

Mathematik III für Physiker*

Sommersemester 2013

Christian Hainzl mit Vorlage von Stefan Teufel, Simon Mayer und Mario Laux
Mathematisches Institut
Uni Tübingen

16. Dezember 2013

*Diese vorläufige Version des Skriptums ist nur zum Gebrauch parallel zum Besuch der Vorlesung gedacht. Das Studium des Skripts kann den Besuch der Vorlesung **nicht** ersetzen! Falls Sie Fehler finden, teilen Sie mir diese (auch die offensichtlichen) bitte mit!

Inhaltsverzeichnis

1	Gewöhnliche Differentialgleichungen und dynamische Systeme	1
1.1	Einleitung	1
1.1.1	Explizite Differentialgleichungen erster Ordnung	3
1.1.2	Lineare Differentialgleichungen erster Ordnung	5
1.1.3	Die exakte Differentialgleichung	6
1.2	Dynamische Systeme und Flüsse	8
2	Fourier-Analysis	23
2.0.1	Einleitung	23
2.1	Fourier-Reihen	24
2.1.1	Integrkern	30
2.1.2	Wiederholung: Hilbertraum und Skalarprodukt	32
2.1.3	L^2 -Konvergenz von Fourier-Reihen	33
2.1.4	Fourier-Transformation	37
2.1.5	Anwendungsbeispiele von Fourier-Transformationen	41
3	Ausflug in Lineare Algebra	43
3.1	Vektorräume mit Skalarprodukt	43
3.2	Symmetrische Operatoren	57
3.3	Klassifikation von Matrizen	62
4	Lineare Differentialgleichungen	71
4.1	Qualitative Theorie autonomer Systeme	80

1 Gewöhnliche Differentialgleichungen und dynamische Systeme

1.1 Einleitung

Wir betrachten die Differentialgleichung erster Ordnung

$$\dot{x} = f(t, x), \tag{1.1}$$

wobei $f(t, x)$ eine stetige Funktion von $D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ ist.

Für eine Lösung $x(t)$ auf dem Intervall I ist definitionsgemäß

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

die Steigung \dot{x} der Lösungskurve $x(t)$ wird also in jedem ihrer Punkte gegeben durch $f(t, x(t))$. Das heißt in (t, x) angekommen wird die sie mit der Steigung $f(t, x)$ weitergeschickt. Die Differentialgleichung dirigiert die Kurve durch ständige Richtungsangaben.

Das Anfangswertproblem

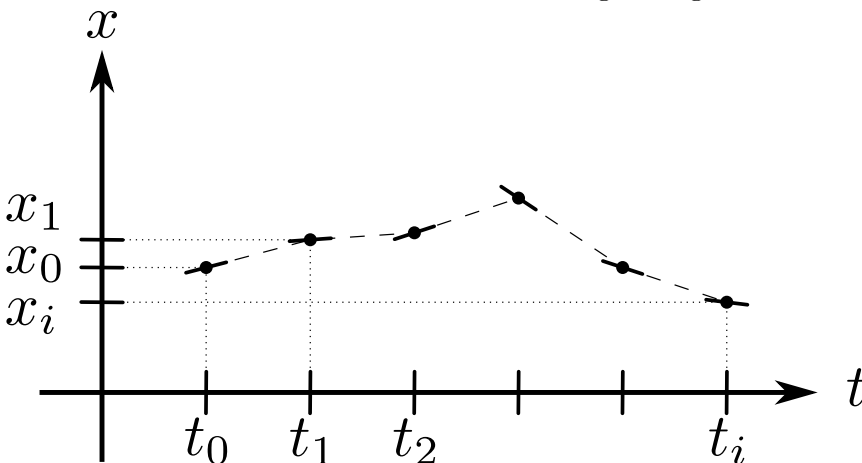
$$\dot{x} = f(t, x), \quad x(t_0) = x_0 \tag{1.2}$$

zu lösen, bedeutet, eine Kurve zu finden, die vom Punkt (t_0, x_0) ausgeht und in ihrem ganzen Verlauf auf das Richtungsfeld passt.

Man kann numerisch leicht näherungsweise eine Lösung konstruieren, indem man I in Teilintervalle gleicher Länge mit den Punkten $t_0, t_1, \dots, t_n, \dots$ teilt, bei t_0 mit dem Wert x_0 beginnt und im Intervall $[t_0, t_1]$ eine Gerade mit der Steigung $f(t_0, x_0)$ legt, sodass

$$x_1 = x_0 + f(t_0, x_0)(t_1 - t_0),$$

und so weiter. So bekommt man eine Näherungslösung.



In höheren Dimensionen ist die rechte Seite der Differentialgleichung ein Vektorfeld. In vielen Anwendungen der Physik ist die rechte Seite unabhängig von der Zeit: Sei $v : G \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Sei $I \subset \mathbb{R}$ ein offenes Intervall. Man nennt eine stetig differenzierbare Kurve $x : I \rightarrow G$, $t \mapsto x(t)$, eine **Lösung der Differentialgleichung**

$$\dot{x} = v(x),$$

wenn für alle $t \in I$ gilt

$$\dot{x}(t) = v(x(t)).$$

Falls $0 \in I$ ist, so heißt $x(t)$ eine **Lösung zum Anfangswert** $x_0 \in G$, wobei $x_0 = x(0)$ ist.

Die Lösungskurven $x(t)$ sind überall tangential an v und heißen **Integalkurven** an das Vektorfeld v . Beim Lösen einer Differentialgleichung spricht man auch vom „Integrieren der Gleichung“.

1.1 Beispiel. Betrachten wir als Beispiel eine lineare Gleichung höherer Ordnung. Die Differentialgleichungen

$$\begin{aligned} \dot{x}_1(t) &= a_{11}x_1(t) + \dots + a_{1n}x_n(t) \\ &\vdots \\ \dot{x}_n(t) &= a_{n1}x_1(t) + \dots + a_{nn}x_n(t) \end{aligned}$$

sind gekoppelt und in Kurzform schreiben wir

$$\dot{x} = Ax.$$

Hier ist also

$$v(x) = Ax.$$

Es sei die Anfangsbedingung $x(0) = a \in \mathbb{R}^n$ gegeben und wir suchen die Lösung für alle anderen Zeiten, also eine Funktion $x : \mathbb{R} \rightarrow \mathbb{R}^n$ welche $\frac{d}{dt}x(t) = Ax(t)$ und $x(0) = x_0$ erfüllt.

Falls A diagonalisierbar ist, also

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

für eine reguläre Matrix S gilt, dann ergibt sich für den Vektor $y(t) = S^{-1}x(t)$

$$\dot{y}(t) = S^{-1}\dot{x}(t) = S^{-1}Ax(t) = S^{-1}ASS^{-1}x(t) = S^{-1}ASy(t) = Dy(t).$$

Wir erhalten also die entkoppelten Differentialgleichungen

$$\dot{y}_1(t) = \lambda_1 y_1(t), \dots, \dot{y}_n(t) = \lambda_n y_n(t).$$

Die Anfangsdaten für y ergeben sich aus denen für x durch

$$y(0) = S^{-1}x(0) = S^{-1}x_0 =: b.$$

Die Lösungen für das entkoppelte System kann man direkt ablesen,

$$y_1(t) = b_1 e^{\lambda_1 t}, \dots, y_n(t) = b_n e^{\lambda_n t}$$

oder kurz

$$y(t) = e^{Dt} b := \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Durch Rücktransformation bekommt man auch wieder

$$x(t) = Sy(t) = Se^{Dt}b = Se^{Dt}S^{-1}a =: e^{At}x_0.$$

Wir haben im letzten Beispiel eine explizite Lösung $x(t) = e^{tA}x_0$ konstruiert, wobei $e^{tA} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ fuer alle t ein Diffeomorphismus ist. Wir werden später

$$\varphi_t = e^{tA},$$

als Fluss betrachten, der die Gruppeneigenschaften

$$\varphi_t \circ \varphi_s = \varphi_{s+t}, \quad \varphi_t \circ \varphi_{-t} = \varphi_0 = \mathbb{1},$$

erfüllt, sodass

$$\dot{\varphi}_t(x_0) = v(\varphi_t(x_0)).$$

1.1.1 Explizite Differentialgleichungen erster Ordnung

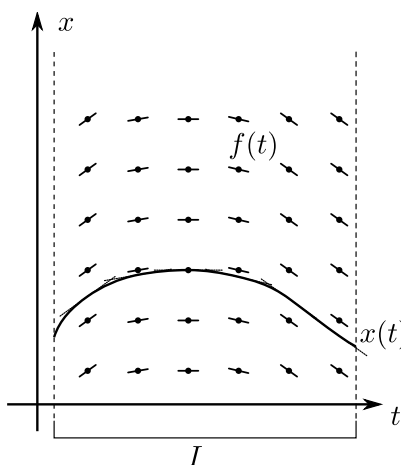
$$\dot{x} = f(t)$$

Sei $f(t)$ eine stetige Funktion auf dem Intervall I . Dann ist das zugehörige Anfangswertproblem

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

wohldefiniert auf

$$D = I \times \mathbb{R}.$$



Mittels Hauptsatz der Integralrechnung

$$x(t) = \int_{t_0}^t f(s)ds + x_0. \quad (1.3)$$

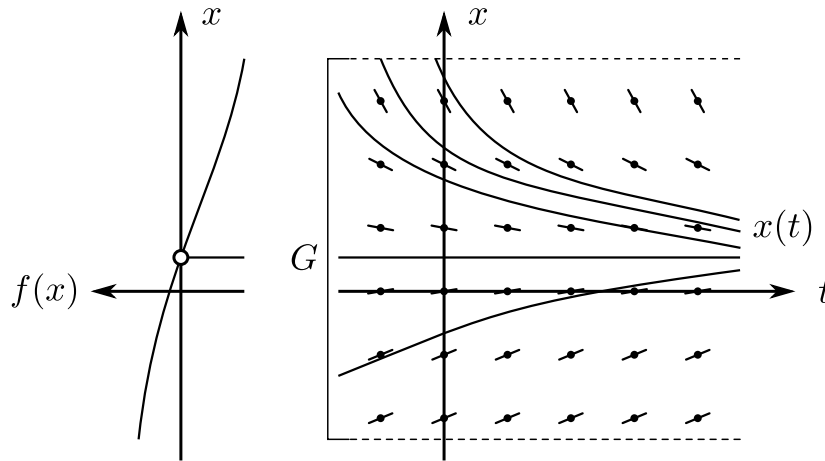
Es ist leicht zu sehen, dass $x(t)$ nicht die Flusseigenschaften erfüllen kann, da die Gleichung explizit von der Zeit abhängt. Man sagt dann das System ist nicht autonom. Man kann aber ein dynamisches System daraus machen, indem man es wie folgt umschreibt:

$$\begin{aligned} \frac{d}{ds}t(s) &= 1 \\ \frac{d}{ds}x(s) &= f(t) \end{aligned}$$

sodass es die Form $\dot{z} = v(z)$, mit $z = (t, x)$, und $v(z) = (1, f(t))$. Das System generiert nun Fluss φ_s , mit $\varphi_s : I \times \mathbb{R} \rightarrow I \times \mathbb{R}$, wobei $\varphi_s(t_0, x_0)$ die Lösung (1.3) liefert.

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

Sei $v(x)$ stetig auf Intervall G , dann ist das Richtungsfeld ähnlich wie zuvor, nur dass die Rollen von x und t vertauscht sind. Dies schlägt vor, die Rollen von x und t zu vertauschen und eine Lösung $t(x)$ hinzuschreiben.



Wenn $v(x) \neq 0$ in einer Umgebung von x_0 , dann kann dort das Anfangswertproblem

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

in

$$\frac{d}{dx}t(x) = 1/v(x), \quad t(x_0) = t_0$$

umgeschrieben und wie in (1.3) gelöst werden, und die Umkehrfunktion $x(t) := t^{-1}$, erfüllt dann

$$\dot{x} = \frac{1}{t'(t^{-1})} = v(x(t)).$$

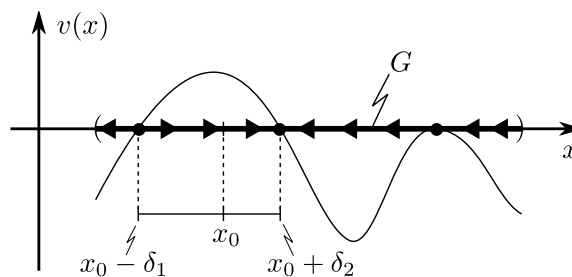
1.2 Bemerkung. Der Hauptsatz zeigt die Berechtigung der heuristischen Rechnung

$$\dot{x} = v(x) \Rightarrow \frac{dx}{dt} = v(x) \stackrel{(*)}{\Rightarrow} \frac{dx}{v(x)} = dt \Rightarrow \int_{x_0}^x \frac{du}{v(u)} = \int_0^t dt \Rightarrow \int_{x_0}^x \frac{du}{v(u)} = t.$$

Den Schritt (*) nennt man oft „Trennung der Variablen“.

Hat $v(x)$ eine Nullstelle η , dann ist die konstante Funktion $x(t) = \eta$ eine stationäre Lösung der Gleichung. Zwischen den Nullstellen lässt sich die Gleichung durch invertieren lösen und so lassen sich alle Lösungen finden, zumindest wenn $v(x)$ so ist, dass die Teilchen in endlicher Zeit nie die Gleichgewichtslagen $v(x) = 0$ erreichen.

In den Nullstellen von v ruht das System. Dazwischen läuft es monoton von einer Nullstelle zur nächsten. Ist v an den Nullstellen differenzierbar, so werden die Gleichgewichtslagen in endlicher Zeit nicht erreicht. An den Rändern des Intervalls G können die Orbits in endlicher Zeit entweichen.



1.3 Beispiele. (a) Sei $a \in \mathbb{R}$ und $v : \mathbb{R} \rightarrow \mathbb{R}$, $v(x) = ax$ ein lineares Vektorfeld. Die Lösung von $\dot{x} = v(x)$ auf $G = \mathbb{R}$ zum Anfangswert $x_0 > 0$ für $a \neq 0$ ist nach Proposition 1.23 für $x > 0$ gegeben durch

$$t(x) = \int_{x_0}^x \frac{du}{au} = \frac{1}{a} (\ln x - \ln x_0) = \frac{1}{a} \ln \frac{x}{x_0}$$

also

$$\ln \frac{x}{x_0} = at \Rightarrow x(t) = x_0 e^{at} \quad \forall t \in \mathbb{R}.$$

Diese Formel gilt auch für $x_0 \leq 0$ und somit ist der Fluss $\varphi^t : \mathbb{R} \rightarrow \mathbb{R}$, $\varphi^t(x_0) = e^{at} x_0$ global.

(b) Sei $v : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $v(x) = \sqrt{|x|}$. Dann erreichen die Lösungen zu Startwerten $x(0) < 0$ in endlicher Zeit die Ruhelage $x = 0$. Dort können Sie beliebig lange verweilen und dann nach rechts wieder aus der Ruhelage herauslaufen. Die Lösungen sind also nicht eindeutig. Siehe Übungen.

