a^b = p(b) \begin{vmatrix} y_1(b) & y(b) \\ y_1'(b) & y'(b) \end{vmatrix} - \underbrace{p(a) \begin{vmatrix} y_1(a) & y(a) \\ y_1'(a) & y'(a) \end{vmatrix}}_{=0 \text{ nachWahl von } y \text{ (s.o.)}}$$

$$\Rightarrow \begin{vmatrix} y_1(b) & y(b) \\ y_1'(b) & y'(b) \end{vmatrix} = 0.$$

Hier ist $\binom{y_1(b)}{y_1'(b)} \neq \binom{0}{0}$, da $y_1 \neq 0$ wegen der eindeutigen Lösbarkeit des Anfangswertproblems. Es gibt also ein $\lambda \in \mathbb{R}$, so dass

$$\begin{pmatrix} y(b) \\ y'(b) \end{pmatrix} = \lambda \begin{pmatrix} y_1(b) \\ y_1'(b) \end{pmatrix}.$$

$\Rightarrow y = \lambda y_1$, denn sowohl y als auch λy_1 lösen das Anfangswertproblem mit den Anfangswerten $\lambda y_1(b)$ und $\lambda y_1'(b)$. $\Rightarrow R_b y = \lambda R_b y_1 = 0$. y ist also Lösung von (3.3). □

Nebenbei haben wir im Verlauf des Beweises von Satz 3.1.6 bewiesen:

3.1.7 Lemma Sind u, v Lösungen von (3.2) mit $R_a u = R_a v = 0$, so gilt:

$$\begin{vmatrix} u(a) & v(a) \\ u'(a) & v'(a) \end{vmatrix} = 0;$$

entsprechend bei b.

3.2 Das Sturm-Liouvillesche Eigenwertproblem

3.2.1 Beispiel Die Differentialgleichung der schwingenden Saite führt auf folgendes Problem: Gesucht sind die $\lambda \in \mathbb{R}$, für welche die Differentialgleichung $y'' + \lambda y = 0$ auf $I = [0, \pi]$ unter den Randbedingungen $y(0) = y(\pi) = 0$ eine Lösung $y \neq 0$ hat. D. h. wir interessieren uns für die Fälle, in denen das Randwertproblem keine eindeutig bestimmte Lösung hat. Ist $y \neq 0$ eine Lösung dieses Problems, so ist

$$\lambda \int_0^\pi y^2(x)\,dx = -\int_0^\pi y''(x)\,y(x)\,dx = \underbrace{\left[-y'(x)\,y(x)\right]_0^\pi}_{=0 \text{ nach RB}} + \int_0^\pi \left(y'(x)\right)^2 dx,$$

also $\lambda \geqslant 0$, und wegen der Randbedingungen ist notwendig $\lambda > 0$. Dann hat die Differentialgleichung nach Beispiel 0.2.6 genau die Lösungen $y(x) = \alpha\cos\sqrt{\lambda}x + \beta\sin\sqrt{\lambda}x$. Die Randbedingungen liefern: $y(0) = 0 \Leftrightarrow \alpha = 0$ und $y(0) = y(\pi) = 0 \Leftrightarrow y(x) = \beta\sin\sqrt{\lambda}y$, wobei $\beta\sin\sqrt{\lambda}\pi = 0$. Hier gilt wegen $\lambda > 0$:

$$\sin\sqrt{\lambda}\pi = 0 \Leftrightarrow \sqrt{\lambda} \in \mathbb{N} \Leftrightarrow \exists\, n \in \mathbb{N} \quad \lambda = n^2 \in \mathbb{N}.$$

Ergebnis: Das Randwertproblem $y'' + \lambda y = 0$, $y(0) = y(\pi) = 0$ ist nur für die Zahlen $\lambda = n^2$ mit $n \in \mathbb{N}$ nicht-trivial lösbar, und dann sind $y_n(x) = \sin nx$ mit $n \in \mathbb{N}$ bis auf konstante Faktoren genau alle Lösungen. Die Eigenfunktionen y_n sind orthogonal in bezug auf das Skalarprodukt

$$\langle f, g\rangle := \int_0^\pi f g\,dx$$

auf $C[0,\pi]$ oder besser $L^2(0,\pi)$, denn

$$\begin{aligned}\int_0^\pi \sin mx \sin nx\,dx &= \int_0^\pi \frac{1}{2}\left(\cos(m-n)x - \cos(m+n)x\right)dt \\ &= \begin{cases} \frac{\pi}{2} - \frac{1}{4n}\left[\sin 2nx\right]_0^\pi = \frac{\pi}{2} & \text{für } m = n, \\ \left[\frac{1}{2}\left(\frac{1}{m-n}\sin(m-n)x - \frac{1}{m+n}\sin(m+n)x\right)\right]_0^\pi = 0 & \text{für } m \neq n,\end{cases}\end{aligned}$$

d. h.

$$\left\langle\sqrt{\frac{2}{\pi}}y_m, \sqrt{\frac{2}{\pi}}y_n\right\rangle = \delta_{m,n}.$$

Frage: Kann man jeden Ton der Saite durch Überlagerung harmonischer Schwingungen herstellen? Das führt zu Fourier-Reihen.

Das Randwertproblem der schwingenden Saite ist ein Beispiel für ein reguläres Sturm-Liouvillesches Eigenwertproblem, das wir in Problem 3.2.2 formulieren. Das Problem ist benannt nach Charles Sturm (1803-1855), einem Schweizer Mathematiker, der ab 1830 in Paris lebte und dort gemeinsam mit seinem französischen Freund und Mathematiker-Kollegen Joseph Liouville (1809-1882) die Theorie des sog. regulären Sturm-Liouvilleschen Eigenwertproblems schuf. Die Theorie hat im 20. Jh. durch Hermann Weyl, Edward Charles Titchmarsh und Kunihiko Kodaira erhebliche Ausdehnungen erfahren. Wir beschäftigeun uns nur mit einem relativ einfachen Spezialfall (kompaktes Intervall, Differentialgleichung der Ordnung 2, „glatte" Koeffizientenfunktionen). Im folgenden sei stets $I = [a, b]$ ein kompaktes Intervall.

3.2.2 Problem (Reguläres Sturm-Liouvillesches Eigenwertproblem) Auf dem kompakten Intervall $[a,b]$ seien die Funktionen

$$p,q,r\colon I \to \mathbb{R} \quad \text{alle stetig, } p \text{ stetig differenzierbar, } p>0,\ r>0. \tag{3.7}$$

Für $u \in C^2(I)$ sei

$$Lu := \left(pu'\right)' + qu. \tag{3.8}$$

Ferner seien $\alpha_1,\ldots,\beta_2 \in \mathbb{R}$, $(\alpha_1,\alpha_2) \neq (0,0) \neq \left(\beta_1,\beta_2\right)$, und für $u \in C^1(I)$ seien

$$\begin{aligned} R_a u &:= \alpha_1 u(a) + \alpha_2 u'(a) \\ R_b u &:= \beta_1 u(b) + \beta_2 u'(b). \end{aligned} \tag{3.9}$$

Gesucht sind die $\lambda \in \mathbb{R}$ und $y \in C^2(I)$, für welche das homogene Randwertproblem

$$\begin{aligned} Ly + \lambda r y &= 0 \\ R_a y + R_b y &= 0 \end{aligned} \tag{3.10}$$

eine Lösung $y \neq 0$ hat. D.h.: Gesucht sind die Eigenwerte λ und Eigenfunktionen y des Differentialoperators $-\frac{1}{r}L$, $-\frac{1}{r}Ly = \lambda y$ unter den Randbedingungen (3.10).

3.2.3 Folgerung a) Nach der Alternative Lemma 3.1.3 gilt: λ ist ein Eigenwert

$\Longleftrightarrow$ Das homogene Randwertproblem $Ly + \lambda r y = 0$, $R_a y + R_b y = 0$ hat eine Lösung $y \neq 0$

$\Longleftrightarrow$ Das Randwertproblem $Ly + \lambda r y = g$, $R_a y = R_b y = 0$ ist für kein $g \in C(I)$ eindeutig lösbar.

$\Longleftrightarrow$ Für jedes Hauptsystem y_1, y_2 von $Ly + \lambda r y = 0$ gilt:

$$\begin{vmatrix} R_a y_1 & R_a y_2 \\ R_b y_1 & R_b y_2 \end{vmatrix} = 0.$$

Frage: Gibt es stets Eigenwerte und Eigenfunktionen? Die Antwort liefert folgender

3.2.4 Satz (Existenzsatz) *Das reguläre Sturm-Liouvillesche Eigenwertproblem* (3.2.4) *hat eine streng wachsende Folge* $(\lambda_n)_{n\geqslant 0}$,

$$\lambda_0 < \lambda_1 < \lambda_2 < \ldots, \quad \lambda_n \xrightarrow{n\to\infty} \infty$$

von Eigenwerten. Jeder Eigenwert ist einfach, d.h. der zugehörige Eigenraum ist eindimensional. Ist $u_n \neq 0$ *eine Eigenfunktion zum Eigenwert* λ_n *und*

$$v_n := \left(\int_a^b r u_n^2\, dx\right)^{-\frac{1}{2}} u_n \qquad (n \geqslant 0), \tag{3.11}$$

so ist $(v_n)_{n\geqslant 1}$ ein Orthonormalsystem in bezug auf das Skalarprodukt

$$\langle f,g\rangle_r := \int_I fgr\,dx \qquad \big(f,g\in C(I)\big), \tag{3.12}$$

d. h. es gilt

$$\langle v_m, v_n\rangle_r = \delta_{mn} \qquad (m,n\geqslant 0). \tag{3.13}$$

v_n hat in $]a,b[$ genau n Nullstellen. Zwischen je zwei aufeinander folgenden Nullstellen von v_n liegt eine Nullstelle von v_{n+1}.

Der Beweis folgt nach Satz 3.2.15.