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

Das Verfahren der Trennung der Variable funktioniert auch für den Fall, dass v zeitabhängig und von der Form

$$v : I \times G \rightarrow \mathbb{R}, \quad v(t, x) = a(t)b(x)$$

ist, wobei $I \subset \mathbb{R}$ ein offenes Intervall und a und b jeweils stetig sind. Dann liefert Auflösen von

$$\int_{t_0}^{t(x)} a(\tau) d\tau = \int_{x_0}^x \frac{du}{b(u)} \quad (1.4)$$

nach $t(x)$ analog zu oben eine lokale Lösung von

$$\dot{x}(t) = v(t, x(t)) \quad \text{zum Anfangswert } x(t_0) = x_0.$$

1.1.2 Lineare Differentialgleichungen erster Ordnung

Eindimensionale lineare Differentialgleichungen, die Wachstumsraten und andere Vorgänge aus Biologie oder Physik, beschreiben, sind von der Form

$$\dot{x} = g(t)x, \quad (1.5)$$

oder

$$\dot{x} = g(t)x + f(t), \quad (1.6)$$

wobei (1.5) *homogen* und (1.6) *inhomogen* genannt wird. Die Funktion $f(t)$ wird Störfunktion genannt.

Sei nun $g(t)$ auf einem Intervall I stetig, so besitzt sie dort auch eine Stammfunktion und nach (1.1.1) lässt sich die Lösung des zugehörigen Anfangswertproblems $\dot{x} = g(t)x$, $x(t_0) = x_0$, als

$$x(t) = x_0 e^{\int_{t_0}^t g(\tau) d\tau}$$

schreiben. Etwas allgemeiner lässt sich sagen, dass genau die Funktionen

$$z(t) = C e^{\int g(t) dt}$$

alle Lösungen von (1.5) sind. Denn ist $x(t)$ eine beliebige Lösung von (1.5) dann gilt

$$\frac{d}{dt} \frac{x}{z} = \frac{x'z - z'x}{z^2} = \frac{g(xz) - g(zx)}{z^2} = 0.$$

Also muss x/z eine Konstante sein.

Nun nehmen wir uns die inhomogene Gleichung (1.6) vor. Ist nun $x_p(t)$ irgend eine (partikuläre) Lösung von (1.6), so gilt für jede andere Lösung $x(t)$ von (1.6) dass, für $y = x - x_p$

$$y'(t) = (x(t) - x_p(t))' = g(t)x(t) + f(t) - (g(t)x_p(t) + f(t)) = g(t)(x - x_p) = g(t)y(t),$$

also y eine Lösung von (1.5) ist. Infolgedessen muss $x(t)$ die Gestalt

$$x(t) = x_p(t) + C e^{\int g(t) dt}.$$

Das Problem **alle** Lösungen für (1.6) zu finden reduziert sich auf die Aufgabe, eine spezielle Lösung für (1.6) zu finden. Wenn f stetig ist, kann man das mittels der Methode der Variation der Konstante machen.

Ansatz:

$$x_p(t) = C(t) e^{\int_{t_0}^t g(s) ds}.$$

Da x_p die Gleichung erfüllen muss, gilt

$$g(t)x_p(t) + f(t) = \dot{x}_p(t) = \dot{C}(t) e^{\int_{t_0}^t g(s) ds} + C(t)g(t) e^{\int_{t_0}^t g(s) ds}$$

sodass

$$\dot{C}(t) = f(t)e^{-\int_{t_0}^t g(s)ds}$$

gelten muss, was wiederum durch die Stammfunktion

$$C(t) = \int_{t_0}^t f(\tau)e^{-\int_{t_0}^{\tau} g(s)ds} d\tau$$

integriert werden kann. Wir bekommen. Somit wissen wir, dass *allgemeine Lösung der inhomogenen Gleichung = partikuläre Lösung der inhomogenen Gleichung + allgemeine Lösung der zugehörigen homogenen Gleichung.*

1.4 Lemma. Sind $f(t)$ und $g(t)$ auf dem Intervall I stetig, so besitzt das Anfangswertproblem

$$\dot{x} = g(t)x + f(t), \quad x(t_0) = x_0 \tag{1.7}$$

eine genau eine Lösung.

1.5 Beispiel.

$$\dot{x} = (\sin t)x + \sin t, \quad x(0) = 0.$$

Step 1: Die allgemeine Lösung der homogenen Gleichung ist offensichtlich $x_h(t) = Ce^{-\cos t}$.

Step 2: Bestimmung der partikulären Lösung durch $x_p = C(t)e^{-\cos t}$, mit $\dot{C}(t) = \sin te^{\cos t}$, und damit $C(t) = -e^{\cos t}$, und $x_p(t) = -1$.

Step 3: Anpassung der freien Konstanten der homogenen Lösung an die Anfangsbedingung. Sei $x(t) = Ce^{-\cos t} - 1$, dann gilt $x(0) = Ce - 1 = x_0$, also $C = e^{-1}(x_0 + 1)$.

1.1.3 Die exakte Differentialgleichung

Sei $U : \mathbb{R}^2 \rightarrow \mathbb{R}$. Sei $(x(t), y(t))$ eine Paramterisierung der Äquipotentiallinien von $U(x, y)$, sprich der Kurven $U(x, y) = c$. Dann gilt

$$\frac{d}{dt}U(x(t), y(t)) = U_x \dot{x} + U_y \dot{y} = 0,$$

und wenn, zum Beispiel $\dot{x} \neq 0$ ist, dann gilt auch

$$U_x + \frac{dy}{dx}U_y = 0.$$

Beziehungsweise,

$$\frac{dx}{dy}U_x + U_y = 0$$

falls $\dot{y} \neq 0$. Sei nun

$$\nabla U = F = (P, Q) : \mathbb{R}^2 \rightarrow \mathbb{R}$$

dann lassen sich die beiden Gleichungen aus Symmetriegründen am besten durch

$$P(x, y)dx + Q(x, y)dy = 0 \tag{1.8}$$

schreiben.

Normalerweise ist die Problemstellung umgekehrt. Wir haben eine Diffgleichung der Form

$$\frac{dy}{dx}U_x + U_y = 0$$

oder

$$P(x, y)dx + Q(x, y)dy = 0$$

vorgegeben, und wir fragen ob wir die Lösungskurven mittels einer Potentialfunktion U angeben können.

Man nennt nun ganz allgemein die Gleichung (1.8) **exakt**, falls $F = (P, Q)$ eine Potentialfunktion U hat. Lösungen dieser Gleichung sind durch die Äquipotentiallinien gegeben.

1.6 Beispiel.

$$2x \sin y dx + x^2 \cos y dy = 0, \quad y(1) = \pi/4.$$

Die Differentialgleichung ist auf \mathbb{R}^2 exakt, denn sie hat dort die Stammfunktion $U(x, y) = x^2 \sin y$. Infolgedessen geht es nur darum die Gleichung

$$x^2 \sin y = U(1, \pi/4) = 1/\sqrt{2}$$

differenzierbar nach y aufzulösen, oder nach x aufzulösen.

Es drängen sich nun drei Fragen auf:

1. Wie kann man der Gleichung (1.8) ansehen, dass sie exakt ist?
2. Wie kann man, falls sie exakt ist, eine Stammfunktion bestimmen?
3. Unter welchen Voraussetzungen kann man $U(x, y) = c$ differenzierbar nach y auflösen?

Die Fragen (1) und (3) wurden in Mathe 2 für Physiker beantwortet und sind in dem folgenden Satz zusammengefasst.

1.7 Satz. Sind die partiellen Ableitungen von P, Q auf einem sternförmigen Gebiet $D \subset \mathbb{R}^2$ vorhanden und stetig, dann ist die Differentialgleichung

$$P(x, y)dx + Q(x, y)dy = 0$$

genau dann exakt, wenn die Integrabilitätsbedingung $P_y = Q_x$ erfüllt ist.

Weiters ist die Anfangswertaufgabe

$$P(x, y)dx + Q(x, y)dy = 0, \quad y(x_0) = y_0$$

in einer hinreichenden Umgebung von x_0 eindeutig lösbar, wenn $Q(x_0, y_0) \neq 0$ ist. Eine analoge Aussage gilt für das Anfangswertproblem

$$P(x, y)dx + Q(x, y)dy = 0, \quad x(y_0) = x_0.$$

Beweis. Hier eine Wiederholung des des Argumentes, wie man die Sei a so dass jeder Punkt in D durch Gerade erreichbar ist. Bezeichne $x = (x_1, x_2)$, und $\mathbf{F} = (F_1, F_2) = (P, Q)$. Die Gerade sei parametrisiert durch $\gamma(t) = a + t(x - a)$.

$$\begin{aligned} U(x) &= \int_a^x \mathbf{F}(z) \cdot dz \\ &= \int_0^1 \mathbf{F}(a + t(x - a)) \cdot (x - a) dt \\ &= \sum_{k=1}^2 \int_0^1 F_k(a + t(x - a)) \cdot (x_k - a_k) dt \end{aligned}$$

$$\begin{aligned}
 \frac{\partial}{\partial x_j} U(x) &= \int_0^1 F_j(a + t(x - a)) dt + \sum_{k=1}^2 \int_0^1 \left(\frac{\partial}{\partial x_j} F_k \right) (a + t(x - a)) t \cdot (x_k - a_k) dt \\
 &= \int_0^1 F_j(a + t(x - a)) dt + \sum_{k=1}^2 \int_0^1 \left(\frac{\partial}{\partial x_k} F_j \right) (a + t(x - a)) \cdot (x_k - a_k) t dt \\
 &= \int_0^1 F_j(a + t(x - a)) dt + \int_0^1 \left(\frac{d}{dt} F_j(a + t(x - a)) \right) t dt \\
 &= \int_0^1 F_j(a + t(x - a)) dt + F_j(a + t(x - a)) t \Big|_{t=0}^1 - \int_0^1 F_j(a + t(x - a)) dt \\
 &= F_j(x).
 \end{aligned}$$

Der zweite Teil der Aussage ist eine Konsequenz des Hauptsatzes über implizite Funktionen. $U(x, y)$ ist nach y auflösbar in Umgebung von (x_0, y_0) wenn $Q(x_0, y_0) = U_y(x_0, y_0) \neq 0$ ist. \square

1.8 Beispiel.

$$(12xy + 3)dx + 6x^2dy = 0, \quad y(1) = 1.$$

Ist die Differentialgleichung

$$P(x, y)dx + Q(x, y)dy = 0$$

nicht exakt, so kann man mittels Multiplikation einer nichtverschwindenden Funktion, eines integrierenden Faktors, $M(x, y)$ versuchen, sie zu einer zu machen

$$M(x, y)P(x, y)dx + M(x, y)Q(x, y)dy = 0,$$

wobei

$$\frac{d}{dy}(M(x, y)P(x, y)) = \frac{d}{dx}(M(x, y)Q(x, y)). \tag{1.9}$$

1.9 Beispiel.

$$(4x + 3y^2)dx + 2xydy = 0$$

Die Gleichung ist nicht exakt. Man kann nun versuchen einen integrierenden Faktor M einzuführen. Es empfiehlt sich eine Funktion zu nehmen, die nur von einer Variable abhängt, z. B., $M(x)$. Um (1.9) zu erfüllen, muss M der Gleichung $xM' = 2M$ genügen, mit Lösung $M = x^2$. Damit wird die Gleichung exakt mit der Stammfunktion $U = x^4 + x^3y^2$. Die Potentiallinien erfüllen die ursprüngliche Gleichung, solange $x \neq 0$ ist, also auf den beiden Halbebenen.

1.2 Dynamische Systeme und Flüsse

Wir betrachten im folgenden dynamische Systeme, die von Differentialgleichungen erster Ordnung erzeugt werden. Wir werden primär an qualitativen Aussagen interessiert sein, wie in der Mathematik so üblich. Dass dies einen Grossteil der Systeme in der Physik abdeckt, folgt aus der Tatsache, dass Gleichungen höherer Ordnung wie die Newton's einfach auf ein System erster Ordnung reduziert werden kann, wie im folgenden Beispiel ersichtlich.

1.10 Beispiel. Klassische Mechanik

Bewegt sich ein Teilchen der Masse $m > 0$ in einem Gebiet $D \subset \mathbb{R}^3$ unter dem Einfluss einer Kraft $K : D \rightarrow \mathbb{R}^3$, so löst die Bahnkurve des Teilchens $t \mapsto x(t)$ die Gleichung

$$m \ddot{x}(t) = F(x(t)) \quad (\text{Masse} \times \text{Beschleunigung} = \text{Kraft}).$$

Setzen wir $G : D \times \mathbb{R}^3 \subset \mathbb{R}^6$ und $v : G \rightarrow \mathbb{R}^6$,

$$v(x, y) = \begin{pmatrix} y \\ \frac{1}{m}F(x) \end{pmatrix}, \quad \text{sowie} \quad z(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix},$$

so löst $z(t)$ die Gleichung

$$\dot{z}(t) = v(z(t)),$$

denn

$$\dot{z}(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{\dot{x}}(t) \end{pmatrix} = \begin{pmatrix} \dot{x}(t) \\ \frac{1}{m}F(x(t)) = v(z(t)) \end{pmatrix}.$$

Es ist also z die Lösung der gewöhnlichen Differentialgleichung

$$\dot{z} = v(z).$$

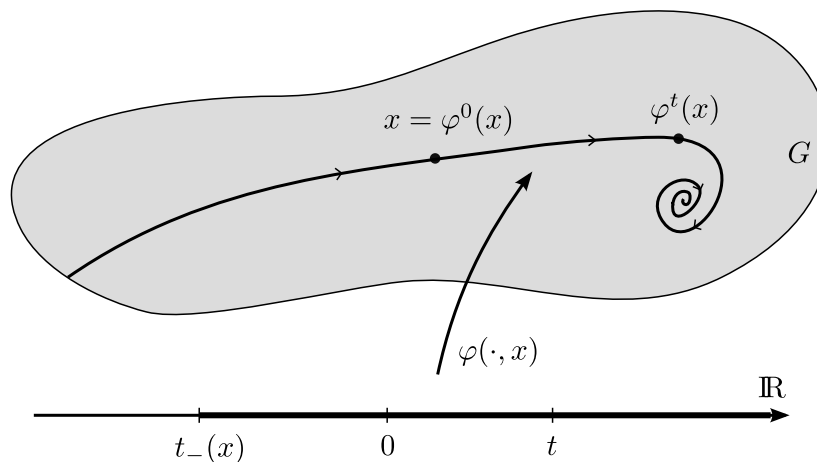
Bestimmt man also zum Vektorfeld $v : G \rightarrow \mathbb{R}^6$ alle Lösungskurven $z : I(z) \rightarrow G$, so erhält man durch Projektion auf die ersten 3 Koordinaten alle Teilchenbahnen $x : I(z) \rightarrow D$ im Kraftfeld F .

Wie wir sehen werden bestimmen diese Lösungskurven ein dynamisches System φ . Oder, besser, einen Fluss am Phasenraum $G \subset \mathbb{R}^6$.

In einem deterministischen dynamischen System kann man jedem Zustand \cong Punkt x im Phasenraum G eine eindeutige zeitliche Entwicklung zuordnen: Ist das System zum Zeitpunkt $t_0 = 0$ im Zustand $x \in G$, so sei der Zustand zum Zeitpunkt t gleich $\varphi(t, x)$. Dabei kann das "Existenzintervall" $I(x) \ni 0$, auf dem die Abbildung $\varphi(\cdot, x) : I(x) \rightarrow G$ erklärt ist, von x abhängen. Wir schreiben oft $\varphi^t(x) := \varphi(t, x)$ und verlangen $\varphi^0(x) = x$ und die Verträglichkeitsbedingung

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$$

für alle $t \in I(x)$ und $s \in I(\varphi^t(x))$.



1.11 Definition. Dynamisches System

Ein **dynamisches System** auf einem Gebiet $G \subset \mathbb{R}^n$ ist gegeben durch eine stetig differenzierbare Abbildung $\varphi : \Omega \rightarrow G$, $(t, x) \mapsto \varphi^t(x)$ mit folgenden Eigenschaften:

- (a) $\Omega \subset \mathbb{R} \times G$ ist ein Gebiet, so dass $\{0\} \times G \subset \Omega$ ist und für jedes $x \in G$ ist

$$I(x) := \{t \in \mathbb{R} \mid (t, x) \in \Omega\} = \{t \mid \varphi^t(x) \in G\} \subset \mathbb{R}$$

ein Intervall, das sogenannte Existenzintervall zum Startpunkt x .

- (b) (i) Für alle $x \in G$ ist $\varphi^0(x) = x$;
(ii) für alle $x \in G$ und $t \in I(x)$ gilt:

Es ist $s \in I(\varphi^t(x))$ genau wenn $s + t \in I(x)$. In diesem Fall gilt

$$\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x).$$

1.12 Bemerkung. (a) Man nennt $G \subset \mathbb{R}^n$ den **Phasenraum** des dynamischen Systems φ .

(b) Weil $\Omega \subset \mathbb{R} \times G \subset \mathbb{R}^{n+1}$ offen ist, ist auch $I(x) \subset \mathbb{R}$ offen und wegen Definition 1.11 (a) ein offenes Intervall,

$$I(x) =: (t_-(x), t_+(x)).$$

Hierbei ist $t_-(x) \in [-\infty, 0)$ und $t_+(x) \in (0, +\infty]$. Wir nennen $\varphi(x) := \{\varphi^t(x) \mid t \in I(x)\}$ den Orbit von x und $I(x)$ sei Existenzintervall.

(c) Setzen wir für jedes $t \in \mathbb{R}$

$$G_t := \{x \in G \mid (t, x) \in \Omega\}$$

(G_t ist also die Menge derjenigen $x \in G$, deren "Zeitentwicklung" mindestens bis zur Zeit t existiert), so folgt aus Definition 1.11 (b), dass

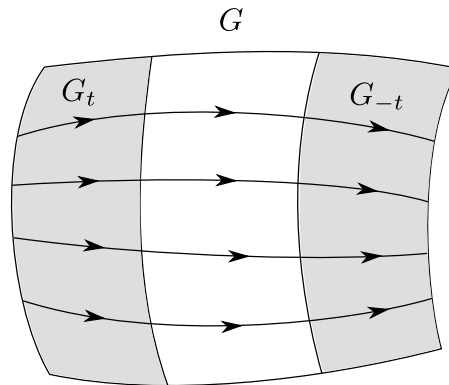
$$\varphi^t(G_t) = \{\varphi^t(x) \mid x \in G_t\} = G_{-t}$$

gilt und, dass

$$\varphi^t : G_t \rightarrow G_{-t}$$

ein Diffeomorphismus ist, denn $\varphi^{-t} : G_{-t} \rightarrow G_t$ ist sein Inverses:

$$\begin{aligned} \varphi^{-t} \circ \varphi^t(x) &= \varphi^{-t+t}(x) = \varphi^0(x) = x \\ &= \varphi^t \circ \varphi^{-t}(x). \end{aligned}$$



Man nennt die Familie $(\varphi^t)_{t \in \mathbb{R}}$ von Diffeomorphismen den **Fluss** von φ .

Statt $\varphi^t(x)$ schreibt man oft auch kurz $x(t)$. Nachteil: In dieser Notation ist der Anfangspunkt $\varphi^0(x) = x = x(0)$ nicht mehr explizit enthalten.

(d) Besonders schön ist die Situation, wenn $I(x) = \mathbb{R}$ ist, für alle $x \in G$, also $G_t = G$ für alle $t \in \mathbb{R}$. In diesem Fall ist $\Omega = \mathbb{R} \times G$ und man spricht von einem **globalen dynamischen System** bzw. einem globalen Fluss auf G . Es ist dann also

$$\mathbb{R} \rightarrow \text{Diff}(G), t \mapsto \varphi^t$$

ein Gruppenhomomorphismus von $(\mathbb{R}, +)$ in die **Diffeomorphismengruppe** von G ,

$$\text{Diff}(G) = \{\varphi : G \rightarrow G \mid \varphi \text{ ist Diffeomorphismus}\}.$$

Umgekehrt definiert jeder stetig differenzierbare Homomorphismus $\varphi : \mathbb{R} \rightarrow \text{Diff}(G)$ ein globales dynamisches System auf G .

1.13 Definition. Vektorfeld eines Flusses

Sei φ ein dynamisches System auf einem Gebiet $G \subset \mathbb{R}^n$. Dann heißt

$$v : G \rightarrow \mathbb{R}^n, \quad x \mapsto v(x) = \left. \frac{d}{dt} \varphi^t(x) \right|_{t=0} =: \dot{x}(0)$$

das **Vektorfeld** von φ auf G .

1.14 Bemerkung. Sei φ ein dynamisches System auf $G \subset \mathbb{R}^n$ mit Vektorfeld $v : G \rightarrow \mathbb{R}^n$. Dann gilt für alle $x \in G$ und alle $t \in I(x)$

$$\frac{d}{dt} \varphi^t(x) = v(\varphi^t(x))$$

oder kurz

$$\dot{x}(t) = v(x(t)).$$

Das Vektorfeld selbst hängt also **nicht** von der Zeit ab.

Beweis. Seien $x \in G$ und $t \in I(x)$ beliebig. Mit $y = \varphi^t(x)$ gilt nun:

$$v(\varphi^t(x)) = v(y) = \frac{d}{ds} \varphi^s(y) \Big|_{s=0} = \frac{d}{ds} \varphi^s(\varphi^t(x)) \Big|_{s=0} = \frac{d}{ds} \varphi^{s+t}(x) \Big|_{s=0} = \frac{d}{dt} \varphi^{s+t}(x) \Big|_{s=0} = \frac{d}{dt} \varphi^t(x).$$

□

1.15 Bemerkung. (a) Jede der Kurven $I(x) \rightarrow \mathbb{R}^n$, $t \mapsto \varphi^t(x) =: x(t)$ erfüllt also das System von n Gleichungen

$$\begin{aligned} \dot{x}_1(t) &= v_1(x_1(t), \dots, x_n(t)) \\ &\vdots \\ \dot{x}_n(t) &= v_n(x_1(t), \dots, x_n(t)) \end{aligned}$$

oder kurz eben

$$\dot{x} = v(x).$$

Man nennt ein solches System ein **autonomes System gewöhnlicher Differentialgleichungen**. „Differentialgleichung“, weil in den Gleichungen sowohl die Funktionen x_1, \dots, x_n als auch ihre Ableitungen vorkommen; „gewöhnlich“, weil die Funktionen nur von **einer** reellen Veränderlichen t abhängen (nicht von mehreren, in welchem Fall man von „partiellen“

Differentialgleichungen spricht); „autonom“, weil das Vektorfeld v nicht selbst von der Zeit t abhängt.

(b) In den meisten Fällen ist es nun so, dass man das dynamische System φ auf $G \subset \mathbb{R}^n$ gerne kennen würde, aber zunächst nur das zugehörige Vektorfeld $v : G \rightarrow \mathbb{R}^n$ kennt. Es stellt sich also die fundamentale Frage: Inwieweit legt das zugehörige Vektorfeld v das dynamische System φ fest? Wir werden zeigen, dass jedes stetig differenzierbare Vektorfeld v auf $G \subset \mathbb{R}^n$ genau ein (maximales) dynamisches System φ auf G festlegt.

1.16 Definition. Lösung einer Differentialgleichung

Sei $v : G \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Sei $I \subset \mathbb{R}$ ein offenes Intervall. Man nennt eine stetig differenzierbare Kurve $x : I \rightarrow G$, $t \mapsto x(t)$, eine **Lösung der Differentialgleichung**

$$\dot{x} = v(x),$$

wenn für alle $t \in I$ gilt

$$\dot{x}(t) = v(x(t)).$$

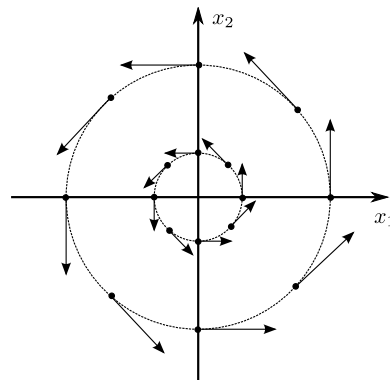
Falls $0 \in I$ ist, so heißt $x(t)$ eine **Lösung zum Anfangswert** $x_0 \in G$, wobei $x_0 = x(0)$ ist.

1.17 Bemerkung. Phasendiagramm

Man verdeutlicht sich das Vektorfeld $v : G \rightarrow \mathbb{R}^n$ oft durch das sogenannte **Phasendiagramm**. Sei beispielsweise $v : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x_1, x_2) \mapsto (-x_2, x_1)$ und

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

Die Lösungskurven $x(t)$ sind überall tangential an v und heißen **Integralkurven** an das Vektorfeld v . Beim Lösen einer Differentialgleichung spricht man auch vom „Integrieren der Gleichung“.



1.18 Beispiele. (a) Der freie Fall

$$F(x) = -mge_3, \quad g > 0, \quad e_3 = (0, 0, 1).$$

Der zugängliche Ortsraum (= Konfigurationsraum) ist $D = \mathbb{R} \times \mathbb{R} \times (0, \infty)$ und der Phasenraum ist $G = D \times \mathbb{R}^3$. Wir suchen Lösungen von

$$\ddot{x} = \frac{1}{m}F(x) = -ge_3,$$

d.h. wir müssen das System $\dot{z} = v(z)$ mit $v(x, y) = (y, -ge_3)$ integrieren:

$$\begin{aligned} \dot{x}_1 &= y_1 & \dot{y}_1 &= 0 \\ \dot{x}_2 &= y_2 & \dot{y}_2 &= 0 \\ \dot{x}_3 &= y_3 & \dot{y}_3 &= -g \end{aligned}$$

Hier bekommt man die Lösungen wirklich unmittelbar durch Integration:

$$\begin{aligned} y_1 &= b_1 & x_1 &= b_1t + a_1 \\ y_2 &= b_2 & x_2 &= b_2t + a_2 \\ y_3 &= -gt + b_3 & x_3 &= -\frac{g}{2}t^2 + b_3t + a_3 \end{aligned}$$

mit Integrationskonstanten $a_1, \dots, b_3 \in \mathbb{R}$ für alle $t \in \mathbb{R}$. Es ist dann $x(0) = a$ und $\dot{x}(0) = b$ und das zu $v : G \rightarrow \mathbb{R}^6, v(x, y) = (y, -ge_3)$ gehörende dynamische System φ auf $\Omega \subset \mathbb{R} \times G$ ist gegeben durch $\varphi : \Omega \rightarrow G$,

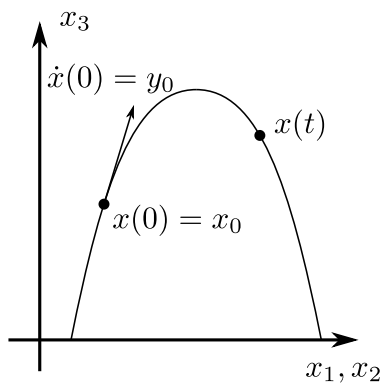
$$\varphi^t(x, y) = \left(-\frac{g}{2}t^2 e_3 + ty + x, -gte_3 + y\right).$$

Die Existenzintervalle sind durch die Bedingung $x_3(t) > 0$ eingeschränkt, es gilt also

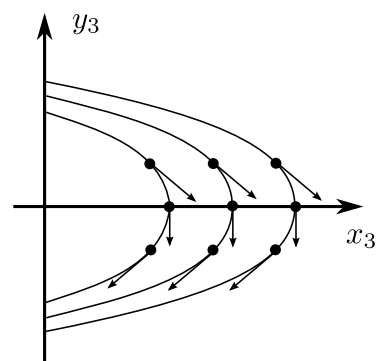
$$I(x, y) = \left(\frac{y_3}{g} - \sqrt{\frac{y_3^2}{g^2} + \frac{2x_3}{g}}, \frac{y_3}{g} + \sqrt{\frac{y_3^2}{g^2} + \frac{2x_3}{g}}\right)$$

und dementsprechend $\Omega = \{(t, x, y) \in \mathbb{R} \times G \mid t \in I(x, y)\}$. Die Lösung von $m\ddot{x} = F(x)$ auf $D \subset \mathbb{R}^3$ zum Anfangswert $(x(0), \dot{x}(0)) = (x_0, y_0)$ ist somit

$$x(t) = -\frac{g}{2}t^2 e_3 + ty_0 + x_0. \quad (\text{Parabelbahn})$$



Lösung im Ortsraum



Phasenraumdiagramm

(b) Das Hookesche Gesetz

Ein Gewicht der Masse m hänge an einer Feder der Federkonstante $k > 0$ und $x \in \mathbb{R}$ bezeichne die vertikale Auslenkung aus der Ruhelage. Dann wirkt auf das Gewicht die Kraft $F(x) = -kx$ und die Bewegungsgleichung lautet diesmal

$$m\ddot{x} = -kx \quad \text{bzw.} \quad \ddot{x} = -\omega^2 x$$

mit $\omega = \sqrt{\frac{k}{m}}$. Mit $G = \mathbb{R} \times \mathbb{R}$ und $v : G \rightarrow \mathbb{R}^2$, $v(x, y) = (y, -\omega^2 x)$ ist also

$$\begin{aligned}\dot{x} &= y \\ \dot{y} &= -\omega^2 x\end{aligned}$$

zu lösen, was wir auch als

$$\dot{z} = Az,$$

schreiben können, mit

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} = \omega \begin{pmatrix} 0 & 1/\omega \\ -\omega & 0 \end{pmatrix} = \omega B.$$

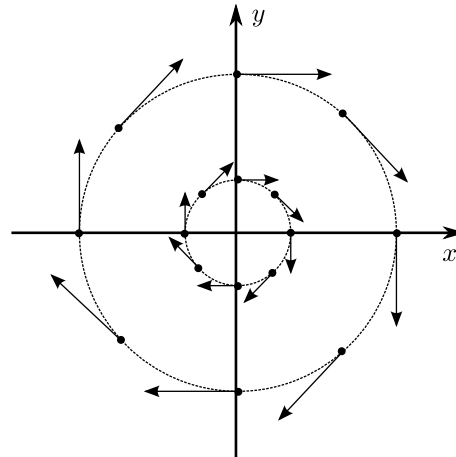
Nun gilt $B^2 = -\mathbb{1}$. Deshalb spielt B die Rolle von i , und man kann leicht sehen, durch Reihenentwicklung der Exponentialfunktion, dass

$$e^{\omega t B} = \mathbb{1} \cos(\omega t) + B \sin(\omega t),$$

gilt.