Frage: Wir „groß" ist der von den Eigenfunktionen aufgespannte Unterraum? Die Antwort liefert folgender

3.2.5 Satz (Entwicklungssatz) *In der Situation von Satz 3.2.4 gilt: Jedes $f\in C^1(I)$ mit $R_af = R_bf = 0$ lässt sich in eine gleichmäßig absolut konvergente Reihe nach den Eigenfunktionen entwickeln:*

$$f = \sum_{n=0}^{\infty} \langle f, v_n\rangle_r\, v_n. \tag{3.14}$$

Mehr dazu im nächsten Kapitel, Satz 3.3.1 und zugehörige Bemerkungen. Die Beweise dieser Sätze erfordern eine Menge Arbeit. Vorab zeigen wir einen kleinen Teil des Beweises von Satz 3.2.4:

3.2.6 Satz *Die Eigenwerte des Eigenwertproblems* (3.10) *sind alle einfach.*

Beweis Es seien $u\neq 0\neq v$ zwei Eigenfunktionen zum Eigenwert λ. Aus $R_a u = R_a v = 0$, d. h.

$$\begin{aligned} \alpha_1 u(a) + \alpha_2 u'(a) &= 0,\\ \alpha_1 v(a) + \alpha_2 v'(a) &= 0,\\ (\alpha_1,\alpha_2) &\neq (0,0) \end{aligned}$$

folgt dann (Lemma 3.1.7):

$$\begin{vmatrix} u(a) & u'(a)\\ v(a) & v'(a)\end{vmatrix} = 0.$$

Hier steht die Wronskische Determinante von u und v, und da diese an der Stelle a verschwindet, sind u, v nach Satz 2.4.3 linear abhängig, also ist λ einfacher Eigenwert. $\square$

3.2.7 Satz *Die Eigenwerte des Eigenwertproblems* (3.10) *sind nach unten beschränkt, d. h. es gibt ein $M\in\mathbb{R}$, so dass für jeden Eigenwert λ von* (3.2.4) *gilt: $\lambda\geqslant M$.*

Beweis Sei u eine Eigenfunktion zum Eigenwert λ:

$$\begin{aligned}\left(pu'\right)' + qu + \lambda r u &= 0,\\ R_a u + R_b u &= 0.\end{aligned}$$

Dann gilt:

$$\begin{aligned}\lambda \int_I r u^2\, dx &= \int_I \left(-\left(pu'\right)' - qu\right) u\, dx\\ &\overset{\text{part. Int.}}{=} -\left[pu'u\right]_a^b + \int_I pu'^2\, dx - \int_I qu^2\, dx.\end{aligned}$$

Hier hat wegen $R_a u = R_b u = 0$ die eckige Klammer die Form

$$-\left[pu'u\right]_a^b = c_1 u^2(a) + c_2 u^2(b)$$

mit $c_1, c_2 \in \mathbb{R}$, und mit $c := |c_1| + |c_2| \geqslant 0$ (c hängt nicht von u ab) gilt:

$$\lambda \int_I r u^2\, dx \geqslant -c\left(u^2(a) + u^2(b)\right) + \int_I pu'^2\, dx - \int_I qu^2\, dx. \tag{3.15}$$

Zur Abschätzung des ersten Terms verwenden wir das

3.2.8 Lemma Sei $r\colon I \to \mathbb{R}$ stetig, $r > 0$, $m := \min_{x\in I} r(x)$, $\varepsilon > 0$ und

$$M(\varepsilon) := \frac{1}{\varepsilon m} + \frac{1}{(b-a)\, m}.$$

Dann gilt für alle $u \in C^1(I)$ und alle $x \in I$:

$$u^2(x) \leqslant \varepsilon \int_I u'^2\, dx + M(\varepsilon) \int_I r u^2\, dx. \tag{3.16}$$

Wir setzen den Beweis von Satz 3.2.7 fort: (3.15) und (3.16) liefern:

$$\begin{aligned}\lambda \int_I r u^2\, dx &\geqslant -2c\left(\varepsilon \int_I u'^2\, dx + M(\varepsilon) \int_I r u^2\, dx\right) + \int_I pu'^2\, dx - \int_I qu^2\, dx\\ &= \int_I \left(p - 2\varepsilon c\right) u'^2\, dx - 2cM(\varepsilon) \int_I r u^2\, dx - \int_I qu^2\, dx.\end{aligned}$$

Hier wählen wir $\varepsilon > 0$ so klein, dass $0 \leqslant 2\varepsilon c \leqslant \min_{x\in I} p(x)$. Dann hängt ε nicht von u ab. Das liefert:

$$\begin{aligned}\lambda \int_I r u^2\, dx &\geqslant -2cM(\varepsilon) \int_I r u^2\, dx - \int_I qu^2\, dx\\ &\geqslant -\left(2cM(\varepsilon) + \frac{Q}{m}\right) \int_I r u^2\, dx\end{aligned}$$

mit $Q := \max_{x\in I} q(x)$, m wie oben. $u \neq 0$ und $r > 0$

$$\Rightarrow \lambda \geqslant -\left(2cM(\varepsilon) + \frac{Q}{m}\right).$$

Das ist die Behauptung von Satz 3.2.7. □

Beweis von Lemma 3.2.8: Sei $\varepsilon > 0$, $u \in C^1(I)$, $x \in I$ und $x_0 \in I$ so gewählt, dass $u^2(x_0) = \min_{x \in I} u^2(x)$;

$$\begin{aligned}
\Rightarrow u^2(x) &\leqslant u^2(x_0) + \left| \int_{x_0}^{x} 2u'u\,dt \right| \\
&\overset{\text{CSU}}{\leqslant} u^2(x_0) + 2\left(\varepsilon \left| \int_{x_0}^{x} u'^2\,dt \right|\right)^{\frac{1}{2}} \left(\frac{1}{\varepsilon}\left| \int_{x_0}^{x} u^2\,dt \right|\right)^{\frac{1}{2}} \\
&\overset{(3.17)}{\leqslant} u^2(x_0) + \varepsilon \left| \int_{x_0}^{x} u'^2\,dt \right| + \frac{1}{\varepsilon}\left| \int_{x_0}^{x} u^2\,dt \right| \\
&\leqslant \varepsilon \int_a^b u'^2\,dt + \underbrace{\left(\frac{1}{(b-a)\,m} + \frac{1}{\varepsilon m}\right)}_{=M(\varepsilon)} \int_a^b r u^2\,dt
\end{aligned}$$

wegen

$$2\sqrt{\alpha\beta} \leqslant \alpha + \beta \qquad \text{für } \alpha, \beta \geqslant 0. \tag{3.17}$$

□

Die Frage nach der Existenz von Eigenwerten wird durch das folgende Lemma zurück gespielt auf die Frage nach der Existenz von Eigenfunktionen eines Fredholmschen Integraloperators, benannt nach dem schwedischen Mathematiker Ivar Fredholm (1866-1927), der durch seine Theorie der Integralgleichungen einen wesentlichen Anstoß gab zur Entwicklung der Funktionalanalysis durch Hilbert, von Neumann, Riesz etc.

3.2.9 Satz *Es sei* 0 *kein Eigenwert des Eigenwertproblems* (3.10). *Dann gilt:*

a) Ist $\lambda \in \mathbb{R}$ und $y \in C^2(I)$, $y \neq 0$ eine Lösung des Eigenwertproblems (3.10), *so ist $\lambda \neq 0$ und*

$$w := \sqrt{r}\,y \in C(I) \tag{3.18}$$

ist eine Lösung der Fredholmschen Integralgleichung

$$\int_I K(x,t)\,w(t)\,dt = \mu w(x) \qquad (x \in I) \tag{3.19}$$

mit

$$\mu := \frac{1}{\lambda} \neq 0 \tag{3.20}$$

und dem symmetrischen stetigen Kern

$$K(x,t) := -\sqrt{r(x)\,r(t)}\,G(x,t) \tag{3.21}$$

mit der Greenschen Funktion G aus Satz 3.1.5.

b) *Umgekehrt: Ist* $w \in C(I)$, $w \neq 0$ *eine Lösung von* (3.19), *so ist* $\mu \neq 0$, *und*

$$y := \frac{1}{\sqrt{r}} w \in C^2(I) \tag{3.22}$$

ist eine Eigenfunktion von (3.10) *zum Eigenwert*

$$\lambda := \frac{1}{\mu}. \tag{3.23}$$

Beweis a) 0 ist kein Eigenwert von (3.10), also existiert nach Satz 3.1.5 die Greensche Funktion $G \colon I \times I \to \mathbb{R}$. Diese hat die folgende Eigenschaft: Für jedes beliebige $g \in C(I)$ hat das Randwertproblem

$$\begin{aligned} (pu')' + qu &= g \\ R_a u = R_b u &= 0 \end{aligned} \tag{3.24}$$

die eindeutig bestimmte Lösung

$$u(x) = \int_I G(x,t)\, g(t)\, dt \qquad (x \in I).$$

Laut Voraussetzung ist (3.24) erfüllt mit $u = y$, $g = -\lambda r y$; dabei ist $\lambda \neq 0$, da 0 kein Eigenwert ist. $\Rightarrow$ y genügt der Integralgleichung

$$y(x) = -\lambda \int_I G(x,t)\, r(t)\, y(t)\, dt \qquad (x \in I), \tag{3.25}$$

d. h. $w := \sqrt{r}\, y \in C(I)$ genügt der Fredholmschen Integralgleichung (3.19) mit $\mu = \frac{1}{\lambda}$ und K aus (3.21). K ist stetig und symmetrisch, da G stetig und symmetrisch ist.

b) Angenommen, $w \in C(I)$ und

$$\int_I K(x,t)\, w(t)\, dt = 0 \qquad (x \in I).$$

Dann ist $u := \frac{1}{\sqrt{r}} w \in C(I)$ und

$$\int_I G(x,t)\, u(t)\, dt = 0 \qquad (x \in I). \tag{3.26}$$

Andererseits: Nach Satz 3.1.5 ist

$$y(x) = \int_I G(x,t)\, u(t)\, dt \qquad (x \in I) \tag{3.27}$$

die eindeutig bestimmte Lösung des Randwertproblems

$$\begin{aligned} (py')' + qy &= u, \\ R_a y = R_b y &= 0. \end{aligned}$$

Da aber nach (3.26) und (3.27) $y = 0$ ist, folgt: $u = 0$. Also ist $\mu = 0$ kein Eigenwert der Integralgleichung (3.19). Sei nun $0 \neq w \in C(I)$ eine Lösung von (3.19). $\Rightarrow \mu \neq 0$, und $0 \neq y := \frac{1}{\sqrt{r}} w \in C(I)$ genügt der Integralgleichung (3.25) mit $\lambda := \frac{1}{\mu}$, d. h. es ist

$$y(x) = \int_I G(x,t)\, g(t)\, dt \qquad (x \in I)$$

mit $g := -\lambda r y \in C(I)$. Satz 3.1.5 liefert: Da 0 kein Eigenwert ist, ist y die zweimal stetig differenzierbare Lösung von (3.10). (Die zweimalige stetige Differenzierbarkeit von y kommt einfach durch Anwendung von Satz 3.1.5 heraus.) □

Ergebnis: Mit Satz 3.2.7 können wir (s.u.) den Beweis der Existenz von Eigenwerten auf den Fall reduzieren, dass 0 kein Eigenwert ist. Mit Satz 3.2.9 wird die Frage äquivalent umgeformt zur Frage nach der Existenz von Eigenwerten für die Fredholmsche Integralgleichung. Letztere Frage werden wir mit Hilfe der Spektraltheorie kompakter selbstadjungierter linearer Operatoren in einem Prähilbertraum lösen.