Die qualitative Form der Lösungen liest man direkt aus dem Phasenraumdiagramm ab, und muss eine Drehung im Phasenraum sein. Explizit erhält man $z(t) = (x(t), y(t)) = e^{\omega t B} z(0) = (x_0, y_0) \cos(\omega t) + B(x_0, y_0) \sin(\omega t)$, sodass wir für $x(t)$ folgendes bekommen,

$$x(t) = x_0 \cos(\omega t) + \frac{y_0}{\omega} \sin \omega t.$$



1.19 Definition. Systeme gewöhnlicher DGLen m -ter Ordnung

Sei $m \in \mathbb{N}$. Ein **System gewöhnlicher Differentialgleichungen m -ter Ordnung** auf einem Gebiet $D \subset \mathbb{R}^n$ ist gegeben durch eine stetige Funktion

$$f : D \times \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{(m-1)\text{-mal}} \rightarrow \mathbb{R}^n$$

und die Gleichung

$$x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)}). \quad (*)$$

Unter einer Lösung von (*) versteht man eine m -mal stetig differenzierbare Kurve $x : I \rightarrow D$, $I \subset \mathbb{R}$ ein offenes Intervall, so dass für alle $t \in I$ gilt

$$x^{(m)}(t) = f(x(t), \dot{x}(t), \dots, x^{(m-1)}(t)).$$

Falls $0 \in I$ ist, heißt $x : I \rightarrow D$ Lösung zum Anfangswert $(x_0, y_1, \dots, y_{m-1})$ wobei $x(0) = x_0$ und $x^{(j)} = y_j$ für $j = 1, \dots, m-1$.

1.20 Bemerkung. Reduktion auf ein System erster Ordnung

Man kann jedes System m -ter Ordnung auf einem Gebiet $D \subset \mathbb{R}^n$,

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

immer auf ein System erster Ordnung auf dem Gebiet $G := D \times \mathbb{R}^{n(m-1)} \subset \mathbb{R}^{nm}$ zurückführen. Man definiert dazu einfach das Vektorfeld $v : G \rightarrow \mathbb{R}^{nm}$ durch

$$v(x, y_1, \dots, y_{m-1}) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ f(x, y_1, \dots, y_{m-1}) \end{pmatrix}$$

und löst dann mit $z = (x, y_1, \dots, y_{m-1}) \in G$ das System

$$\dot{z} = v(z).$$

Beweis. Ist $x : I \rightarrow D$, $t \mapsto x(t)$, Lösung von $x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)})$, so ist $z : I \rightarrow G$

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

Lösung von $\dot{z} = v(z)$. Ist umgekehrt $z(t) = (x(t), y_1(t), \dots, y_{m-1}(t))$ Lösung von $\dot{z} = v(z)$, so ist $t \mapsto x(t)$, $I \rightarrow D$ Lösung von $x^{(m)} = f(x, \dot{x}, \dots, x^{(m-1)})$. \square

1.21 Definition. Autonome und nicht-autonome Systeme

Sei $I \subset \mathbb{R}$ ein offenes Intervall. Ein stetiges **zeitabhängiges Vektorfeld** f auf einem Gebiet $D \subset \mathbb{R}^n$ ist eine stetig differenzierbare Abbildung

$$f : I \times D \rightarrow \mathbb{R}^n, \quad (t, x) \mapsto f(t, x).$$

Man nennt dann das System gewöhnlicher Differentialgleichungen

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

nicht-autonom, wenn f explizit von $t \in I$ abhängt. Ist f (wie bisher) unabhängig von t , so heißt $\dot{x} = f(x)$ **autonom**. Sei $J \subset I$ ein Teilintervall. Eine stetig differenzierbare Kurve $x : J \rightarrow D$ heißt **Lösung von (*)**, wenn für alle $t \in J$ gilt

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

Für $t_0 \in J$ heißt x Lösung zum Anfangswert $x_0 = x(t_0)$ zur Anfangszeit t_0 .

1.22 Bemerkung. Reduktion auf ein autonomes System

Man kann jedes nicht-autonome System $\dot{x} = f(t, x)$ auf einem Gebiet $D \subset \mathbb{R}^n$ auf ein autonomes System auf $G := I \times D \subset \mathbb{R}^{n+1}$ wie folgt zurückführen:

Man setzt $v : G \rightarrow \mathbb{R}^{n+1}$, $v(s, y) = (1, f(s, y))$, $z := (s, y) \in G$, und löse dann $\dot{z} = v(z)$ auf G zum Anfangswert $(s, y) = (t_0, x_0)$. Sei $z(t) = (\underbrace{s(t)}_{=t_0+t}, y(t))$ diese Lösung, dann ist $x(t) = y(t - t_0)$ Lösung von

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

zum Anfangswert x_0 zur Anfangszeit t_0 .

Beweis. Es ist $x(t_0) = y(0) = x_0$ und es gilt

$$\dot{x}(t) = \dot{y}(t - t_0) = f(\underbrace{s(t - t_0)}_{=t}, y(t - t_0)) = f(t, x(t)).$$

\square

1.23 Bemerkung. (a) Im allgemeinen hat man keine Chance, ein System gewöhnlicher Differentialgleichungen

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

auf ein Gebiet $G \subset \mathbb{R}^n$ bei Vorgabe von $v : G \rightarrow \mathbb{R}^n$ zu integrieren, d.h. den zugehörigen Fluss φ explizit zu bestimmen. Man ist daher schon damit zufrieden, qualitative Aussagen über die Bahn $t \mapsto x(t)$ zu bekommen, wie z.B. Antworten auf die Fragen:

- (i) Konvergiert $t \mapsto x(t)$ für $t \rightarrow \infty$ gegen eine Gleichgewichtslage $x_* \in G$, also einen Punkt mit $v(x_*) = 0$?
- (ii) Ist $t \mapsto x(t)$ periodisch, also $x(t+T) = x(t)$ für ein $T > 0$ und alle $t \in \mathbb{R}$?
- (iii) Oder auch nur: für welche $x \in G$ ist überhaupt $I(x) = \mathbb{R}$, existiert also eine Lösung für alle Zeiten?

Man nennt das Studium dieser Fragen die „Qualitative Theorie gewöhnlicher Differentialgleichungen“.

(b) Hat man spezielle Informationen über v oder G , z.B. Symmetrieaussagen oder $G \subset \mathbb{R}$ so kann man $\dot{x} = f(x)$ in manchen Fällen doch explizit integrieren.

1.24 Proposition. Integration eindimensionaler DGLen

Sei $G \subset \mathbb{R}$ ein Intervall, $v : G \rightarrow \mathbb{R}$ stetig und $x_0 \in G$ mit $v(x_0) \neq 0$. Seien weiter $\delta_1, \delta_2 > 0$ derart, dass $(x_0 - \delta_1, x_0 + \delta_2) \subset G$ und $v(x) \neq 0$ für $x \in (x_0 - \delta_1, x_0 + \delta_2)$ ist. Definiere $\tau : (x_0 - \delta_1, x_0 + \delta_2) \rightarrow \mathbb{R}$ durch

$$\tau(x) = \int_{x_0}^x \frac{du}{v(u)}.$$

Dann ist $I := \text{Bild}(\tau) \subset \mathbb{R}$ ein offenes Intervall und $\tau : (x_0 - \delta_1, x_0 + \delta_2) \rightarrow I$ ist ein Diffeomorphismus. Seine Inverse $\varphi := \tau^{-1} : I \rightarrow (x_0 - \delta_1, x_0 + \delta_2)$ ist Lösungskurve von $\dot{x} = v(x)$ auf G zum Anfangswert x_0 .

Beweis. von Proposition 1.24. Es ist $\tau'(x) = \frac{1}{v(x)} \neq 0$ und nach dem Zwischenwertsatz für alle $x \in (x_0 - \delta_1, x_0 + \delta_2)$ von gleichem Vorzeichen. Damit ist τ streng monoton und deshalb bijektiv auf sein Bild. Nach dem Umkehrsatz ist τ ein Diffeomorphismus auf sein offenes Bild I . Für $\varphi = \tau^{-1}$ gilt dann $\dot{\varphi}(t) = \frac{1}{\tau'(\varphi(t))} = \frac{1}{v(\varphi(t))}$ für alle $t \in I$. Wegen $\tau(x_0) = 0$ ist $\varphi(0) = x_0$. \square

1.25 Bemerkung. (a) Man kann also eindimensionale Gleichungen durch „Quadratur“ lösen, d.h. durch die Prozesse „Stammfunktion bilden“ und „Invertieren“.

(b) Das Verfahren der Trennung der Variable funktioniert auch für den Fall, dass v zeitabhängig und von der Form

$$v : I \times G \rightarrow \mathbb{R}, \quad v(t, x) = a(t)b(x)$$

ist, wobei $I \subset \mathbb{R}$ ein offenes Intervall und a und b jeweils stetig sind. Dann liefert Auflösen von

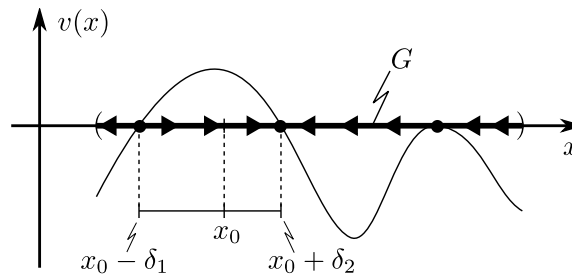
$$\int_{t_0}^{\tau(x)} a(t) dt = \int_{x_0}^x \frac{du}{b(u)}$$

nach $\tau(x)$ analog zu Proposition 1.24 eine lokale Lösung von

$$\dot{x}(t) = v(t, x(t)) \quad \text{zum Anfangswert } x(t_0) = x_0.$$

- (c) Ganz allgemein sind die Punkte $x_0 \in G$, wo ein gegebenes Vektorfeld $v : G \rightarrow \mathbb{R}^n$ verschwindet, $v(x_0) = 0$, für das dynamische System von $\dot{x} = v(x)$ besonders einfach: Es ist $I(x_0) = \mathbb{R}$ und $x_0(t) = \varphi^t(x_0) = x_0$ für alle $t \in \mathbb{R}$. Man nennt diese Punkte Gleichgewichtslagen oder stationäre Punkte.
- (d) Mit Proposition 1.24 ist das dynamische System φ von $\dot{x} = v(x)$ auf $G \subset \mathbb{R}$ vollständig bestimmt. Sein Phasendiagramm sieht so aus:

In den Nullstellen von v ruht das System. Dazwischen läuft es monoton von einer Nullstelle zur nächsten. Ist v an den Nullstellen differenzierbar, so werden die Gleichgewichtslagen in endlicher Zeit nicht erreicht. An den Rändern des Intervalls G können die Orbits in endlicher Zeit entweichen.



1.26 Beispiele. (a) Sei $a \in \mathbb{R}$ und $v : \mathbb{R} \rightarrow \mathbb{R}, v(x) = ax$ ein lineares Vektorfeld. Die Lösung von $\dot{x} = v(x)$ auf $G = \mathbb{R}$ zum Anfangswert $x_0 > 0$ für $a \neq 0$ ist nach Proposition 1.23 für $x > 0$ gegeben durch

$$t(x) = \int_{x_0}^x \frac{du}{au} = \frac{1}{a} (\ln x - \ln x_0) = \frac{1}{a} \ln \frac{x}{x_0}$$

also

$$\ln \frac{x}{x_0} = at \Rightarrow x(t) = x_0 e^{at} \quad \forall t \in \mathbb{R}.$$

Diese Formel gilt auch für $x_0 \leq 0$ und somit ist der Fluss $\varphi^t : \mathbb{R} \rightarrow \mathbb{R}, \varphi^t(x_0) = e^{at} x_0$ global.

(b) Sei $v : \mathbb{R} \rightarrow \mathbb{R}, v(x) = 1 + x^2$ quadratisch in x . Die Lösungskurve $t \rightarrow x(t)$ zum Anfangswert $x_0 = 0$ ist

$$t(x) = \int_0^x \frac{du}{1 + u^2} = \arctan(x)$$

also

$$x(t) = \tan(t).$$

Somit ist $I(0) = (-\frac{\pi}{2}, \frac{\pi}{2})$ und $t \mapsto x(t)$ läuft „in endlicher Zeit nach Unendlich“.

(c) Sei $v : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $v(x) = \sqrt{|x|}$. Dann erreichen die Lösungen zu Startwerten $x(0) < 0$ in endlicher Zeit die Ruhelage $x = 0$. Dort können Sie beliebig lange verweilen und dann nach rechts wieder aus der Ruhelage herauslaufen. Die Lösungen sind also nicht eindeutig. (vgl. Übungen).

1.27 Definition. Lipschitz-Stetigkeit

Sei $G \subset \mathbb{R}^n$ ein Gebiet und $v : G \rightarrow \mathbb{R}^n$ ein Vektorfeld.

(a) Es heißt v **Lipschitz-stetig**, wenn es ein $L > 0$ gibt, so dass für alle $x, y \in G$ gilt

$$\|v(x) - v(y)\| \leq L\|x - y\|.$$

Es heißt dann L eine **Lipschitz-Konstante** für v .

(b) Es heißt v **lokal-Lipschitz-stetig**, wenn jedes $x \in G$ eine offene Umgebung $U \subset G$ besitzt, so dass $v|_U$ Lipschitz-stetig ist.

1.28 Bemerkung. Differenzierbar \Rightarrow lokal Lipschitz

Sei $v : G \rightarrow \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einem Gebiet $G \subset \mathbb{R}^n$. Dann ist v lokal Lipschitz-stetig.

Beweis. Übungsaufgabe (leicht). □

1.29 Proposition. Sei $G \subset \mathbb{R}^n$ ein Gebiet und $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld. Für jedes Kompaktum $K \subset G$ ist $v|_K$ Lipschitz-stetig.

Beweis. Dies folgt aus folgender Tatsache: Es sei $\cup_x U_x$ die Vereinigung dieser Umgebungen U_x von x , wo $v(x)$ Lipschitzstetig ist. Es gilt natürlich $K \subset \cup_x U_x$. Nun verwenden wir die Tatsache, dass aus jeder offenen Überdeckung einer kompakten Menge K man eine endliche Teilüberdeckung $\cup_{x_i} U_{x_i}$ für $i = 1, \dots, M$, auswählen kann. Für separable Räume (es gibt eine abzählbare dichte Teilmenge) ist diese Eigenschaft

äquivalent zu dem Begriff der Kompaktheit, die wir kennen, nämlich dem dass jede Folge eine konvergente Teilfolge besitzt.

Auf Grund der Separabilität kann man nämlich zuerst schliessen, dass es eine abzählbare Überdeckung $\cup_{x_j} U_{x_j}$ geben muss. Angenommen man kann keine endliche Teilüberdeckung auswählen, dann findet man Folge $x_N \in K \setminus \cup_{x_j}^N U_{x_j}$. Aber da K kompakt ist gibt es konvergente Teilfolge x_{N_j} mit Limes $x \in K$. Dies x liegt in mindestens einem U_{x_l} , wo aber auch alle bis auf endlich viele Elemente von x_{N_j} liegen, damit einen Widerspruch zu obiger Annahme.

Mit dieser endlichen Überdeckung lässt sich jetzt leicht die Aussage zeigen. □

1.30 Bemerkung. Sei $D \subset \mathbb{R}^n$ kompakt. Dann wissen wir aus Mathe 1 dass $X = (C^0(D), \|\cdot\|_\infty)$ ein Banachraum ist, wobei $C^0(D)$ die Menge der stetigen Funktionen mit Bild in \mathbb{R}^n sind, und

$$\|f\|_\infty = \sup_{x \in D} \|f(x)\|,$$

und $\|\cdot\|$ die Euklidische Norm eines Vektors in \mathbb{R}^n . Es sei erinnert, dass Banachraum bedeutet, dass der Raum vollständig ist bezüglich der Norm $\|\cdot\|_\infty$. Das heisst, dass alle Cauchyfolgen konvergieren.

Beweis. Sei f_n eine Cauchyfolge, also $\|f_n - f_m\| \rightarrow_{n,m \rightarrow \infty} 0$. Da für alle $x \in D$ $f_n(x)$ beschränkt ist, gibt es konvergente Teilfolge, die gegen einen Grenzwert, sagen wir $f(x)$ konvergiert. Also für jedes $\varepsilon > 0$ und jedes x gibt es ein $n(x)$ sodass $\|f(x) - f_{n(x)}(x)\| < \varepsilon/2$. Da f_n Cauchyfolge ist können wir m und auch $n(x)$ so gross genug wählen, sodass $\|f_{n(x)} - f_m\|_\infty < \varepsilon/2$ ist. Also gilt

$$\|f(x) - f_m(x)\| \leq \|f(x) - f_{n(x)}(x)\| + \|f_{n(x)}(x) - f_m(x)\| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

und zwar unabhängig von x . Das heisst $\|f - f_m\|_\infty \rightarrow 0$. Und da wir wissen, dass der gleichmäßige Limes stetiger Funktionen stetig ist, gilt dass f stetig ist und somit ist X ein Banachraum. Die Stetigkeit von f sieht man durch

$$\|f(x) - f(y)\| \leq \|f(x) - f_n(x)\| + \|f_n(x) - f_n(y)\| + \|f_n(y) - f(y)\|$$

und der Tatsache, dass der erste und dritte Term wegen der Gleichmäßigkeit kleiner als jedes ε gewählt werden kann für n genügend gross und der mittlere Term auf Grund der gleichmäßigen Stetigkeit von $f_n(x)$ kleiner als ε wird für alle y in einer geeigneten δ -Umgebung. □

1.31 Theorem. Existenz- und Eindeutigkeitssatz von Picard-Lindelöf

Sei $G \subset \mathbb{R}^n$ ein Gebiet, $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld und $x_0 \in G$.

Dann existiert ein $\delta > 0$, sodass es genau eine stetig differenzierbare Kurve $x : (-\delta, \delta) \rightarrow G$ gibt, die Lösung von $\dot{x} = v(x)$ zum Anfangswert x_0 ist, also

$$\left. \begin{aligned} \dot{x}(t) &= v(x(t)), & \forall t \in (-\delta, \delta) \\ x(0) &= x_0 \end{aligned} \right\} (*)$$

erfüllt.

1.32 Bemerkung. Die Grundidee des Beweises ist die Folgende: Ist $x : (-\delta, \delta) \rightarrow G$ Lösung von (*), so folgt durch Integration und den Hauptsatz

$$x(t) - \underbrace{x(0)}_{= x_0} = \int_0^t \dot{x}(s) ds = \int_0^t v(x(s)) ds,$$

also

$$x(t) = x_0 + \int_0^t v(x(s)) ds. \quad (**)$$

Jede Lösung von (*) ist also Fixpunkt der Abbildung Φ , die jeder Funktion φ die Funktion $\Phi[\varphi]$ zuordnet, gegeben durch

$$\Phi[\varphi](t) = x_0 + \int_0^t v(\varphi(s)) ds.$$

Ist umgekehrt x Fixpunkt von Φ und stetig, so liefert Differentiation von (**), dass $\dot{x}(t) = v(x(t))$ für alle $t \in (-\delta, \delta)$, also x Lösung von $\dot{x} = v(x)$ zum Anfangswert $x(0) = x_0$ ist.

Man muss nun lediglich den Funktionenraum auf dem Φ operiert so definieren, dass er ein Banachraum stetiger Funktionen ist und Φ eine Kontraktion darauf. Dann liefert der Banachsche Fixpunktsatz das Resultat.

Beweis. von Theorem 1.31. Weil G offen ist, existiert ein $r > 0$, so dass $K = \overline{B}_r(x_0) \subset G$ ist.

Da K kompakt und $v|_K : K \rightarrow \mathbb{R}^n$ stetig ist, existiert ein $M > 0$ so, dass für alle $x \in K$ gilt $\|v(x)\| \leq M$.

Nach Proposition 1.29 existiert auch ein $L > 0$, sodass $v|_K$ Lipschitz-stetig ist mit Lipschitz-Konstante L .

Setze nun $\delta := \min\{\frac{1}{L}, \frac{r}{M}\}$ und sei $0 < \delta_0 < \delta$ beliebig. Wir betrachten den Vektorraum

$$X = \{\varphi : [-\delta_0, \delta_0] \rightarrow \mathbb{R}^n \mid \varphi \text{ ist stetig}\} = C([-\delta_0, \delta_0], \mathbb{R}^n)$$

mit der Supremumsnorm

$$\|\varphi\|_\infty = \sup_{t \in [-\delta_0, \delta_0]} \|\varphi(t)\|.$$

Es ist dann $(X, \|\cdot\|_\infty)$ ein Banachraum und der Teilraum

$$A = \{\varphi \in X \mid \varphi(0) = x_0 \text{ und } \varphi([-\delta_0, \delta_0]) \subset K\} \subset X$$

ist abgeschlossen. Für $\varphi \in A$ sei nun $\Phi[\varphi] : [-\delta_0, \delta_0] \rightarrow \mathbb{R}^n$ gegeben durch

$$\Phi[\varphi](t) := x_0 + \int_0^t v(\varphi(s)) ds.$$

Es ist dann $\Phi[\varphi]$ wieder in A , denn $\Phi[\varphi]$ ist stetig (sogar stetig differenzierbar), $\Phi[\varphi](0) = x_0$ und $\Phi[\varphi](t) \in K$ für alle $t \in [-\delta_0, \delta_0]$. Letzteres folgt aus

$$\|\Phi[\varphi](t) - x_0\| \leq \left\| \int_0^t v(\varphi(s)) ds \right\| \leq \int_0^t \underbrace{\|v(\underbrace{\varphi(s)}_{\in K})\|}_{\leq M} ds \leq \delta_0 M < \frac{r}{M} M = r.$$

Damit ist Φ eine Abbildung von A nach A , $\Phi : A \rightarrow A$. Schließlich ist Φ eine Kontraktion mit Kontraktionskonstante $0 < \theta < 1$, $\theta := \delta_0 L$. Denn für alle $\varphi, \psi \in A$ gilt

$$\begin{aligned} \|\Phi[\varphi] - \Phi[\psi]\|_\infty &= \sup_{|t| \leq \delta_0} \left\| \int_0^t (v(\varphi(s)) - v(\psi(s))) ds \right\| \leq \sup_{|t| \leq \delta_0} \int_0^t \|v(\varphi(s)) - v(\psi(s))\| ds \\ &\leq \sup_{|t| \leq \delta_0} \int_0^{\delta_0} L \|\varphi(s) - \psi(s)\| ds \leq L \delta_0 \|\varphi - \psi\|_\infty = \theta \|\varphi - \psi\|_\infty \end{aligned}$$

mit $\theta = \delta_0 L < 1$, weil $\delta_0 < \frac{1}{L}$ ist.

Nach dem Banachschen Fixpunktsatz existiert also genau ein $\varphi \in A$ mit $\Phi[\varphi] = \varphi$.

Setze nun $x : (-\delta, \delta) \rightarrow G$, $x(t) = \varphi(t)$, wobei für $t \in (-\delta, \delta)$ erst ein $|t| < \delta_0 < \delta$ gewählt sei und φ dann der eindeutige Fixpunkt von $\Phi = \Phi_{\delta_0}$ sei. Da jeder Fixpunkt von Φ_{δ_0} auch Fixpunkt von Φ_{δ_1} mit $\delta_1 \leq \delta_0$ ist, hängt diese Definition nicht von der Wahl von δ_0 ab. Es folgt, dass x stetig ist und $\forall t \in (-\delta, \delta)$ gilt:

$$x(t) = x_0 + \int_0^t v(x(s)) ds.$$

Also ist x sogar stetig differenzierbar mit

$$\dot{x}(t) = v(x(t)) \quad \text{und} \quad x(0) = x_0.$$

Also ist x Lösung von (*). Da jede Lösung von (*) auch (**) erfüllt und die Lösung von (**) ja nach dem Banachschen Fixpunktsatz eindeutig ist, ist x auch die eindeutige Lösung von (*). \square

1.33 Bemerkung. (a) Man beachte, dass der Existenzsatz nur die Existenz einer Lösung von $\dot{x} = v(x)$, $x(0) = x_0$ „für kurze Zeiten“ liefert, $x : (-\delta, \delta) \rightarrow G$. Man spricht daher von lokaler- bzw. Kurzzeiteistenz.

(b) Der Beweis gibt auch eine untere Schranke für die Lebensdauer der Lösung x , nämlich

$$\delta(x_0) = \min \left\{ \frac{1}{L}, \frac{r(x_0)}{M} \right\},$$

wobei $r(x_0)$ so ist, dass $K = \overline{B_r(x_0)} \subset G$ ist und $M > 0$ eine Schranke für die „Geschwindigkeit“ $\|v\|$ auf K und $L > 0$ eine Lipschitzkonstante (z.B. eine Schranke auf $\|Dv\|$, falls v stetig differenzierbar ist) auf K ist.

1.34 Definition. Maximale Lösung

Sei $G \subset \mathbb{R}^n$ ein Gebiet und $v : G \rightarrow \mathbb{R}^n$ ein stetiges Vektorfeld. Eine Lösung $x : I \rightarrow G$, $I \subset \mathbb{R}$ ein offenes Intervall, von $\dot{x} = v(x)$ auf G heißt **maximal**, wenn gilt: Ist $\tilde{I} \subset \mathbb{R}$ ein Intervall mit $\tilde{I} \supset I$ und $\tilde{x} : \tilde{I} \rightarrow G$ eine Lösung von $\dot{x} = v(x)$ mit $\tilde{x}|_I = x$, so gilt bereits $\tilde{I} = I$ und somit $\tilde{x} = x$.

Man sagt auch, eine maximale Lösung kann nicht fortgesetzt werden.

1.35 Proposition. Sei $G \subset \mathbb{R}^n$ ein Gebiet und $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld. Seien $I, J \in \mathbb{R}$ offene Intervalle mit $0 \in I \cap J$ und $x : I \rightarrow G$ und $y : J \rightarrow G$ Lösungen der Differentialgleichung $\dot{x} = v(x)$ mit Anfangswert $x(0) = y(0) = x_0 \in G$. Dann gilt für alle $t \in I \cap J$: $x(t) = y(t)$.

Beweis. Sei $A = \{t \in I \cap J \mid x(t) = y(t)\}$. Wir zeigen, dass $A = I \cap J$ ist. Sei dazu $I = (t_-, t_+)$ und $J = (s_-, s_+)$ und o.B.d.A. $t_+ \leq s_+$ und

$$t_0 := \sup\{\tau \in (0, t_+) \mid x(t) = y(t) \forall 0 \leq t < \tau\}.$$

Nach dem Eindeutigkeitsatz existiert ein $\delta_{x_0} > 0$, sodass $x(t) = y(t)$ für alle $0 \leq t < \delta_{x_0}$. Also ist $0 < \delta \leq t_0 \leq t_+$. Angenommen $t_0 < t_+$. Die Stetigkeit von x liefert dann $x(t_0) = y(t_0) =: \tilde{x}_0$. Wiederum nach Picard-Lindelöf existiert dann aber ein $\delta_{\tilde{x}_0} > 0$ so, dass $x(t) = y(t)$ für alle $t \in (t_0 - \delta_{\tilde{x}_0}, t_0 + \delta_{\tilde{x}_0})$, was im Widerspruch zur Definition vor t_0 steht. Also ist $t_0 = t_+$ und mit einem analogen Argument für die andere Intervallgrenze ergibt sich $A = I \cap J$. □

1.36 Satz. Existenz und Eindeutigkeit der maximalen Lösung

Sei $G \subset \mathbb{R}^n$ ein Gebiet und $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld. Dann existiert zu jedem $x_0 \in G$ genau eine maximale Lösung $x : I \rightarrow G$ von $\dot{x} = v(x)$ mit $x(0) = x_0$.

Beweis. Ist $y : J \rightarrow G$ eine Lösung von $\dot{x} = v(x)$ mit $y(0) = x_0$ und $0 \in J$, so notieren wir diese mit (y, J) . Sei nun

$$I := \bigcup_{\substack{J \subset \mathbb{R} \\ (y, J) \text{ ist Lösung}}} J$$

Dann ist $I \subset \mathbb{R}$ ein offenes Intervall. Ist $t \in I$ beliebig, so wählen wir eine Lösung (y, J) mit $t \in J$ und setzen $x(t) := y(t)$. Wegen Proposition 1.35 ist dies unabhängig von der Wahl (y, J) . Es ist dann die so definierte Funktion $x : I \rightarrow G$ eine Lösung von $\dot{x} = v(x)$ mit $x(0) = x_0$. Nach Konstruktion ist x maximal und zwar die einzige maximale Lösung. □

1.37 Bemerkung. (a) Wir schreiben für das Definitionsintervall I der maximalen Lösung von $\dot{x} = v(x)$, $x(0) = x_0$,

$$I(x_0) = (t_-(x_0), t_+(x_0)),$$

wobei $t_-(x_0) \in [-\infty, 0]$ bzw. $t_+(x_0) \in [0, \infty]$ die linke bzw. rechte Intervallgrenze bezeichnet.

(b) Wir setzen nun weiter

$$\Omega := \{(t, x) \in \mathbb{R} \times G \mid t \in I(x)\}$$

und $\varphi : \Omega \rightarrow G$, $(t, x) \mapsto \varphi^t(x) = x(t)$, wobei $I(x) \rightarrow G$, $t \mapsto x(t)$ die maximale Lösung von $\dot{x} = v(x)$ zum Anfangswert $x \in G$ ist.