Exkurs (Spektraltheorie kompakter selbstadj. linearer Operatoren im Prähilbertraum)
$H := C(I)$ ist ein Prähilbertraum über $\mathbb{R}$, d. h. ein $\mathbb{R}$-Vektorraum mit dem Skalarprodukt

$$\langle f, g\rangle := \int_I f g\, dx.$$

Zu diesem Skalarprodukt gehört die Norm

$$\|f\| := \langle f, f\rangle^{\frac{1}{2}} = \left(\int_I f^2\, dx\right)^{\frac{1}{2}}.$$

Die Konvergenz in H ist die Konvergenz in der Norm: Für $f_n, f \in H$ gilt:

$$f_n \xrightarrow[H]{n\to\infty} f \Leftrightarrow \|f_n - f\| \xrightarrow{n\to\infty} 0 \qquad \text{(Konvergenz im quadratischen Mittel)}$$

Diese Konvergenz ist von der punktweisen und der gleichmäßigen Konvergenz sorgfältig zu unterscheiden; es gilt z. B.: $f_n \xrightarrow[\text{glm.}]{n\to\infty} f \Rightarrow f_n \xrightarrow[H]{n\to\infty} f$, aber die Umkehrung ist hier falsch.
Sei $K\colon I \times I \to \mathbb{R}$ stetig. Dann definiert K einen linearen Operator

$$T\colon H \to H, \quad (Tf)(x) := \int_I K(x,y)\, f(y)\, dy \qquad (x \in I).$$

Diese Konstruktion klappt auch in $L^2(I)$, falls K messbar ist und $\iint_{I\times I} |K(x,y)|^2\, dx\, dy < \infty$.

Ersichtlich ist T linear und hat Werte in H, und T ist beschränkt, d. h. es gibt eine Konstante $C \geqslant 0$, so dass $\|Tf\| \leqslant C \|f\|$ für alle $f \in H$. Begründung:

$$\begin{aligned}
\|Tf\|^2 &= \int_I ((Tf)(x))^2 \, dx \\
&= \int_I \left(\int_I K(x,y) f(y) \, dy \right)^2 dx \\
&\overset{\text{CSU}}{\leqslant} \int_I \left(\int_I (K(x,y))^2 \, dy \underbrace{\int_I f^2(y) \, dy}_{=\|f\|^2} \right) dx \\
&= \left(\int_I \int_I K^2 \, dx \, dy \right) \|f\|^2 .
\end{aligned}$$

Analoges gilt im $L^2(I)$. Insbesondere ist T stetig. – Umgekehrt zeigt man leicht: Jede stetige lineare Abbildung $S\colon H \to H$ ist beschränkt. Für jedes solche S setzt man

$$\|S\| := \sup\left\{ \|Sf\| : f \in H,\ \|f\| \leqslant 1 \right\},$$

und dann gilt:

$$\|Sf\| \leqslant \|S\| \, \|f\| \qquad (f \in H).$$

Insbesondere gilt für das obige T:

$$\|T\| \leqslant \left(\iint_{I \times I} K^2 \, dx \, dy \right)^{\frac{1}{2}} .$$

3.2.10 Satz *Ist K symmetrisch, d. h. $K(x,y) = K(y,x)$, so ist T selbstadjungiert, d. h.*

$$\langle Tf, g \rangle = \langle f, Tg \rangle \qquad (f, g \in H).$$

Beweis Für $f, g \in H$ ist

$$\langle Tf, g \rangle = \int_I \left(\int_I \underbrace{K(x,y)}_{=K(y,x)} f(y) \, dy \right) g(x) \, dx \overset{\text{Fubini}}{=} \int_I f(y) \left(\int_I K(y,x) g(x) \, dx \right) dy = \langle f, Tg \rangle . \quad \square$$

Anmerkung: Ist K nicht notwendig symmetrisch, so hat die Adjungierte T^* von T den Kern $K^*(x,y) = K(y,x)$ nach obigem Beweis.

3.2.11 Satz *T ist kompakt („vollstetig"), d. h.: Für jede bezüglich der Norm beschränkte Folge $(f_n)_{n \geqslant 1}$ in H hat $(Tf_n)_{n \geqslant 1}$ eine bezüglich der Norm konvergente Teilfolge.*

Beweis Sei $f_n \in H$, $\|f_n\| \leqslant C$ für alle $n \in \mathbb{N}$, und seien $\varepsilon > 0$ und δ das Delta der gleichmäßigen Stetigkeit von K auf $I \times I$, d. h.

$$|K(x,t) - K(x',t')| < \varepsilon \qquad \forall\ (x,t), (x',t') \in I \times I \quad \|(x,t) - (x',t')\|_2 < \delta .$$

Dann gilt $\forall\, x, x' \in I,\ \left|x - x'\right| < \delta \quad \forall\, n \in \mathbb{N}$:

$$\begin{aligned}
\left|\left(Tf_n\right)(x) - \left(Tf_n\right)\left(x'\right)\right| &\leqslant \int_I \underbrace{\left|K(x,t) - K\left(x',t\right)\right|}_{<\varepsilon} \left|f_n(t)\right| \, dt \\
&\leqslant \langle \varepsilon \chi_I, \left|f_n\right| \rangle \\
&\overset{\text{CSU}}{\leqslant} \varepsilon \left\|\chi_I\right\| \left\|f_n\right\| \\
&\leqslant C\sqrt{b-a}\varepsilon.
\end{aligned}$$

$\Rightarrow \left(Tf_n\right)$ ist gleichstetig. Wir wollen Satz 2.1.5 anwenden und müssen noch prüfen, ob $\left(Tf_n\right)$ gleichmäßig beschränkt ist. Dazu zeigen wir: $\left|K\left(x,y\right)\right| \leqslant M$ für alle $x, y \in I$. $\Rightarrow \forall\, x \in I,\ n \in \mathbb{N}$

$$\left|Tf_n(x)\right| \leqslant \int_I \underbrace{\left|K\left(x,y\right)\right|}_{\leqslant M} \left|f_n\left(y\right)\right| \, dy \leqslant M \langle \chi_I, \left|f_n\right| \rangle \overset{\text{CSU}}{\leqslant} \sqrt{b-a}M \left\|f_n\right\| \leqslant \sqrt{b-a}MC.$$

$\Rightarrow \left(Tf_n\right)$ ist gleichmäßig beschränkt. Satz 2.1.5 liefert nun: $\left(Tf_n\right)$ hat eine gleichmäßig konvergente Teilfolge $\left(Tf_{n_k}\right)_{k\geqslant 1}$, und diese konvergiert a forteriori im quadratischen Mittel. □

Die folgenden Überlegungen gelten für beliebige beschränkte selbstadjungierte lineare Operatoren S in einem beliebigen Prähilbertraum H über $\mathbb{R}$ mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und zugehöriger Norm $\|\cdot\|$.

3.2.12 Satz *Sei S ein selbstadjungierter linearer Operator im Prähilbertraum $H \neq \{0\}$. Dann gilt:*

a) Sind u, v Eigenvektoren von S zu den Eigenwerten λ bzw. μ, $\lambda \neq \mu$, so ist $u \perp v$.

b) Ist S beschränkt, so gilt:

$$\|S\| = \sup\left\{\left|\langle Sf, f\rangle\right| : f \in H,\ \left\|f\right\| = 1\right\}.$$

Beweis a)

$$\lambda \langle u, v\rangle = \langle Su, v\rangle = \langle u, Sv\rangle = \mu \langle u, v\rangle.$$

$\lambda \neq \mu \Rightarrow \langle u, v\rangle = 0$.

b) Für $f \in H$, $\left\|f\right\| = 1$ ist

$$\left|\langle Sf, f\rangle\right| \overset{\text{CSU}}{\leqslant} \left\|Sf\right\| \left\|f\right\| \leqslant \|S\|,$$

also

$$\alpha := \sup\left\{\left|\langle Sf, f\rangle\right| : f \in H,\ \left\|f\right\| = 1\right\} \leqslant \|S\|.$$

Umgekehrt: S selbstadjungiert, also folgt (wir bewegen uns in $\mathbb{R}$!)

$$\langle Sf, g\rangle = \frac{1}{4}\left(\langle S\left(f+g\right), f+g\rangle - \langle S\left(f-g\right), f-g\rangle\right).$$

Das liefert

$$\left|\langle Sf, g\rangle\right| \leqslant \frac{1}{4}\alpha\left(\|f+g\|^2 + \|f-g\|^2\right) = \frac{1}{2}\alpha\left(\|f\|^2 + \|g\|^2\right).$$

$$\Rightarrow \sup_{\|f\|=\|g\|=1} \left|\langle Sf, g\rangle\right| \leqslant \alpha. \tag{3.28}$$

Hier ist

$$\sup_{\|g\|=1} \left|\langle Sf, g\rangle\right| = \|Sf\|,$$

also besagt (3.28): $\sup_{\|f\|=1} \|Sf\| \leqslant \alpha$, d. h. $\|S\| \leqslant \alpha$. □

3.2.13 Satz *Ist S ein kompakter selbstadjungierter linearer Operator im Prähilbertraum $H \neq \{0\}$, so existiert ein Eigenwert λ von S mit $|\lambda| = \|S\|$.*

Bemerkung Für jeden Eigenwert μ von S gilt trivialerweise $|\mu| \leqslant \|S\|$. Der Satz garantiert also, dass es einen Eigenwert von größtmöglichem Betrag gibt.

BEWEIS (i) $S = 0 \Rightarrow 0$ ist Eigenwert vom Betrag $\|S\|$.

(ii) $S \neq 0 \Rightarrow \|S\| > 0$, und Satz 3.2.12 b) liefert, da S selbstadjungiert ist:

$$\exists f_n \in H, \ \|f_n\| = 1 \quad \left|\langle Sf_n, f_n\rangle\right| \xrightarrow{n\to\infty} \|S\|.$$

Gegebenenfalls nach Übergang zu einer Teilfolge kann oBdA angenommen werden:

$$\langle Sf_n, f_n\rangle \xrightarrow{n\to\infty} \lambda, \qquad |\lambda| = \|S\|.$$

S ist kompakt, also hat (Sf_n) nach Satz 3.2.11 eine konvergente Teilfolge. Nach erneutem Übergang zu einer Teilfolge kann also oBdA angenommen werden, dass $(Sf_n)_{n\geqslant 1}$ konvergiert. Wegen $|\lambda| = \|Sf\| > 0$ kann man den Limes von (Sf_n) in der Form λf mit $f \in H$ schreiben und hat

$$Sf_n \xrightarrow[H]{n\to\infty} \lambda f. \tag{3.29}$$

Nun gilt:

$$\begin{aligned} 0 &\leqslant \|Sf_n - \lambda f_n\|^2 \\ &= \underbrace{\|Sf_n\|^2}_{\leqslant \|S\|^2 = \lambda^2} - 2\lambda \underbrace{\langle Sf_n, f_n\rangle}_{\xrightarrow{n\to\infty} \lambda} + \lambda^2 \underbrace{\|f_n\|^2}_{=1} \\ &\leqslant 2\lambda^2 - 2\lambda\langle Sf_n, f_n\rangle \xrightarrow{n\to\infty} 0. \end{aligned}$$

$$\Rightarrow \|\lambda f - \lambda f_n\| \leqslant \|\lambda f - Sf_n\| + \|Sf_n - \lambda f_n\| \xrightarrow{n\to\infty} 0.$$

Da hier $\lambda \neq 0$ ist, folgt $f_n \xrightarrow{n\to\infty} f$, speziell $\|f\| = 1$. S ist stetig, also folgt $Sf_n \xrightarrow{n\to\infty} Sf$, nach (3.29) gilt auch $Sf_n \xrightarrow{n\to\infty} \lambda f$, also wegen der Eindeutigkeit des Limes bereits $Sf = \lambda f$, $\|f\| = 1$. $\Rightarrow f$ ist Eigenvektor zum Eigenwert λ. □

Kommentar: Jeder endlich-dimensionale euklidische Vektorraum ist (reeller) Prähilbertraum. Auf diesem ist jeder lineare Operator $T\colon V \to V$ kompakt (nach Bolzano-Weierstraß), denn jeder lineare Operator auf V ist beschränkt. In Satz 3.2.13 ist also ein wohlbekannter Satz der Linearen Algebra II über die Existenz von Eigenwerten selbstadjungierter linearer Abbildungen enthalten. Der obige Beweis kommt ganz ohne charakteristisches Polynom und Fundamentalsatz der Algebra aus! – Wie im Beweis zu Satz 3.2.15 kann man dann die Existenz einer Orthonormalbasis von Eigenvektoren beweisen.

3.2.14 Satz *Sei S ein kompakter selbstadjungierter linearer Operator im Prähilbertraum H. Dann gilt:*

a) *Für jedes $\varepsilon > 0$ gibt es höchstens endlich viele Eigenwerte λ von S mit $|\lambda| \geqslant \varepsilon$.*

b) *Jeder Eigenwert $\lambda \neq 0$ von S hat endliche Vielfachheit.*

Beweis a) Gäbe es unendlich viele verschiedene Eigenwerte λ_n von S mit $|\lambda_n| \geqslant \varepsilon$ und mit zugehörigen Eigenvektoren v_n, $\|v_n\| = 1$, so gilt nach Satz 3.2.12: $\langle v_m, v_n\rangle = \delta_{mn}$ für alle $m, n \in \mathbb{N}$. Die Folge $(Sv_n)_{n\geqslant 1} = (\lambda_n v_n)_{n\geqslant 1}$ müsste eine Norm-konvergente Teilfolge enthalten. Es ist aber wegen $v_m \perp v_n$ für $m \neq n$:

$$\|\lambda_m v_m - \lambda_n v_n\|^2 = \lambda_m^2 + \lambda_n^2 \geqslant 2\varepsilon^2.$$

Das aber ist ein Widerspruch zum Cauchy-Kriterium.

b) Gäbe es unendlich viele linear unabhängige Eigenvektoren zum Eigenwert λ, so gäbe es ein Orthonormalsystem (v_n) von Eigenvektoren zum Eigenwert λ. Da S kompakt ist, müsste (Sv_n) eine konvergente Teilfolge enthalten. Aber: Wegen $Sv_n = \lambda v_n$ ist für $m \neq n$:

$$\|Sv_m - Sv_n\|^2 = 2\lambda^2 \neq 0.$$

Auch das ist ein Widerspruch zum Cauchy-Kriterium. □

3.2.15 Satz (Spektralsatz für kompakte selbstadjungierte Operatoren) *Es seien H ein unendlich-dimensionaler Prähilbertraum (über $\mathbb{R}$) und $S\colon H \to H$ ein selbstadjungierter, kompakter linearer Operator.*

a) *Dann hat S eine (nach Vielfachheit gezählte) unendliche Folge $(\lambda_n)_{n\geqslant 0}$ (reeller) Eigenwerte*

$$|\lambda_0| \geqslant |\lambda_1| \geqslant |\lambda_2| \geqslant \ldots, \qquad \lambda_n \xrightarrow{n\to\infty} 0,$$

in der jeder von 0 verschiedene Eigenwert von S vorkommt, mit einem zugehörigen Orthonormalsystem $\big(f_n\big)$ von Eigenvektoren:

$$Sf_n = \lambda_n f_n, \qquad \langle f_m, f_n\rangle = \delta_{mn} \qquad (m, n \in \mathbb{N}_0).$$

Dabei gilt: Ist

$$H_n := \left(\operatorname{span}\left(f_0, \dots, f_{n-1}\right)\right)^{\perp} = \left\{g \in H \colon \langle f_j, g\rangle = 0 \text{ für } j = 0, \dots, n-1\right\},$$

so ist H_n S-invariant, und es ist

$$|\lambda_n| = \sup_{\substack{g \in H_n \\ \|g\|=1}} \left|\langle Sg, g\rangle\right| = \left\|S|_{H_n}\right\| \qquad (n \geqslant 0).$$

b) *Ist $g = Sf \in S(H)$ ($f \in H$), so hat g die folgende Entwicklung im Sinne der Norm von H:*

$$g = \sum_{k=0}^{\infty} \langle g, f_k\rangle f_k = \sum_{k=0}^{\infty} \lambda_k \langle f, f_k\rangle f_k.$$

Bemerkung Ein kompakter selbstadjungierter linearer Operator in einem Hilbertraum (das ist ein vollständiger Prähilbertraum) hat ein vollständiges Orthonormalsystem von Eigenvektoren (und die Folge der von 0 verschiedenen Eigenwerte ist eine Nullfolge). Das kann leicht aus Satz 3.2.15 gefolgert werden. Wenn man ein vollständiges Orthonormalsystem von Eigenvektoren hat, kann man natürlich jedes $g \in H$ entwickeln analog zu b). Aber: In Prähilberträumen gilt Entwickelbarkeit nur unter einschränkenden Voraussetzungen wie unter b). Zu dieser Problematik siehe Dieudonné (1978), 11.5.7 und 11.5.8 und dazu Exercise 14.

Beweis a) Es ist $H_0 = H$, und nach Satz 3.2.13 gibt es einen Eigenwert λ_0 von S mit

$$|\lambda_0| = \|S\| = \sup\left\{\left|\langle Sg, g\rangle\right| : g \in H,\ \|g\| = 1\right\}$$

(siehe auch Satz 3.2.12 b).) Sei f_0 ein normierter Eigenvektor zu λ_0. Damit sind λ_0, f_0 konstruiert. Sei weiter $n \geqslant 1$, und $\lambda_0, \dots, \lambda_{n-1}$ mit zugehörigen orthonormierten Eigenvektoren $f_0, \dots, f_{n-1}$ seien schon konstruiert.

$$H_n := \left(\operatorname{span}\left(f_0, \dots, f_{n-1}\right)\right)^{\perp}.$$

$\Rightarrow$ H_n ist ein unendlich-dimensionaler Prähilbertraum, und H_n ist S-invariant, denn: Sei $g \in H_n$, $0 \leqslant j \leqslant n-1$.

$$\Rightarrow \langle Sg, f_j\rangle = \langle g, Sf_j\rangle = \langle g, \lambda_j f_j\rangle = 0.$$

$\Rightarrow Sg \in H_n$. $\Rightarrow S|_{H_n}$ ist ein kompakter selbstadjungierter linearer Operator im Prähilbertraum H_n, und H_n ist unendlich-dimensional. Satz 3.2.13 liefert dann: $S|_{H_n}$ hat einen Eigenwert λ_n mit $|\lambda_n| = \sup_{\substack{g \in H_n \\ \|g\|=1}} \left|\langle Sg, g\rangle\right|$ und eine zugehörige Eigenfunktion f_n mit $\left|f_n\right| = 1$. Satz 3.2.14 liefert: $|\lambda_n| \xrightarrow{n\to\infty} 0$, und nach Konstruktion gilt:

$$|\lambda_0| \geqslant |\lambda_1| \geqslant |\lambda_2| \geqslant \dots$$

Bemerkung Ist H ein Hilbert-Raum, so kann man (f_n) durch eine Orthonormalbasis von $\left(\operatorname{span}(f_n \colon n \geqslant 0)\right)^{\perp} \subset \ker S$ zu einer Orthonormalbasis von Eigenvektoren von S ergänzen. Vgl. Dieudonné (1978).

b) Sei $f \in H$.

$$\Rightarrow g_n := f - \sum_{k=0}^{n-1} \langle f, f_k \rangle f_k \in H_n,$$

also gilt:

$$\|Sg_n\| \leqslant \underbrace{\|S|_{H_n}\|}_{=|\lambda_n|} \underbrace{\|g_n\|}_{\leqslant \|f\|} \leqslant |\lambda_n| \, \|f\| \xrightarrow{n\to\infty} 0,$$

denn nach Pythagoras ist

$$\|g_n\|^2 = \|f\|^2 - \sum_{k=1}^{n} |\langle f, f_k \rangle|^2 \leqslant \|f\|^2.$$