- (c) Man kann nun zeigen, dass Ω offen ist und $\varphi : \Omega \rightarrow G$ stetig ist. Man spricht von „stetiger Abhängigkeit von den Anfangsdaten“. Falls $v : G \rightarrow \mathbb{R}^n$ stetig differenzierbar ist, so ist auch $\varphi : \Omega \rightarrow G$ stetig differenzierbar.
- (d) Wir zeigen hier noch, dass φ einen Fluss definiert, also die Verträglichkeitsbedingung (b) aus Definition 1.12 erfüllt: Offenbar ist $\varphi^0(x) = x$ für alle $x \in G$. Für $x \in G$ und $t \in I(x)$ lösen nun sowohl

$$\phi : (t_-(x) - t, t_+(x) - t) \rightarrow G, \quad \phi(s) = \varphi^{s+t}(x)$$

als auch

$$\psi : (t_-(\varphi^t(x)), t_+(\varphi^t(x))) \rightarrow G, \quad \psi(s) = \varphi^s(\varphi^t(x))$$

die Differentialgleichung $\dot{x} = v(x)$ zum Anfangswert $\varphi^t(x)$. Außerdem sind beide Lösungen maximal und daher gilt: $s \in I(\varphi^t(x)) \Leftrightarrow s + t \in I(x)$ und für diese s gilt $\phi(s) = \psi(s)$ also $\varphi^s(\varphi^t(x)) = \varphi^{s+t}(x)$.

- (e) Ähnlich wie zuvor sagen wir, dass ein dynamisches System $\varphi : \Omega \rightarrow G$ **maximal** ist, wenn gilt: Ist $\tilde{\varphi} : \tilde{\Omega} \rightarrow G$ ein dynamisches System mit $\tilde{\Omega} \supset \Omega$ und $\tilde{\varphi}|_{\Omega} = \varphi$, so muss bereits $\tilde{\Omega} = \Omega$ sein und damit $\tilde{\varphi} = \varphi$.
- (f) Das dynamische System $\varphi : \Omega \rightarrow G$, das wir in (b) und (c) zu einem Vektorfeld $v : G \rightarrow \mathbb{R}^n$ konstruiert haben, ist offenbar maximal. Sein assoziiertes Vektorfeld ist offenbar gerade v .

1.38 Bemerkung. Insgesamt ergibt sich also, dass jedem maximalen dynamischen System φ auf G ein stetiges Vektorfeld $v = \frac{d}{dt}\varphi^t|_{t=0}$ zugeordnet werden kann. Umgekehrt liefert zumindest jedes stetig differenzierbare Vektorfeld v auf G ein maximales dynamisches System φ auf G .

Dabei ist der Übergang $\varphi \mapsto v$ einfach („Differenzieren kann jeder“) der Übergang $v \mapsto \varphi$ schwer („Integrieren ist eine Kunst“).

Im allgemeinen weiß man nicht einmal, ob bei gegebenem $v : G \rightarrow \mathbb{R}^n$ und $x \in G$ das Ende $t_+(x) \in (0, \infty]$ endlich oder unendlich ist. Der folgende Satz besagt aber immerhin, dass die Bahnkurve $t \mapsto x(t)$ eines Punktes $x \in G$ mit $t_+(x) < \infty$ jedes Kompaktum K in G verlassen muss, wenn $t \rightarrow t_+(x)$ geht, d.h. $t \mapsto x(t)$ strebt für $t \rightarrow t_+(x)$ zum Rand von G oder nach Unendlich. Anders gesagt, $t \mapsto x(t)$ kann sich in endlicher Zeit „nicht einfach in Luft auflösen“.

1.39 Satz. Verhalten für $t \rightarrow t_+(x)$

Sei $G \subset \mathbb{R}^n$ ein Gebiet, $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld und $x \in G$ so, dass $t_+(x) < \infty$. Ist dann $K \subset G$ kompakt, so gibt es ein $0 < \tau < t_+(x)$, sodass für alle $\tau < t < t_+(x)$ gilt: $x(t) \notin K$.

Beweis. Da K kompakt ist, gibt es ein $\varrho > 0$, so dass $B_{\varrho}(x) \subset G$ für alle $x \in K$ gilt. Denn die Distanzfunktion zum Rand von G , $\text{dist} : \mathbb{R}^n \rightarrow [0, \infty)$,

$$\text{dist}(x, \partial G) := \inf \{ \|x - y\| \mid y \in \partial G \}$$

ist stetig und nimmt auf K ihr Minimum ϱ an. Es ist $\varrho > 0$ denn sonst wäre $K \cap \partial G \neq \emptyset$, also $K \not\subset G$. Seien weiter $\|v(x)\| \leq M$ für $x \in \overline{B}_{\varrho}(K)$ und L eine Lipschitzkonstante für v auf $\overline{B}_{\varrho}(K)$. Dann gilt für $t \in (0, t_+(x))$ mit $x(t) \in K$, dass

$$t_+(x) = t_+(x(t)) + t \geq \delta + t$$

mit $\delta := \min \{ \frac{1}{L}, \frac{\varrho}{M} \}$. Für alle $t \in (\tau, t_+(x))$ mit $\tau := t_+(x) - \delta$ muss also gelten, dass $x(t) \notin K$. □

- 1.40 Bemerkung.** (a) Eine entsprechende Aussage gilt natürlich, wenn $t_-(x) > -\infty$ ist.
 (b) Bleibt eine Lösungskurve $t \mapsto x(t)$ in einem Kompaktum, z.B. wenn $\lim_{t \rightarrow t_+} x(t) = p$ für ein $p \in G$ ist, so muss also $t_+(x) = \infty$ sein. In diesem Fall muss dann p eine Gleichgewichtslage sein.

Falls wir wissen dass $v(x)$ höchstens linear wächst, dann kann man garantieren dass $t_+(x) = \infty$. Sei auf $G = \mathbb{R}^n$ $\|v(x)\| \leq L\|x\|$, dann wissen wir für

$$x(t) = x_0 + \int_0^t v(x(s))ds,$$

dass

$$\|x(t)\| \leq \|x_0\| + L \int_0^t \|x(s)\| ds.$$

Mit dem folgenden Lemma wissen wir, dass eine solche Lösung höchstens exponentiell anwachsen kann und somit NICHT in endlicher Zeit divergieren kann. Wir wissen dann dass für alle x gilt $t_+(x) = \infty$.

1.41 Lemma. von Gronwall

Sei $a < b$ und $u : [a, b] \rightarrow [0, \infty)$ eine stetige Funktion. Es gebe Konstanten $L, C \geq 0$ so, dass für alle $t \in [a, b]$ gilt

$$u(t) \leq C + L \int_a^t u(s) ds.$$

Dann ist

$$u(t) \leq C e^{L(t-a)} \quad \text{für alle } t \in [a, b].$$

Beweis. Sei zunächst $C > 0$. Setze $U : [a, b] \rightarrow [0, \infty)$

$$U(t) := C + L \int_a^t u(s) ds.$$

Dann ist U stetig differenzierbar und $\dot{U}(t) = L u(t) \geq 0$, also monoton steigend. Außerdem ist $U(a) = C > 0$, also $U(t) \geq C$ für alle $t \in [a, b]$. Es folgt

$$\frac{d}{dt} \ln(U(t)) = \frac{\dot{U}(t)}{U(t)} = \frac{L u(t)}{U(t)} \leq L$$

da $u \leq U$. Also ist

$$\ln U(t) - \ln C = \ln U(t) - \ln U(a) = \int_a^t \frac{d}{ds} \ln U(s) ds \leq \int_a^t L ds = L(t-a).$$

und somit

$$u(t) \leq U(t) = \exp(\ln U(t)) \leq \exp(\ln C + L(t-a)) = C e^{L(t-a)}.$$

Falls $C = 0$, so liefert unser Argument, dass $u(t) \leq \tilde{C} e^{Lt}$ für jedes $\tilde{C} > 0$, also $u(t) \equiv 0$. □

Damit haben wir folgendes bewiesen.

1.42 Satz. Sei $G = \mathbb{R}^n$, $v : G \rightarrow \mathbb{R}^n$ ein lokal Lipschitz-stetiges Vektorfeld und $\|v(x)\| \leq L\|x\|$ für alle $x \in G$. Dann gilt für alle $x \in G$, dass $I(x) = (-\infty, \infty)$ und

$$\|x(t)\| \leq \|x_0\| e^{tL},$$

falls $x(0) = x_0$ ist.

1.43 Beispiel. Das N -Körper Problem

Betrachte N Punktteilchen im \mathbb{R}^3 die über das Gravitationspotential

$$V(q_i - q_j) = -\frac{1}{|q_i - q_j|}$$

wechselwirken (Masse und Gravitationskonstante = 1). Der Phasenraum des Systems ist der \mathbb{R}^{6N} und wir schreiben

$$z = (q, p) \in \mathbb{R}^{6N}.$$

Die zugehörige Differentialgleichung ist

$$\dot{z} = v(z)$$

mit

$$v(q, p) = \left(p, \nabla \sum_{i < j} \frac{1}{|q_i - q_j|} \right) = \left(p, - \sum_{i < j} \frac{q_i - q_j}{|q_i - q_j|^3} \right).$$

Es ist v lokal Lipschitz-stetig auf dem Gebiet

$$G = \{(q, p) \in \mathbb{R}^{6N} \mid q_i \neq q_j \ \forall i \neq j\}.$$

Für $N = 2$ kann man die Lösungen explizit berechnen (Keplerproblem).

Für $N = 3$ existieren die Lösungen für „fast alle“ Anfangsdaten global, man kennt sie aber nicht. Für $N > 3$ würde man gerne zeigen, dass die Lösungen für „fast alle“ Anfangsdaten existieren.

Problem: Für alle $N \geq 2$ gibt es Lösungen, die nicht global existieren, nämlich die direkte Kollision zweier Teilchen. Für $N \geq 5$ ist bekannt, dass Lösungen in endlicher Zeit nach unendlich laufen können.

Eine naheliegende mathematische Frage ist also: Wieviel „schlechte“ Anfangsdaten gibt es ?

Das quantenmechanische N -Körper Problem ist übrigens kein Problem mehr. Die Lösungen der entsprechenden Schrödingergleichung existieren für alle Zeiten, die zugehörigen Teilchentrajektorien in der Bohmschen Mechanik existieren für fast alle Anfangsdaten global.

2 Fourier-Analyse

2.0.1 Einleitung

Die Wellengleichung ist gegeben durch

$$\partial_t^2 u = \partial_x^2 u.$$

Wir untersuchen nun zwei Lösungsansätze.

Den ersten Lösungsansatz erhalten wir durch „Faktorisieren“ der Wellengleichung, schreibe dazu

$$(\partial_t - \partial_x)(\partial_t + \partial_x)u = 0, \quad (\partial_t + \partial_x)(\partial_t - \partial_x)u = 0.$$

Man liest daran ab, dass Lösungen die Form $u(x, t) = F(x+t) + G(x-t)$ haben, sich also in einen rechts- und einen linkslaufenden Teil aufspalten lassen. Gib die Randbedingungen vor: $u(t, 0) = 0$ und $u(t, \pi) = 0 \forall t \geq 0$. Gibt man jetzt noch eine Anfangsauslenkung sowie eine Anfangsgeschwindigkeit vor, so erhält man folgendes Gleichungssystem:

$$\begin{aligned} f(x) &= u(0, x) = F(x) + G(x) \\ g(x) &= \partial_t u(0, x) = F'(x) - G'(x). \end{aligned}$$

Löst man dieses nach F' und G' auf, so erhält man zunächst

$$\begin{aligned} 2F'(x) &= f'(x) + g(x) \quad \Rightarrow F(x) = \frac{1}{2} \left(f(x) + \int_0^x g(u) du \right) + c_1 \\ 2G'(x) &= f'(x) - g(x) \quad \Rightarrow G(x) = \frac{1}{2} \left(f(x) - \int_0^x g(u) du \right) + c_2. \end{aligned}$$

Addiere nun die beiden erhaltenen Gleichungen:

$$F(x) + G(x) = f(x) + c_1 + c_2 \quad \Rightarrow c_1 + c_2 = 0.$$

Setze nun t ein und erhalte eine Lösung für u :

$$u(t, x) = F(x+t) + G(x-t) = \frac{1}{2} (f(x+t) + f(x-t)) + \frac{1}{2} \int_{x-t}^{x+t} g(u) du.$$

Diese allgemeine Lösung heißt d'Alembert-Lösung. Setze nun die oben vorgegebenen Randbedingungen ein:

$$\begin{aligned} u(t, 0) &= \frac{1}{2} (f(t) + f(-t)) + \frac{1}{2} \int_{-t}^t g(u) du \stackrel{!}{=} 0 \\ u(t, \pi) &= \frac{1}{2} (f(\pi+t) + f(\pi-t)) + \frac{1}{2} \int_{\pi-t}^{\pi+t} g(u) du \stackrel{!}{=} 0. \end{aligned}$$

Also müssen f und g ungerade Funktionen und 2π -periodisch sein.

Als zweiten Lösungsansatz haben wir einen Separationsansatz: $u(x, t) = \varphi(x)\psi(t)$. Setzt man dies in die Wellengleichung ein, so erhält man

$$\begin{aligned} \psi''(t)\varphi(x) &= \psi(t)\varphi''(x) \\ \frac{\psi''(t)}{\psi(t)} &= \frac{\varphi''(x)}{\varphi(x)} \quad \forall (x, t). \end{aligned}$$

Da dies für alle Orte und alle Zeiten gelten muss, können beide Seiten der Gleichung nur konstant sein. Man erhält damit

$$\varphi''(x) = \lambda\varphi(x) \quad \text{und} \quad \psi''(t) = \lambda\psi(t).$$

Löse die DGL mit einem Exponentialansatz, $\varphi(x) = Ae^{\sqrt{\lambda}x} + Be^{-\sqrt{\lambda}x}$. Für eine oszillatorische Lösung muss λ negativ sein, also $\lambda = -m^2$ und $\psi''(t) = -m^2\psi(t)$ sowie $\varphi''(x) = -m^2\varphi(x)$. Dann ist

$$\begin{aligned} \varphi(x) &= A \cos(mx) + B \sin(mx) \\ \psi(t) &= A' \cos(mt) + B' \sin(mt). \end{aligned}$$

Setzt man wieder die Anfangsbedingungen ein, $\varphi(0) = \varphi(\pi) = 0$, so erhält man, dass $\varphi(x) = A \sin(mx)$ mit m ganzzahlig. Also ist

$$u_m(t, x) = (A_m \cos(mt) + B_m \sin(mt)) \sin(mx) \quad m \in \mathbb{Z}$$

eine Lösung der Wellengleichung. In Abbildung 2.1 sind beispielhaft einige Funktionen u_m gezeigt. Die allgemeine Lösung ergibt sich dann als Superposition der Einzellösungen:

$$u(t, x) = \sum_{m=1}^{\infty} (A_m \cos(mt) + B_m \sin(mt)) \sin(mx).$$

Die Koeffizienten A_m und B_m bestimmt man aus den Anfangsbedingungen:

$$\begin{aligned} f(x) &= u(0, x) = \sum_{m=1}^{\infty} A_m \sin(mx) \\ g(x) &= \frac{\partial}{\partial t} u(0, x) = \sum_{m=1}^{\infty} mB_m \sin(mx). \end{aligned}$$

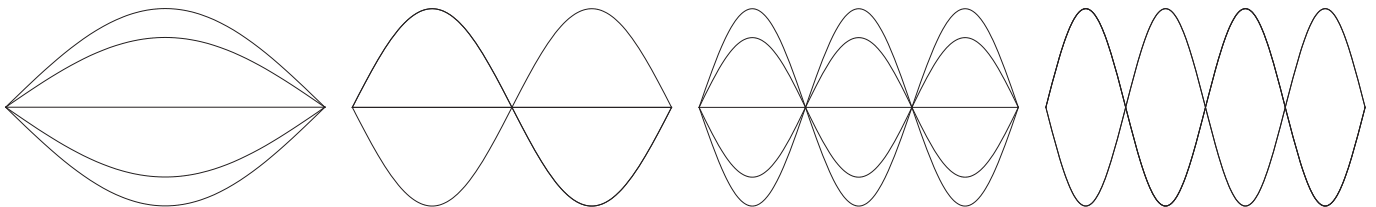


Abbildung 2.1: Funktionen $u_m(t, x) = \cos(mt) \sin(mx)$, die die Wellengleichung lösen. Es wurden die Graphen für $m = 1, \dots, 4$ gezeichnet, jeweils für $x \in [0, \pi]$ und $t = n\frac{\pi}{4}$ mit $n = 0, \dots, 4$.