Wegen

$$Sg_n = g - \sum_{k=0}^{n-1} \lambda_k \langle f, f_k \rangle f_k = g - \sum_{k=0}^{n-1} \langle g, f_k \rangle f_k$$

ist damit alles klar. □

Beweis (zu Satz 3.2.4) Nach Satz 3.2.7 gibt es ein $M \in \mathbb{R}$, so dass für alle Eigenwerte λ von (3.10) gilt: $\lambda \geqslant M$. Sei $\alpha \in \mathbb{R}$, $\alpha < M$. $\Rightarrow \alpha$ ist kein Eigenwert des Problems (3.10). Wir betrachten das Eigenwertproblem

$$\begin{aligned} Lu + \alpha r u + \mu r u &= 0, \\ R_a u = R_b u &= 0. \end{aligned} \tag{3.30}$$

Offenbar gilt: λ ist Eigenwert von (3.10) mit Eigenfunktion u genau dann, wenn $\mu := \lambda - \alpha$ Eigenwert von (3.30) ist mit Eigenfunktion u. Satz 3.2.4 gilt genau dann für das Eigenwertproblem (3.10), wenn er für das Eigenwertproblem (3.30) gilt, und nach Konstruktion ist $\mu = 0$ kein Eigenwert für (3.30). Mehr noch: Alle Eigenwerte von (3.30) sind laut Konstruktion positiv. D. h.: Wir dürfen für den Beweis oBdA gleich annehmen, dass alle Eigenwerte von (3.10) positiv sind. Nun greift Satz 3.2.9: λ ist Eigenwert von (3.10) mit Eigenfunktion u genau dann, wenn $\mu = \frac{1}{\lambda}$ Eigenwert ist von

$$T \colon C(I) \to C(I), \quad Tf(x) := \int_I K(x,t) f(t)\, dt \qquad (x \in I)$$

(K aus (3.21)) mit Eigenfunktion $w = \sqrt{r}u$. Beachte: $\mu = 0$ ist kein Eigenwert von T nach Satz 3.2.9! Nun ist aber T kompakt und selbstadjungiert nach Satz 3.2.10 und Satz 3.2.11. Nach Satz 3.2.15 hat T eine Folge (μ_n) von Eigenwerten mit

$$|\mu_0| \geqslant \|\mu_1| \geqslant \ldots, \quad \mu_n \xrightarrow{n\to\infty} 0,$$

und da alle Eigenwerte positiv sind, gilt hier $\mu_0 \geqslant \mu_1 \geqslant \mu_2 \geqslant \ldots, \mu_n > 0$ für alle $n \in \mathbb{N}$, $\mu_n \xrightarrow{n\to\infty} 0$. Sei (w_n) ein zugehöriges Orthonormalsystem von Eigenfunktionen. $\Rightarrow \lambda_n := \frac{1}{\mu_n}$ für alle $n \geqslant 0$ sind Eigenwerte von (3.10) mit zugehörigen Eigenfunktionen $v_n := \frac{1}{\sqrt{r}} w_n$ für alle $n \geqslant 0$. Dabei gilt im Sinne von (3.12):

$$\langle v_m, v_n \rangle_r = \int_I v_m v_n r \, dx = \int_I w_m w_n \, dx = \delta_{mn},$$

d. h. (v_n) ist ein Orthonormalsystem bezüglich des Skalarprodukts (3.12). Nach Satz 3.2.6 sind alle Eigenwerte einfach, also $\mu_0 > \mu_1 > \mu_2 > \ldots, \mu_n > 0$, $\mu_n \xrightarrow{n\to\infty} 0$, also $\lambda_0 < \lambda_1 < \lambda_2 < \ldots, \lambda_n \xrightarrow{n\to\infty} \infty$. Bleibt zu zeigen:

3.2.16 Satz *Für die Eigenfunktionen v_n $(n \geqslant 0)$ des Sturm-Liouvilleschen Eigenwertproblems Problem 3.2.2 gilt: v_n hat in $]a,b[$ genau n Nullstellen. Zwischen je zwei aufeinanderfolgenden Nullstellen von v_n liegt eine Nullstelle von v_{n+1}. Sind $a < x_1 < x_2 < \ldots < x_n < b$ die Nullstellen von v_n, so liegt in $]a, x_1[$ und in $]x_n, b[$ eine Nullstelle von v_{n+1}.*

Beweis Dass v_n in $]a,b[$ genau n Nullstellen hat, beweist man mit Hilfe der Prüfer-Transformation (vgl. Beweis von Satz 2.6.8) und einigen tüfteligen Argumenten, die wir aus Zeitgründen weglassen müssen. Am besten kann man die Details nachlesen bei Hsieh und Sibuya (1999), S. 159 ff., Beweis von Theorem VI.3.11, und bei Coddington und Levinson (1998), S. 212 ff., Beweis von Theorem 2.1 oder bei Weidmann (1987a). Einen Teil der Aussagen von Satz 3.2.16 kann man sehr durchsichtig mit der Methode des Beweises von Satz 2.6.5 beweisen. Wir zeigen:

(I) Hat v_n in $]a,b[$ genau $k \geqslant 0$ Nullstellen $a < x_1 < x_2 < \ldots < x_k < b$, so hat v_{n+1} in $]a,b[$ $k+1$ Nullstellen $x'_1, \ldots, x'_{k+1}$, so dass $a < x'_1 < x_1 < x'_2 < x_2 < \ldots < x_k < x'_{k+1} < b$.
Begründung:

(i) Zwischen x_j und x_{j+1} $(1 \leqslant j \leqslant k-1)$ liegt eine Nullstelle von v_{n+1} nach Satz 2.6.5.

(ii) In $]a, x_1[$ und in $]x_k, b[$ (Fall $k \geqslant 1$) (bzw. in $]a,b[$ im Fall $k = 0$) liegt eine Nullstelle von v_{n+1}.
Begründung: Sei

$$c := \begin{cases} x_1, & \text{falls } k \geqslant 1, \\ b, & \text{falls } k = 0. \end{cases}$$

Angenommen, die Behauptung sei falsch. Dann dürfen wir oBdA annehmen, dass für $a < x < c$ gilt: $v_n(x) > 0$, $v_{n+1}(x) > 0$. Aus

$$\begin{aligned} Lv_n + \lambda_n r v_n &= 0, \\ Lv_{n+1} + \lambda_{n+1} r v_{n+1} &= 0 \end{aligned}$$

erhalten wir (Multiplikation mit v_{n+1} bzw. v_n und Subtraktion):

$$\begin{aligned}
0 &= v_{n+1} L v_n - v_n L v_{n+1} - (\lambda_{n+1} - \lambda_n)\, r\, v_n v_{n+1} \\
&\overset{\text{Satz 2.6.4}}{=} \frac{d}{dx}\left(p \begin{vmatrix} v_{n+1} & v_n \\ v'_{n+1} & v'_n \end{vmatrix}\right) - \underbrace{(\lambda_{n+1} - \lambda_n)}_{>0}\, \underbrace{r\, v_n v_{n+1}}_{>0 \text{ auf }]a,c[}
\end{aligned}$$

$$\Rightarrow 0 < \int_a^c (\lambda_{n+1} - \lambda_n)\, r\, v_n v_{n+1}\, dx = \left[p \begin{vmatrix} v_{n+1} & v_n \\ v'_{n+1} & v'_n \end{vmatrix}\right]_a^c \tag{3.31}$$

Schon in Lemma 3.1.7 haben wir gesehen, dass die Randbedingungen implizieren, dass $\begin{vmatrix} v_{n+1} & v_n \\ v'_{n+1} & v'_n \end{vmatrix}$ für $x = a$ und für $x = b$ verschwindet.

Fall (A) Für $c = b$ ist die rechte Seite in (3.31) gleich Null: Widerspruch!

Fall (B) Für $c = x_1$ ist die rechte Seite von (3.31) gleich

$$p(x_1) \begin{vmatrix} v_{n+1}(x_1) & \underbrace{v_n(x_1)}_{=0} \\ v'_{n+1}(x_1) & v'_n(x_1) \end{vmatrix} = p(x_1)\, v_{n+1}(x_1)\, v'_n(x_1). \tag{3.32}$$

Hier ist $p(x_1) > 0$, $v_{n+1}(x_1) \geqslant 0$, aber $v'_n(x_1) < 0$, d. h. die rechte Seite von (3.32) ist $\leqslant 0$. Das aber ist ein Widerspruch zur Positivität von (3.31).

Der gleiche Schluss klappt für $]x_k, b[$: Wieder nimmt man oBdA an, dass $v_n(x) > 0$, $v_{n+1}(x) > 0$ auf $]x_k, b[$, und hat

$$\begin{aligned}
0 &< \int_{x_k}^b (\lambda_{n+1} - \lambda_n)\, r\, v_n v_{n+1}\, dx \\
&= \left[p \begin{vmatrix} v_{n+1} & v_n \\ v'_{n+1} & v'_n \end{vmatrix}\right]_{x_k}^b \\
&= -p(x_k)\, v_{n+1}(x_k)\, v'_n(x_k)
\end{aligned} \tag{3.33}$$

(auch hier verschwindet die Determinante an der Stelle b, $v_n(x_k) = 0$). Hier ist $p(x_k) > 0$, $v_{n+1}(x_k) \geqslant 0$, $v'_n(x_k) > 0$, also (3.33)$\leqslant 0$: Widerspruch zu (3.31)!

Damit ist (I) bewiesen. Zum vollständigen Beweis von Satz 3.2.16 fehlen noch zwei Punkte:

(II) Zwischen je zwei aufeinander folgenden Nullstellen von v_{n+1} in $]a, b[$ liegt eine Nullstelle von v_n. ($\Rightarrow$ Hat v_n genau k Nullstellen in $]a, b[$, so hat v_{n+1} dort genau $k+1$ Nullstellen.)

(III) v_0 hat in $]a, b[$ keine Nullstelle. □

Die Punkte (II) und (III) lassen wir aus Zeitgründen offen. □

Bemerkung Das in Satz 3.2.16 angesprochene Nullstellenverhalten liegt bei vielen klassischen Systemen orthogonaler Polynome vor (Legendresche Polynome, Tschebyscheffsche Polynome, Laguerresche Polynome, Jacobische Polynome, ...). Zu diesen Polynomen siehe z. B. Courant und Hilbert (1924, 1937) oder Tricomi (1955) oder Szegö (1998).