2.1 Fourier-Reihen

Wir behandeln die folgende Problemstellung: Gegeben eine periodische (möglicherweise komplexwertige) Funktion $f(x)$ mit, zum Beispiel) Periode 2π , also

$$f(x + 2\pi) = f(x).$$

Wir kennen solche periodischen Funktionen, nämlich e^{imx} , für $x \in \mathbb{Z}$. Also jede Linearkombination $\sum c_m e^{imx}$ ist auch eine periodische Funktion.

Frage: Gilt umgekehrt dass jede periodische Funktion $F(x)$ als trigonometrisches Polynom

$$F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx}$$

dargestellt werden kann. Wenn ja, in welchem Sinne gilt Gleichheit. wir werden folgendes zeigen:

- Wenn F stetig, dann gilt es „fast“
- Definiere $S_N(F)(x) = \sum_{m=-N}^N a_m e^{imx}$, dann gilt $\int_{-\pi}^{\pi} |S_N(F)(x) - F(x)|^2 dx \rightarrow 0$ für $N \rightarrow \infty$. Also konvergiert $S_N(F)(x)$ gegen $F(x)$ im L^2 -Sinne (Lebesgue-Integral):

$$L^2([-\pi, \pi]) = \left\{ f \mid \int_{-\pi}^{\pi} |f(x)|^2 dx < \infty \right\}.$$

- Anders formuliert bedeutet das, wenn F integrierbar, dann konvergiert $S_N(F)(x) \rightarrow F(x)$ fast überall. Um dies zu beweisen werden wir als Zwischenschritt das Mittel, genannt Cesaro-Mittel, $\frac{1}{N} \sum_{i=1}^N S_i(F)$ betrachten und zeigen, dass dies für alle stetigen Funktionen F konvergiert. Dies impliziert dann auch dass alle stetigen Funktionen auf einem endlichen Intervall durch trigonometrische Polynome beliebig genau approximiert werden können.

2.1 Bemerkung. Es gelte nun $F(x) = \sum_m a_m e^{imx}$. Berechne folgendes Integral:

$$\int_{-\pi}^{\pi} e^{imx} e^{-inx} dx = \int_{-\pi}^{\pi} e^{i(m-n)x} dx = \begin{cases} 0 & m \neq n \\ 2\pi & m = n \end{cases} = 2\pi \delta_{mn}.$$

Ersetzt man in diesem Integral nun e^{imx} durch F , so erhält man

$$\int_{-\pi}^{\pi} F(x) e^{-inx} dx = \sum_m a_m \int_{-\pi}^{\pi} e^{i(m-n)x} dx = \sum_m a_m 2\pi \delta_{mn} = 2\pi a_n.$$

Also liefert

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(x) e^{-inx} dx$$

den n -ten Fourier-Koeffizient.

2.1.0.1 Bemerkung.

Es gibt einen natürlichen Zusammenhang zwischen 2π periodischen Funktionen und Funktionen, die auf dem Einheitskreis $e^{i\theta}$ gegeben sind. Ein Punkt auf dem Einheitspreis ist durch $e^{i\theta}$ gegeben, wobei θ eine reelle Zahl ist die eindeutig bis auf ganzzahlige Vielfache von 2π sind. Wenn $F(x)$ eine Funktion am Kreis ist, dann ist die Funktion f , mit

$$f(\theta) = F(e^{i\theta})$$

periodisch mit Periode 2π . Die Integrabilitätsbedingung, Stetigkeit und andere Glattheitsbedingungen von F sind durch f gegeben. Zum Beispiel ist F integrierbar wenn f integrierbar ist auf irgendeinem Intervall der Länge 2π . Also F ist stetig wenn f stetig ist auf \mathbb{R} . Da f Periode 2π hat können wir f auf irgendein Intervall der Länge 2π reduzieren und bekommen dennoch die ursprüngliche Funktion F am Kreis Siehe Abbildung 2.2 für ein Beispiel. Bemerke dass f denselben Wert an den Endpunkten des Intervalls haben muss. Umgekehrt jede Funktion f auf $[0, 2\pi]$ gibt eine stetige Funktion am Kreis genau dann wenn

$$f(0) = f(2\pi).$$

Also, Funktionen f auf \mathbb{R} , die 2π periodisch sind und Funktionen auf einem Intervall 2π die denselben Wert an den Endpunkten annehmen beschreiben ein und das selbe Objekt, nämlich *Funktionen am Kreis*.

2.1.0.2 Definition. Fourier-Koeffizient

Sei $f(x)$ Riemann-integrierbar auf $[a, b]$ mit $b - a = L$, dann ist der n -te Fourier-Koeffizient definiert durch

$$a_n = \hat{f}(n) = \frac{1}{L} \int_a^b f(x) e^{-\frac{2\pi}{L} inx} dx.$$

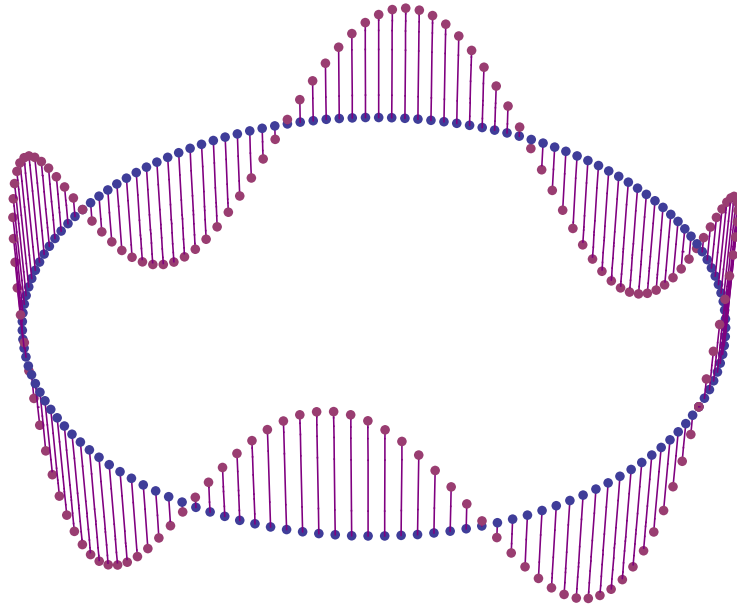


Abbildung 2.2: Eine Funktion auf dem Kreis; eine Sinusfunktion wäre in dieser Darstellung lediglich eine Ellipse, welche gegen das Nullniveau geneigt ist.

2.1.0.3 Bemerkung.

Sei f auf $[-\pi, \pi]$ integrierbar, dann ist der n -te Fourier-Koeffizient

$$\hat{f}(n) = a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta.$$

Rein formal ist die Fourier Reihe von f gegeben durch

$$\sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

Die Frage ist, wie oben erwähnt, in welchem Sinne konvergiert die Partialsumme

$$S_N(f) = \sum_{n=-N}^N a_n e^{in\theta}$$

gegen f ??

2.1.0.4 Beispiel.

Als konkretes Beispiel einer Funktion auf $[-\pi, \pi]$ wählen wir die Sägezahnfunktion, $F(\theta) = \theta$. Die Fourier-Koeffizienten sind

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \theta e^{-in\theta} d\theta = \frac{(-1)^{n+1}}{in}.$$

Die Reihenentwicklung ist

$$F(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta} = \sum_{n=-\infty}^{\infty} \frac{(-1)^{n+1}}{in} e^{in\theta} = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(n\theta).$$

Man bemerkt, dass die Reihe nicht absolut konvergiert, da die a_n mit $\frac{1}{n}$ abfallen. Wir werden später sehen in welchem Sinne = wirklich gilt.

Wenn wir annehmen könnten, dass Die Fourier-Reihe $S_N(f)$ gegen f konvergiert in geeignetem Sinne, dann kann man schliessen, dass eine Funktion eindeutig durch die Fourier-Koeffizienten gegeben ist. Das würde folgendes bedeuten: wenn f und g die selben Fourier-Koeffizienten haben, dann müssen f und g sogar notwendigerweise gleich sein. Oder anders formuliert Falls $\hat{f}(n) = 0$ für alle $n \in \mathbb{Z}$, dann muss $f = 0$ sein.

In der Allgemeinheit kann die Aussage nicht stimmen, denn wenn zwei Funktionen sich nur auf endlich vielen Punkten unterscheiden, dann haben diese dieselben Fourier-Koeffizienten, weil diese durch Integration gewonnen werden. Aber falls f stetig ist, dann gilt die Aussage, wie folgendes Theorem zeigt:

2.1.0.5 Theorem.

Sei f integrierbar am Kreis mit $\hat{f}(n) = 0 \forall n \in \mathbb{Z}$. Dann gilt $f(\theta_0) = 0 \Leftrightarrow f$ stetig im Punkt θ_0 .

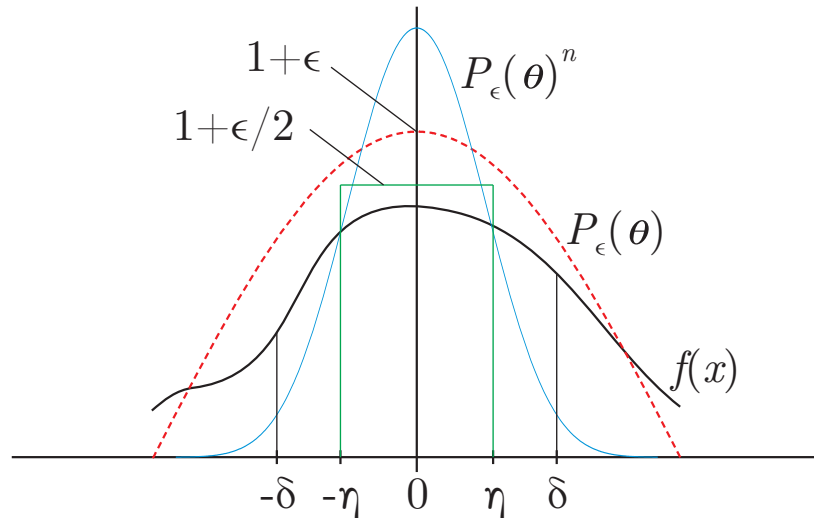


Abbildung 2.3: Veranschaulichung des Beweises zu Theorem 2.1.0.5.

Proof. Wähle o.B.d.A. $\theta_0 = 0$ und nimm an, dass $f(0) > 0$, f reell. Definiere $P_\epsilon(\theta) := \epsilon + \cos \theta$. Wähle nun $\delta, \epsilon, \eta > 0$ entsprechend der folgenden Kriterien (Stetigkeit von f im Punkt $\theta_0 = 0$, vgl. Abbildung 2.3):

1. δ so, dass $f(\theta) > \frac{1}{2}f(0)$ für $|\theta| \leq \delta$
2. ϵ so, dass $|P_\epsilon(\theta)| \leq 1 - \frac{\epsilon}{2}$ für $|\theta| \geq \delta$
3. η so, dass $|P_\epsilon(\theta)| \geq 1 + \frac{\epsilon}{2}$ für $|\theta| \leq \eta$

Da nach Voraussetzung $\hat{f}(n) = 0 \forall n$, gilt $\int_{-\pi}^{\pi} f(\theta)P_\epsilon(\theta)^n d\theta = 0$. Betrachte zuerst den Fall $|\theta| \geq \delta$:

$$\left| \int_{|\theta| \geq \delta} f(\theta)P_\epsilon(\theta)^n d\theta \right| \leq 2\pi \sup_{\theta \in [-\pi, \pi]} |f(\theta)| \left(1 - \frac{\epsilon}{2}\right)^n \xrightarrow{n \rightarrow \infty} 0.$$

Für $|\theta|$ zwischen η und δ ist das Integral sicher positiv, da f stetig und $f(\theta) > 0$ in einer Umgebung von 0:

$$\int_{\eta \leq |\theta| \leq \delta} f(\theta)P_\epsilon(\theta)^n d\theta \geq 0.$$

Für $|\theta|$ kleiner als η gilt dann

$$\int_{|\theta| \leq \eta} f(\theta)P_\epsilon(\theta)^n d\theta \geq \frac{f(0)}{2} \int_{|\theta| \leq \eta} \left(1 + \frac{\epsilon}{2}\right)^n d\theta \xrightarrow{n \rightarrow \infty} \infty.$$

Damit folgt, dass

$$\int_{-\pi}^{\pi} f(\theta)P_\epsilon(\theta)^n d\theta \xrightarrow{n \rightarrow \infty} \infty.$$

Dies ist allerdings ein Widerspruch zur Annahme, dass $f(0) > 0$ und somit kann nur $f(0) = 0$ gelten. \square

2.1.0.6 Korollar.

Wenn f stetig ist am Kreis und $\hat{f}(n) = 0 \forall n \in \mathbb{Z}$, dann ist $f \equiv 0$.

2.1.0.7 Korollar.

Ist f stetig am Kreis und konvergieren die Fourier-Koeffizienten absolut, ist also die Summe $\sum_{n=-\infty}^{\infty} |a_n|$ konvergent, $a_n = \hat{f}(n)$, dann konvergieren die Partialsummen $S_N(f)(\theta)$ gleichmäßig gegen $f(\theta)$. Dies ist äquivalent zur Formulierung $f(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta} \forall \theta \in [-\pi, \pi]$.

Proof. Definiere $g(\theta) := \sum_{n=-\infty}^{\infty} a_n e^{in\theta}$ mit $a_n = \hat{f}(n)$. Die Partialsummen der Fourier-Reihe von f konvergieren gleichmäßig gegen g , denn

$$\begin{aligned} |S_N(f)(\theta) - g(\theta)| &= \left| \sum_{n=-N}^N a_n e^{in\theta} - \sum_{n=-\infty}^{\infty} a_n e^{in\theta} \right| = \left| \sum_{n=N+1}^{\infty} a_n e^{in\theta} + \sum_{n=-\infty}^{-(N+1)} a_n e^{in\theta} \right| \\ &= \left| \sum_{|n|>N} a_n e^{in\theta} \right| \leq \sum_{|n|>N} |a_n| \xrightarrow{N \rightarrow \infty} 0 \quad \forall \theta \in [-\pi, \pi]. \end{aligned}$$

Oder anders formuliert:

$$\forall \epsilon > 0 \exists N_\epsilon \in \mathbb{N} \forall n > N_\epsilon : |S_n(f)(\theta) - g(\theta)| < \epsilon \quad \forall \theta \in [-\pi, \pi].$$

Da der gleichmäßige Limes stetiger Funktionen wieder stetig ist, ist g stetig. Definiere nun $h(\theta) := f(\theta) - g(\theta)$. Die Funktion h ist stetig, da sie eine Verkettung stetiger Funktionen ist. Die Fourierkoeffizienten von f und g sind die selben, also gilt $\hat{h}(n) = \hat{f}(n) - \hat{g}(n) = 0 \forall n \in \mathbb{Z}$. Nach 2.1.0.5 folgt dann $h(\theta) = 0$ und damit $f(\theta) = g(\theta)$. □

2.1.0.8 Korollar.

Sei $f(\theta) \in C^2([-\pi, \pi])$ (zweimal stetig differenzierbar am Kreis). Dann gilt $|\hat{f}(n)| \leq c \frac{1}{n^2}$ und damit $S_N(f)(\theta) \rightarrow f(\theta)$ (denn die Fourier-Koeffizienten konvergieren dann absolut).