Bemerkung Hochstadt (1975) behauptet, dass man ähnlich wie oben auch zeigen kann, dass zwischen je zwei aufeinander folgenden Nullstellen von v_{n+1} auch eine Nullstelle von v_n liegt. Das kann *so* nicht stimmen: Wir haben oben den Vergleichssatz benutzt, und der unterstellt nur die Gültigkeit der Differentialgleichungen, aber keine Randbedingungen. Aus den Differentialgleichungen

$$y'' + \omega^2 y = 0, \qquad \lambda_1 = \omega^2 < 1, \tag{3.34}$$

$$y'' + y = 0, \qquad \lambda_2 = 1 \tag{3.35}$$

folgt aber nur: Die Nullstellen der Lösungen von (3.34)

$$y_1(x) = A\cos(\omega x + \alpha) \qquad (A \neq 0)$$

haben konstanten Abstand $\frac{\pi}{\omega} > 1$, die Nullstellen der Lösungen von (3.35)

$$y_2(x) = B\cos\left(x + \beta\right) \qquad (B \neq 0)$$

haben aber den konstanten Abstand π. Zwischen je zwei aufeinander folgenden Nullstellen von y_1 liegt eine Nullstelle von y_2, eventuell sogar mehrere Nullstellen von y_2, aber zwischen 2 aufeinander folgenden Nullstellen von y_2 braucht keine Nullstelle von y_1 zu liegen. Für die Gültigkeit von Satz 3.2.16 sind daher die Randbedingungen ganz wesentlich!

Literaturhinweise: Bei Heuser (1995b) und Walter (1976b) findet man gut lesbare Darstellungen des Sturm-Liouvilleschen Eigenwertproblems, desgleichen bei Peyerimhoff (1970), Storch und Wiebe (1999). Hervorragend ist auch Coddington und Levinson (1998). Dort wird auch das sog. „singuläre“ Sturm-Liouville-Problem behandelt. Zu diesem gibt es viele weitere Literatur, z. B. Hellwig (1964), Jörgens (1964), Jörgens und Rellich (1976), Weidmann (1987b), Weidmann (2003). Dieses Buch enthält eine sehr schöne Behandlung von regulärem und singulärem Sturm-Liouville-Problem unter systematischem Einsatz der Spektraltheorie (auch unbeschränkter) linearer Operatoren im Hilbert-Raum. Eine gut lesbare, moderne Darstellung der Theorie gewöhnlicher Differentialgleichungen (einschließlich des Sturm-Liouville-Problems) ist Hsieh und Sibuya (1999). Ein klassisches Werk über die Eigenwerttheorie gewöhnlicher Differentialgleichungen ist Titchmarsh (1962, 1958). Den Zusammenhang mit Problemen der mathematischen Physik findet man in dem Klassiker Courant und Hilbert (1924, 1937).

3.3 Entwicklungssätze

Wir behalten die Situation und die Bezeichnungen von Abschnitt 3.2 bei: Sturm-Liouvillesches Eigenwertproblem: $I = [a,b]$, $p,q,r\colon I \to \mathbb{R}$ stetig, p stetig differenzierbar, $p > 0$, $r > 0$,

$Lu := \left(pu'\right)' + qu$ $(u \in C^2(I))$. $(\alpha_1, \alpha_2) \neq (0,0) \neq \left(\beta_1, \beta_2\right)$:

$$R_a u := \alpha_1 u(a) + \alpha_2 u'(a)$$
$$R_b u := \beta_1 u(b) + \beta_2 u'(b)$$

Wir betrachten das Eigenwertproblem

$$Lu + \lambda r u = 0$$
$$R_a u = R_b u = 0 \qquad \left(u \in C^2(I),\ \lambda \in \mathbb{R}\right)$$

Eigenwerte $\lambda_0 < \lambda_1 < \lambda_2 < \ldots,\ \lambda_n \xrightarrow{n\to\infty} \infty$, zugehöriges Orthonormalsystem von Eigenfunktionen v_n $(n \geqslant 0)$, Skalarprodukt:

$$\langle f, g\rangle_r = \int_I f g r\, dx, \qquad \langle v_m, v_n\rangle_r = \delta_{mn}.$$

Ist $\lambda = 0$ kein Eigenwert, G die Greensche Funktion,

$$K(x,t) = -\sqrt{r(x)\, r(t)}\, G(x,t),$$

so ist λ Eigenwert von L mit Eigenfunktion v genau dann, wenn

$$(Tw)(x) := \int_I K(x,t)\, w(t)\, dt = \mu w(x)$$

mit $\mu = \frac{1}{\lambda}$, $w = \sqrt{r}\, v$.

3.3.1 Satz (Entwicklungssatz) *Seien $\lambda_0 < \lambda_1 < \lambda_2 < \ldots,\ \lambda_n \xrightarrow{n\to\infty} \infty$ die Eigenwerte des Sturm-Liouvilleschen Eigenwertproblems und v_n $(n \geqslant 0)$ ein zugehöriges Orthonormalsystem von Eigenfunktionen. Dann hat jedes $f \in C^2(I)$ mit $R_a f = R_b f = 0$ die gleichmäßig absolut und im quadratischen Mittel konvergente Reihenentwicklung*

$$f = \sum_{n=0}^{\infty} \langle f, v_n\rangle_r\, v_n$$

mit

$$\langle f, v_n\rangle_r := \int_I f v_n r\, dx.$$

Bemerkung Der Satz gilt auch für alle $f \in C^1(I)$ mit $R_a f = R_b f = 0$. Siehe dazu Kamke (1950).

Beweis Wie zu Anfang des Beweises von Satz 3.2.4 können wir durch eine „Verschiebung“ erreichen, dass 0 kein Eigenwert ist. Dadurch werden die Eigenfunktionen nicht verändert. Daher kann oBdA angenommen werden: 0 ist kein Eigenwert. Sei $f \in C^2(I)$, $R_a f = R_b f = 0$.

$$\Rightarrow Lf = \left(pf'\right)' + qf =: -\sqrt{r}\, g \in C(I),$$

und 0 ist kein Eigenwert. Dann liefert Satz 3.1.5:

$$f(x) = \int_I G(x,t)\left(-\sqrt{r(t)}g(t)\right) dt,$$

d. h.

$$\sqrt{r(x)}f(x) = \int_I K(x,t)\, g(t)\, dt = Tg(x) \in TH$$

mit $H = C(I)$ und T aus dem Beweis von Satz 3.2.4. Satz 3.2.15 b) liefert nun

$$\sqrt{r}f = \sum_{k=0}^{\infty} \langle \sqrt{r}f, w_k \rangle\, w_k, \tag{3.36}$$

wobei die Reihe im quadratischen Mittel konvergiert und wobei hier

$$\langle u, v \rangle = \int_I uv\, dx \qquad (u, v \in H = C(I)).$$

Nach Definition der $v_k = \frac{1}{\sqrt{r}} w_k$ heißt das:

$$f = \sum_{k=0}^{\infty} \langle f, v_k \rangle_r\, v_k \qquad \text{(Entwicklung im quadratischen Mittel)}$$

Behauptung: Die Entwicklung gilt auch im Sinne der gleichmäßig absoluten Konvergenz.
Begründung: Unter (3.36) ist

$$\langle \sqrt{r}f, w_k \rangle\, w_k(x) = \langle Tg, w_k \rangle\, w_k(x) \overset{T \text{ selbstadj.}}{=} \langle g, Tw_k \rangle\, w_k(x) = \langle g, w_k \rangle\, \mu_k w_k(x),$$

also

$$\begin{aligned} \sum_{k=m}^{n} \left|\langle \sqrt{r}f, w_k \rangle\, w_k(x)\right| &= \sum_{k=m}^{n} \left|\langle g, w_k \rangle\right| \left|\mu_k w_k(x)\right| \\ &\overset{\text{CSU}}{\leqslant} \left(\sum_{k=m}^{n} \langle g, w_k \rangle^2\right)^{\frac{1}{2}} \left(\sum_{k=m}^{n} \left(\mu_k w_k(x)\right)^2\right)^{\frac{1}{2}}. \end{aligned} \tag{3.37}$$

Hier ist für alle $k \geqslant 0$

$$\mu_k w_k(x) = \int_I K(x,t)\, w_k(t)\, dt = \langle K(x,\cdot), w_k \rangle,$$

also nach der Besselschen Ungleichung (siehe z. B. Elstrodt (2002b))

$$\begin{aligned} \sum_{k=m}^{n} \left(\mu_k w_k(x)\right)^2 &= \sum_{k=m}^{n} \left(\langle K(x,\cdot), w_k \rangle\right)^2 \\ &\leqslant \|K(x,\cdot)\|^2 \\ &= \int_I (K(x,t))^2\, dt \\ &\leqslant M^2 (b-a) \qquad (x \in I,\ m, n \in \mathbb{N}), \end{aligned} \tag{3.38}$$

wobei $|K| \leqslant M$ auf $I \times I$. (3.37) und (3.38) besagen dann mit der Besselschen Ungleichung: Das Cauchy-Kriterium für die gleichmäßige Konvergenz von (3.36) ist erfüllt, die Reihe $\sum_{k=0}^{\infty} \langle f, v_k \rangle_r \, v_k$ konvergiert also gleichmäßig absolut auf I. Da die Reihe im quadratischen Mittel gegen f konvergiert und f stetig ist, folgt die Behauptung. □

3.3.2 Korollar (v_n) ist ein vollständiges Orthonormalsystem in

$$L_r^2(I) = \left\{ f \colon I \to \mathbb{R} \text{ messbar: } \int_I |f|^2 \, r \, dx < \infty \right\}$$

(mit Skalarprodukt $\langle f, g \rangle_r = \int_I f g r \, dx$), d. h. Für jedes $f \in L_r^2(I)$ gilt:

$$f = \sum_{n=0}^{\infty} \langle f, v_n \rangle_r \, v_n \qquad \text{(Konvergenz im quadratischen Mittel).}$$

Insbesondere gelten für $f, g \in L_r^2(I)$ die Parsevalsche Gleichung

$$\langle f, g \rangle_r = \sum_{n=0}^{\infty} \langle f, v_n \rangle_r \langle v_n, g \rangle_r$$

und die Vollständigkeitsrelation

$$\|f\|_r^2 = \sum_{n=0}^{\infty} \left| \langle f, v_n \rangle_r \right|^2 .$$

Beweis

$$U := \left\{ h \in C^2(I) : R_a h = R_b h = 0 \right\}$$

liegt dicht in $L_r^2(I)$, und $\operatorname{span}(v_n \colon n \geqslant 0) =: V$ liegt dicht in U. $\Rightarrow$ V liegt dicht in $L_r^2(I)$. Der Satz von der besten Approximation (siehe Elstrodt (2002c), dieser Satz gilt auch für Prähilberträume!) liefert nun die Behaptung. □

3.3.3 Satz (James Mercer, 1909) *Es sei* 0 *kein Eigenwert des Sturm-Liouvilleschen Eigenwertproblems. Dann hat die Greensche Funktion G die gleichmäßig absolut konvergente Entwicklung*

$$G(x, t) = -\sum_{n=0}^{\infty} \frac{1}{\lambda_n} v_n(x) \, v_n(t) \qquad (x, t \in I).$$

Beweis Sei wieder

$$K(x, t) := -\sqrt{r(x) \, r(t)} \, G(x, t)$$

und dazu

$$\big(Tf\big)(x) = \int_I K(x, t) \, f(t) \, dt \qquad \big(f \in C(I)\big),$$

Eigenwerte μ_n, Eigenfunktionen $w_n = \sqrt{r}\, v_n$. Offenbar ist die Behauptung äquivalent zur gleichmäßig absoluten Konvergenz von

$$K(x,t) = \sum_{n=0}^{\infty} \mu_n w_n(x)\, w_n(t).$$

Wegen $\lambda_n \xrightarrow{n\to\infty} \infty$, $\lambda_n = \frac{1}{\mu_n}$ gibt es ein $N \in \mathbb{N}$ mit $\mu_n > 0$ für alle $n \geqslant N$. N sei im folgenden fest gewählt. Wir setzen

$$K_n(x,t) := K(x,t) - \sum_{k=0}^{n-1} \mu_k w_k(x)\, w_k(t)$$

und dazu

$$T_n\colon H \to H, \quad (T_n u)(x) = \int_I K_n(x,t)\, u(t)\, dt \qquad (u \in H = C(I)).$$

Eine kleine Rechnung liefert:

$$\langle T_n u, u\rangle = \langle Tu, u\rangle - \sum_{k=0}^{n-1} \mu_k \langle u, w_k\rangle^2.$$

Andererseits ist nach Satz 3.2.15 b)

$$Tu = \sum_{k=0}^{\infty} \mu_k \langle u, w_k\rangle\, w_k \qquad \text{(Konvergenz im quadratischen Mittel),}$$

also

$$\langle Tu, u\rangle = \sum_{k=0}^{\infty} \mu_k \langle u, w_k\rangle^2.$$

$$\begin{aligned} \Rightarrow \langle T_n u, u\rangle &= \langle Tu, u\rangle - \sum_{k=0}^{n-1} \mu_k \langle u, w_k\rangle^2 \\ &= \sum_{k=n}^{\infty} \mu_k \langle u, w_k\rangle^2. \end{aligned} \tag{3.39}$$

Behauptung: Für alle $n \geqslant N$ und $x \in I$ ist $K_n(x,x) \geqslant 0$.
Begründung: Angenommen $n \geqslant N$ und $K_n(x_0,x_0) < 0$ für ein $x_0 \in I$. Dann gibt es ein $\delta > 0$, so dass $K_n(x,t) < 0$ für alle $x,t \in I$ mit $|x - x_0| < \delta$, $|t - x_0| < \delta$. Wählt man $u \in C(I)$, so dass $u \geqslant 0$, $u \neq 0$, $\{x \in I\colon u(x) > 0\} \subset \,]x_0 - \delta, x_0 + \delta[$, so ist für $n \geqslant N$: $\langle T_n u, u\rangle < 0$, und das ist ein Widerspruch zu (3.39)
Für alle $n \geqslant N$, $x \in I$ ist also

$$0 \leqslant K_n(x,x) = K(x,x) - \sum_{k=0}^{n-1} \mu_k \big(w_k(x)\big)^2,$$

und wegen $\mu_k > 0$ für $k \geqslant N$ folgt:

$$\sum_{k=0}^{\infty} \mu_k \left(u_k(x)\right)^2 \text{ konvergiert auf } I, \tag{3.40}$$

denn die Teilsummen sind für $n \geqslant N$ monoton und beschränkt. Weiter ist für alle $n \geqslant m \geqslant N$, $t \in I$

$$\sum_{k=m}^{n} \mu_k \left(w_k(t)\right)^2 \leqslant K(t,t) - \sum_{j=0}^{N-1} \mu_j \left(w_j(t)\right)^2 \leqslant M$$

mit geeignetem $M > 0$, also für alle $x, t \in I$:

$$\begin{aligned}\left(\sum_{k=m}^{n} \mu_k \left|w_k(x)\, w_k(t)\right|\right)^2 &\overset{\text{CSU}}{\leqslant} \left(\sum_{k=m}^{n} \mu_k \left(w_k(x)\right)^2\right)\left(\sum_{k=m}^{n} \mu_k \left(w_k(t)\right)^2\right) \\ &\leqslant M \sum_{k=m}^{n} \mu_k \left(w_k(x)\right)^2 .\end{aligned} \tag{3.41}$$

(3.40) zusammen mit dem Cauchy-Kriterium hat die absolute Konvergenz der Reihe

$$\sum_{k=0}^{\infty} \mu_k w_k(x)\, w_k(t)$$

zur Folge für alle $x, t \in I$, und für jedes x ist die Konvergenz bezüglich t gleichmäßig absolut.

$$\Rightarrow K_n(x,t) \xrightarrow{n\to\infty} K(x,t) - \sum_{k=0}^{\infty} \mu_k w_k(x)\, w_k(t) =: L(x,t)\,, \tag{3.42}$$

und bei festem x ist die Konvergenz gleichmäßig bzgl. t. Für alle $x \in I$ gilt nun nach der Vollständigkeitsrelation

$$\begin{aligned}\int_I K_n^2(x,t)\, dt &= \langle K_n(x,\cdot)\,, K_n(x,\cdot)\rangle \\ &= \langle K(x,\cdot)\,, K(x,\cdot)\rangle - 2\sum_{k=0}^{n-1} \mu_k \langle K(x,\cdot)\,, w_k\rangle\, w_k(x) + \sum_{k=0}^{n-1} \mu_k^2 w_k^2(x) \\ &= \|K(x,\cdot)\|^2 - \sum_{k=0}^{n-1} \left(\langle K(x,\cdot)\,, w_k\rangle\right)^2 \xrightarrow{n\to\infty} 0\end{aligned} \tag{3.43}$$

wegen $\langle K(x,\cdot)\,, w_k\rangle = \mu_k w_k(x)$. Da bei festem x die Konvergenz bzgl. t in (3.42) gleichmäßig ist, folgt:

$$\forall\, x \in I \quad \int_I \left(L(x,t)\right)^2 dt = 0, \tag{3.44}$$

d. h. wegen (3.43)

$$\|K(x,\cdot)\|^2 = \sum_{k=0}^{\infty} \mu_k^2 \left(w_k(x)\right)^2 . \tag{3.45}$$

Da $L(x,\cdot)$ stetig ist, liefert (3.44)

$$\forall\, x \in I \quad L(x,\cdot) = 0, \qquad \text{d. h. } \forall\, x,t \in I \quad L(x,t) = 0,$$

d. h. wegen (3.44)

$$K(x,t) = \sum_{k=0}^{\infty} \mu_k w_k(x)\, w_k(t) \qquad (x,t \in I)\,.s \tag{3.46}$$

Das ist zwar die behauptete Reihenentwicklung, aber bisher wissen wir nur, dass die Reihe für jedes $x \in I$ bzgl. t gleichmäßig absolut konvergiert. Deshalb müssen wir noch einmal ausholen zum Beweis der Gleichmäßigkeit der Konvergenz bzgl. $x,t \in I$: (3.46) liefert:

$$K(x,x) = \sum_{k=0}^{\infty} \mu_k \big(w_k(x)\big)^2 \qquad (x \in I)\,. \tag{3.47}$$

Die Teilsummen der Reihe sind für $n \geqslant N$ monoton (wegen $\mu_k > 0$ für $k \geqslant N$), alle Terme sind stetig, und $x \mapsto K(x,x)$ ist stetig. Der Satz von Dini liefert dann: (3.47) konvergiert gleichmäßig auf I. Nochmalige Anwendung der Abschätzung (3.41) liefert: Das Cauchy-Kriterium für die gleichmäßig absolute Konvergenz von (3.46) ist erfüllt, d. h. (3.46) konvergiert gleichmäßig absolut, wie zu zeigen war. □

3.3.4 Korollar Sei 0 kein Eigenwert des Sturm-Liouvilleschen Eigenwertproblems. Dann gilt:

a)

$$-\int_I G(x,x)\, r(x)\, dx = \sum_{n=0}^{\infty} \frac{1}{\lambda_n}.$$

Die Reihe konvergiert absolut, da fast alle Terme positiv sind.

b)

$$\iint_{I\times I} (G(x,t))^2\, r(x)\, r(t)\, dx\, dt = \sum_{n=0}^{\infty} \frac{1}{\lambda_n^2}.$$

Beweis a) klar, da $w_k = \sqrt{r}\, v_k$ und

$$\int_I v_k^2(x)\, r(x)\, dx = 1,$$

und gliedweise Integration in (3.47) ist zulässig wegen der gleichmäßigen Konvergenz.

b)

$$\begin{aligned}\iint_{I\times I} (G(x,t))^2\, r(x)\, r(t)\, dx\, dt &= \int_I \|K(x,\cdot)\|^2\, dx \\ &\overset{(3.45)}{=} \int_I \sum_{k=0}^{\infty} \mu_k^2 \big(w_k(x)\big)^2\, dx = \sum_{k=0}^{\infty} \mu_k^2 = \sum_{k=0}^{\infty} \frac{1}{\lambda_k^2}.\end{aligned}$$

Die Reihe in der letzten Zeile ist dabei gleichmäßig absolut konvergent, da (3.47) gleichmäßig absolut konvergent ist und $\mu_k \xrightarrow{k\to\infty} 0$. □

3.3.5 Korollar Für die Eigenwerte des Sturm-Liouvilleschen Eigenwertproblems gilt:

$$\sum_{n=0}^{\infty}{}' \frac{1}{\lambda_n} \quad \text{konvergiert}$$

(auch wenn 0 Eigenwert ist!). Der Strich am Summenzeichen bedeutet: Der Eigenwert 0 ist von der Summation auszuschließen, wenn 0 Eigenwert ist.

BEWEIS Parameterverschiebung um α wie zu Beginn des Beweises von Satz 3.2.4 liefert mit Korollar 3.3.4:

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n - \alpha} \quad \text{konvergiert,}$$

und wegen $\lambda_n \xrightarrow{n\to\infty} \infty$ folgt:

$$\sum_{n=0}^{\infty}{}' \frac{1}{\lambda_n} \quad \text{konvergiert.} \qquad \square$$

Interessante Aufgaben und Beispiele finden sich noch bei Heuser (1995a), S. 419f.

Literaturverzeichnis

[Arnol'd 1979] ARNOL'D, Vladimir I.: *Satz 4.3.2.* Kap. 4, S. 223. Berlin : Deutscher Verlag der Wissenschaften, 1979 (Hochschulbücher für Mathematik 83)

[Barner und Flohr 2000] BARNER, Martin ; FLOHR, Friedrich: *De-Gruyter-Lehrbuch.* Bd. 2: *Analysis.* 3., durchges. Aufl. Berlin [u.a.] : de Gruyter, 2000

[Braun 1975] BRAUN, Martin: *Differential equations and their applications.* New York [u.a.] : Springer, 1975 (Applied mathematical sciences 15)

[Coddington und Levinson 1998] CODDINGTON, Earl A. ; LEVINSON, Norman: *Theory of ordinary differential equations.* 16. repr. Neu Delhi : Tata McGraw-Hill, 1998 (TMH edition)

[Courant und Hilbert 1924, 1937] COURANT, Richard ; HILBERT, David: *Heidelberger Taschenbücher.* Bd. I, II: *Methoden der mathematischen Physik.* Berlin u.a. : Springer, 1924, 1937. – Wesentlich veränderte englische Ausgabe 1953 bzw. 1962 unter dem Titel „Methods of mathematical physics“ Vol. I,II im Verlag Interscience Publishers, New York, erschienen

[Dieudonné 1978] DIEUDONNÉ, Jean Alexandre E.: *Treatise on Analysis.* New York u.a. : Academical Press, 1978. – aus dem Französischen, Originaltitel „Eléments d'analyse“

[Elstrodt 2002a] ELSTRODT, Jürgen: *Analysis.* 2002. – URL `http://www.muenster.de/~j-engel/Mathe/Infini.pdf`

[Elstrodt 2002b] ELSTRODT, Jürgen: *Besselsche Ungleichung.* Kap. VI, S. 236. In: *Maß- und Integrationstheorie.* Berlin; Heidelberg; New York; Barcelona; Hongkong; London; Mailand; Paris; Tokio : Springer, 2002 (Springer-Lehrbuch)

[Elstrodt 2002c] ELSTRODT, Jürgen: *Satz von der besten Approximation.* Kap. VI, S. 236. Siehe (Elstrodt, 2002b)

[Elstrodt 2003] ELSTRODT, Jürgen: *Funktionentheorie.* 2003. – URL `http://www.muenster.de/~j-engel/Mathe/FT.pdf`. – Vorlesungsmitschrift 2002/2003

[Forster 1977] FORSTER, Otto: S. 60. In: *Riemannsche Flächen.* Berlin [u.a.] : Springer, 1977 (Heidelberger Taschenbücher 184)

[Forster 1996] FORSTER, Otto: *Vieweg Studium: Grundkurs Mathematik.* Bd. 3: *Integralrechnung im $\mathbb{R}^n$ mit Anwendungen.* Kap. 18. In: *Analysis* Bd. 3. Braunschweig [u.a.] : Vieweg, 1996

[Forster 1999] FORSTER, Otto: *Vieweg Studium: Grundkurs Mathematik*. Bd. 2: *Differentialrechnung im $\mathbb{R}^n$, gewöhnliche Differentialgleichungen.* Kap. I, §6, Satz 5. In: *Analysis* Bd. 2. Braunschweig [u.a.] : Vieweg, 1999

[Gantmacher 1965] GANTMACHER, Feliks R.: *Matrizenrechnung*. 2., berichtigte Aufl. Berlin : Deutscher Verlag der Wissenschaften, 1965 (Hochschulbücher für Mathematik 36)

[Gerthsen und Meschede 2004] GERTHSEN, Christian ; MESCHEDE, Dieter: S. 63. In: *Gerthsen Physik*. Berlin [u.a.] : Springer, 2004

[Hellwig 1964] HELLWIG, G.: *Differentialoperatoren der mathematischen Physik*. Berlin u.a. : Springer, 1964

[Heuser 1990] HEUSER, Harro: *Mathematische Leitfäden*. Bd. I: *Lehrbuch der Analysis*. 7. Aufl. Stuttgart : Teubner, 1990

[Heuser 1995a] HEUSER, Harro: *Gewöhnliche Differentialgleichungen : Einführung in Lehre und Gebrauch*. 3., durchges. Aufl. Stuttgart : Teubner, 1995 (Mathematische Leitfäden)

[Heuser 1995b] HEUSER, Harro: *Gewöhnliche Differentialgleichungen : Einführung in Lehre und Gebrauch*. Kap. VI. Siehe (Heuser, 1995a)

[Hochstadt 1975] HOCHSTADT, H.: *Differential equations, A modern approach*. S. 149, Dover Publishing, 1975

[Hsieh und Sibuya 1999] HSIEH, Po-Fang ; SIBUYA, Yasutaka: *Basic theory of ordinary differential equations*. New York, Berlin, Heidelberg, Barcelona, Hong Kong, London, Milan, Paris, Singapore, Tokyo : Springer, 1999 (Universitext)

[Jörgens 1964] JÖRGENS, Konrad (Hrsg.) ; Matematisk Institut, Aarhus Universitet (Veranst.): *Spectral theory of second order ordinary differential operators*. Aarhus, 1964. – Lectures delivered at Aarhus Universitet 1962/63

[Jörgens und Rellich 1976] JÖRGENS, Konrad ; RELLICH, Franz: *Eigenwerttheorie gewöhnlicher Differentialgleichungen*. Berlin, Heidelberg, New York : Springer, 1976 (Hochschultext)

[Kamke 1950] KAMKE, Erich: *Differentialgleichungen reeller Funktionen*. 2. verb. Aufl. Leipzig : Geest & Portig, 1950 (Mathematik und ihre Anwendungen in Physik und Technik : Reihe A 7)

[Knobloch und Kappel 1974a] KNOBLOCH, Hans W. ; KAPPEL, Franz: *Gewöhnliche Differentialgleichungen*. 1. Aufl. Stuttgart : Teubner, 1974 (Mathematische Leitfäden)

[Knobloch und Kappel 1974b] KNOBLOCH, Hans W. ; KAPPEL, Franz: *Satz 4.1*. Kap. 4, S. 129–132. Siehe (Knobloch und Kappel, 1974a)

[Königsberger 1903] KÖNIGSBERGER, Leo: *Gesammelte Schriften*. Bd. 2. Braunschweig : Vieweg, 1903

[Magnus u. a. 1966] MAGNUS, Wilhelm ; OBERHETTINGER, Fritz ; SONI, Raj P.: *Formulas and theorems for the special functions of mathematical physics.* 3., enl. ed. Berlin, Heidelberg, New York : Springer, 1966 (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen 52)

[Peyerimhoff 1970] PEYERIMHOFF, Alexander: *Studien-Texte.* Bd. II: . Kap. VI. In: *Gewöhnliche Differentialgleichungen* Bd. II. Frankfurt a. M. : Akademische Verlagsgesellschaft, 1970

[Roelcke 1961] ROELCKE, Walter: *Differentialgleichungen.* 1961. – Vorlesungsausarbeitung

[Stepanov 1976] STEPANOV, Vjačeslav V.: *Lehrbuch der Differentialgleichungen.* 4. Aufl. Berlin : Deutscher Verlag der Wissenschaften, 1976 (Hochschulbücher für Mathematik 20)

[Storch und Wiebe 1999] STORCH, Uwe ; WIEBE, Hartmut: *Lineare Algebra.* Bd. 2. Kap. 20.D, 20.E. In: *Lehrbuch der Mathematik* Bd. 2. Heidelberg [u.a.] : Spektrum Akademischer Verlag, 1999

[Stromberg 1996] STROMBERG, Karl R.: *An introduction to classical real analysis.* London [u.a.] : Chapman & Hall, 1996

[Szegö 1998] SZEGÖ, Gábor: *Orthogonal polynomials.* Providence, RI : American Mathematical Society, 1998 (Colloquium publications 23)

[Titchmarsh 1962, 1958] TITCHMARSH, Edward C.: *Eigenfunction expansions associated with second order differential equations.* Bd. I und II. Oxford : Clarendon Press, 1962, 1958

[Tricomi 1955] TRICOMI, Francesco G.: *Vorlesungen über Orthogonalreihen.* Berlin u.a. : Springer, 1955 (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete 76). – Aus d. Italien. übers.

[Walter 1976a] WALTER, Wolfgang: *Gewöhnliche Differentialgleichungen: eine Einführung.* 2., korr. Aufl. Berlin [u.a.] : Springer, 1976 (Heidelberger Taschenbücher 110)

[Walter 1976b] WALTER, Wolfgang: *Gewöhnliche Differentialgleichungen: eine Einführung.* Kap. V. Siehe (Walter, 1976a)

[Walter 1976c] WALTER, Wolfgang: *Satz.* Kap. XI, S. 102f. Siehe (Walter, 1976a)

[Walter 1976d] WALTER, Wolfgang: *Satz X.* Kap. 13, S. 101. Siehe (Walter, 1976a)

[Weidmann 1987a] WEIDMANN, Joachim: *Spectral theory of ordinary differential operators.* Kap. 13, S. 194ff. Siehe (Weidmann, 1987b)

[Weidmann 1987b] WEIDMANN, Joachim: *Spectral theory of ordinary differential operators.* Berlin, Heidelberg, New York : Springer, 1987 (Lecture notes in mathematics 1258)

[Weidmann 2003] WEIDMANN, Joachim: *Anwendungen.* Kap. II. In: *Lineare Operatoren in Hilberträumen.* Stuttgart, Leipzig, Wiesbaden : Teubner, 2003

[Weise 1966] WEISE, Karl H.: Kap. IV. In: *Differentialgleichungen.* Göttingen : Vandenhoeck & Ruprecht, 1966 (Studia mathematica 17)

Index

Abbildungsverzeichnis

Inhaltsverzeichnis