Proof. Integriere zweimal partiell:

$$\begin{aligned} 2\pi \hat{f}(n) &= \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta \\ &= \underbrace{\frac{1}{-in} f(\theta) e^{-in\theta} \Big|_{-\pi}^{\pi}}_{=0} - \frac{1}{-in} \int_{-\pi}^{\pi} f'(\theta) e^{-in\theta} d\theta \\ &= \underbrace{\frac{1}{n^2} f'(\theta) e^{-in\theta} \Big|_{-\pi}^{\pi}}_{=0} - \frac{1}{n^2} \int_{-\pi}^{\pi} f''(\theta) e^{-in\theta} d\theta. \end{aligned}$$

Also gilt für den Betrag der Fourier-Koeffizienten

$$|\hat{f}(n)| \leq \frac{1}{2\pi n^2} \int_{-\pi}^{\pi} |f''(\theta)| d\theta = c \frac{1}{n^2}.$$

Nach Theorem 2.1.0.7 konvergieren die Partialsummen dann gegen die Funktion, da die Fourier-Koeffizienten absolut konvergieren. Bemerke, dass dies z.B. bei der Sägezahnfunktion nicht der Fall ist, diese ist nicht zweimal stetig differenzierbar und die Fourier-Koeffizienten fallen nur ab wie $\frac{1}{n}$. □

2.1.0.9 Bemerkung.

Bildet man die Fourier-Koeffizienten von der Ableitung einer Funktion die stetig differenzierbar am Kreis ist, so erhält man durch partielle Integration

$$\hat{f}'(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(\theta) e^{-in\theta} d\theta = \frac{in}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta = in \hat{f}(n).$$

Wir sehen damit, dass die Glattheit von Funktionen in x in direktem Zusammenhang zur Schnelligkeit des Abfalls von $\hat{f}(n)$ stehen.

Wir wissen jetzt auch dass für 2-mal stetigen Funktionen die Fourier-Reihe punktweise gegen f konvergiert. Wie können die Partialsumme etwas eleganter schreiben und definieren nun die Faltung von Funktionen.

2.1.0.10 Definition. Faltung

Seien f, g 2π -periodische Funktionen auf \mathbb{R} . Die FALTUNG von f mit g ist definiert durch

$$(f * g)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)dy.$$

2.1.0.11 Satz.

Seien f, g und h 2π -periodische, integrierbare Funktionen auf \mathbb{R} . Dann gilt

- (a) $(f * g) * h = f * (g * h)$
- (b) $f * g(x) = g * f(x)$
- (c) $f * (g + h) = f * g + f * h$
- (d) $f * g$ ist stetig.
- (e) $\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$

Proof. Für (b): Substituiere im Integral $y = x - z$ und nutze die Periodizität der Funktionen:

$$\begin{aligned} f * g(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)dy = \frac{1}{2\pi} \int_{x+\pi}^{x-\pi} g(z)f(x-z)dz \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(z)f(x-z)dz = g * f(x). \end{aligned}$$

Beweis für (e):

$$\begin{aligned} \widehat{f * g}(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)g(x-y)dy e^{-inx} dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)e^{-iny} g(x-y)e^{iny} e^{-inx} dy dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)e^{-iny} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x-y)e^{-in(x-y)} dx dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y)e^{-iny} \frac{1}{2\pi} \int_{-\pi}^{\pi} g(u)e^{-inu} du dy = \hat{f}(n)\hat{g}(n). \end{aligned}$$

Der schwierigste Teil ist (d). Dies sagt aus, dass das Falten die Regularität verbessert. Wir skizzieren hier nur kurz den Beweis. Wir zeigen in den Übungen, dass die Faltung zweier stetiger Funktionen wieder stetig ist. Nun verwenden wir, dass weiters jede Riemann-integrierbare Funktion f durch eine Folge von stetigen Funktionen f_n approximiert werden kann, sodass

$$\int_{-\pi}^{\pi} |f(x) - f_n(x)| dx \rightarrow_{n \rightarrow \infty} 0.$$

Um dies zu sehen kann man die Riemann-integrierbarkeit verwenden und bis auf ein ε die Funktion f mittels Stufenfunktion f^* approximieren, auf einer Zerlegung $\{x_1, \dots, x_N\}$. Dieses f^* lässt sich wiederum bis auf ε durch eine stetige Funktion \tilde{f} approximieren, wobei wir f^* von $x_i - \delta$ bis $x_i + \delta$ linear ergänzen und so \tilde{f} erzeugen. Da δ beliebig klein werden kann und N fest ist, läßt sich die Differenz $\int_{-\pi}^{\pi} |f^*(x) - \tilde{f}(x)| dx$ kleiner als das vorgegebene ε machen. Als letzter Schritt finden wir nun stetige f_n, g_n , die f, g approximieren. Es genügt nun zu zeigen dass

$$\lim_{n \rightarrow \infty} \|f_n * g_n - f * g\|_{\infty} = \sup_{x \in [-\pi, \pi]} \lim_{n \rightarrow \infty} \|f_n * g_n(x) - f * g(x)\| \rightarrow 0,$$

Denn da der gleichmäßige Limes stetiger Funktionen stetig ist, muss $f * g$ stetig sein. Da

$$\|f_n * g_n - f * g\|_\infty \leq \|f_n * g_n - f_n * g\|_\infty + \|f_n * g - f * g\|_\infty,$$

und zum Beispiel

$$\|f_n * g - f * g\|_\infty = \sup_x \int_{-\pi}^{\pi} |f_n(x-y) - f(x-y)| |g(y)| dy \leq \sup_y |g(y)| \int_{-\pi}^{\pi} |f_n(x-y) - f(x-y)| dy \xrightarrow{n \rightarrow \infty} 0,$$

woraus obige Behauptung folgt. □

2.1.1 Integralkerne

Die Partialsumme einer Fourier-Reihe ist darstellbar als Faltung mit einer neuen Funktion:

$$\begin{aligned} S_N(f)(x) &= \sum_{|n| \leq N} a_n e^{inx} = \sum_{|n| \leq N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iny} dy e^{inx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) \sum_{|n| \leq N} e^{in(x-y)} dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) D_N(x-y) dy = f * D_N(x) \end{aligned}$$

Dabei heißt $D_N(x) := \sum_{|n| \leq N} e^{inx}$ Dirichlet-Kern.

2.1.1.1 Definition. Familie guter Kerne

Es sei $\{K_n(x)\}_{n=1}^\infty$ eine Familie von Integralkernen auf dem Kreis. $\{K_n(x)\}$ heißt FAMILIE GUTER KERNE, wenn

- (a) $\forall n \geq 1 : \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(x) dx = 1$
- (b) $\exists M > 0 : \forall n \geq 1 : \frac{1}{2\pi} \int_{-\pi}^{\pi} |K_n(x)| dx \leq M$
- (c) $\forall \delta > 0 : \frac{1}{2\pi} \int_{\delta \leq |x| \leq \pi} |K_n(x)| dx \rightarrow 0$ für $n \rightarrow \infty$.

2.1.1.2 Beispiel.

Betrachte eine Funktion, die $\int g(x) dx = 1$ und $g(x) \geq 0$ erfüllt. Definiere $K_n(x) := ng(nx)$. Dann ist $\int K_n(x) dx = n \int g(nx) \frac{d(nx)}{n} = \int g(x) dx = 1$.

2.1.1.3 Theorem.

Sei $\{K_n(x)\}$ Familie guter Kerne, f integrierbare Funktion auf dem Kreis. Dann gilt $f * K_n(x) \rightarrow f(x)$ für $n \rightarrow \infty \forall x$ mit $f(x)$ stetig. Wenn f stetig auf dem Kreis ist die Konvergenz sogar gleichmäßig.

Proof. f stetig in x bedeutet $\forall \epsilon > 0 \exists \delta > 0 : |f(x-y) - f(x)| < \epsilon \forall |y| \leq \delta$. Unter Ausnutzung der Tatsache, dass $\int_{-\pi}^{\pi} K_n(x) dx = 1$ und $f * g = g * f$ schreibt man

$$\begin{aligned} f * K_n(x) - f(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} K_n(y) (f(x-y) - f(x)) dy \\ &= \frac{1}{2\pi} \int_{|y| \leq \delta} K_n(y) (f(x-y) - f(x)) dy \\ &\quad + \frac{1}{2\pi} \int_{\delta \leq |y| \leq \pi} K_n(y) (f(x-y) - f(x)) dy. \end{aligned}$$

Nun lässt sich der Betrag dieses Ausdrucks abschätzen mithilfe der Stetigkeitsforderung an f :

$$\begin{aligned} |f * K_n(x) - f(x)| &\leq \frac{1}{2\pi} \int_{|y| \leq \delta} |K_n(y)| |f(x-y) - f(x)| dy \\ &\quad + \frac{1}{2\pi} \int_{\delta \leq |y| \leq \pi} |K_n(y)| 2 \sup_{z \in [-\pi, \pi]} |f(z)| dy \\ &\leq M\epsilon + F(n) \text{ mit } F(n) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

□

Lassen sie uns bemerken, dass die Dirichlet Kerne D_N keine Familie guter Kerne bildet. Falls wir dies Kerne aber mitteln, dann bekommen wir eine Familie guter Kerne und dies werden wir später nutzen.

2.1.1.4 Definition. Fejér-Kerne

Der FEJÉR-KERN ist definiert durch

$$F_N(x) := \frac{1}{N} \sum_{k=0}^{N-1} D_k(x).$$

Als CESÀRO-MITTEL bezeichnet man

$$\sigma_N(f)(x) := \frac{1}{N} \sum_{k=1}^N S_k(f)(x) = f * F_N(x).$$

2.1.1.5 Beispiel.

Betrachte die Folge $a_n = (-1)^n = (-1, 1, -1, 1, \dots)$, deren Partialsumme lautet $S_N = \sum_{n \leq N} a_n = (-1, 0, -1, 0, \dots)$. Bildet man das Cesàro-Mittel davon erhält man $\sigma_N = (-1, 0, -\frac{1}{3}, 0, -\frac{1}{5}, 0, \dots) \rightarrow 0$.

2.1.1.6 Lemma.

Die diskrete Darstellung des Fejér-Kerns ist gegeben durch

$$F_N(x) = \frac{1}{N} \frac{\sin^2\left(\frac{Nx}{2}\right)}{\sin^2\left(\frac{x}{2}\right)}$$

und F_N ist eine Familie guter Kerne.

Proof. Die diskrete Darstellung erhält man durch Anwendung der Summenformel der geometrischen Reihe und wird hier nicht näher ausgeführt (Übungen). Die drei Eigenschaften der guten Kerne zeigt man folgendermaßen:

- (a) $\forall n \geq 1 : \frac{1}{2\pi} \int_{-\pi}^{\pi} F_N(x) dx = \frac{1}{2\pi N} \sum_{k=1}^N \int_{-\pi}^{\pi} D_k(x) dx = 1$.
- (b) Da die F_N stets positiv sind, ist $|F_N(x)| = F_N(x)$ und damit die obere Grenze durch 1 gegeben.
- (c) Wähle $\delta > 0$ beliebig aber fest.

$$\begin{aligned} \frac{1}{2\pi} \int_{\delta \leq |x| \leq \pi} F_N(x) dx &= \frac{1}{2\pi N} \int_{\delta \leq |x| \leq \pi} \frac{\sin^2\left(\frac{Nx}{2}\right)}{\sin^2\left(\frac{x}{2}\right)} dx \\ &\leq \frac{1}{N} \int_{\delta \leq |x| \leq \pi} \frac{1}{\sin^2\left(\frac{x}{2}\right)} dx \\ &= \frac{1}{N} M \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

□

2.1.1.7 Bemerkung.

Die Konsequenz aus 2.1.1.6 ist, dass $\{\sigma_N\}$ eine Familie guter Kerne ist, also $\sigma_N(f)(x) \rightarrow f(x)$ gleichmäßig, wenn f stetig auf dem Kreis ist.

2.1.1.8 Theorem.

Stetige Funktionen am Kreis können gleichmäßig approximiert werden durch trigonometrische Polynome.

Proof. f ist stetig, daraus folgt direkt, dass $\sigma_N(f)(x) = (f * F_N)(x) \rightarrow f(x)$ gleichmäßig. Dabei ist $\sigma_N(f)(x) = \sum_{|n| \leq N} c_n (e^{ix})^n$ trigonometrisches Polynom, dessen Koeffizienten c_n nicht notwendigerweise gleich den Fourier-Koeffizienten $a_n = \hat{f}(n)$ der Funktion sein müssen. □

2.1.2 Wiederholung: Hilbertraum und Skalarprodukt

2.1.2.1 Definition. Skalarprodukt

Ein SKALARPRODUKT (inneres Produkt) auf einem komplexen Vektorraum \mathcal{H} (welcher dann Prä-Hilbertraum heißt) ist eine Abbildung $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$ für die $\forall \phi, \psi, \chi \in \mathcal{H}$ und $\alpha \in \mathbb{C}$ folgendes gilt:

- (a) $\langle \phi, \phi \rangle \geq 0$, $\langle \phi, \phi \rangle = 0 \Leftrightarrow \phi = 0$ (pos. Definitheit)
- (b) $\langle \phi, \chi + \psi \rangle = \langle \phi, \chi \rangle + \langle \phi, \psi \rangle$
- (c) $\langle \phi, \alpha \psi \rangle = \alpha \langle \phi, \psi \rangle$ (Linearität in 2. Komponente)
- (d) $\langle \phi, \psi \rangle = \overline{\langle \psi, \phi \rangle}$ (Hermitezität)

2.1.2.2 Bemerkung.

Aus der Definition des Skalarprodukts folgt, dass

$$\langle \phi + \chi, \psi \rangle = \langle \phi, \psi \rangle + \langle \chi, \psi \rangle \quad \text{und} \quad \langle \alpha \phi, \psi \rangle = \bar{\alpha} \langle \phi, \psi \rangle.$$

Wegen der Linearität in der zweiten Komponente und der Antilinearität in der ersten Komponente nennt man solche Abbildungen Sesquilinearformen.

2.1.2.3 Definition. Orthogonalität

Zwei Vektoren $\phi, \psi \in \mathcal{H}$ heißen ORTHOGONAL, falls $\langle \phi, \psi \rangle = 0$. Eine Folge $(\phi_j)_{j=1}^\infty$ heißt ORTHONORMALFOLGE (ONF), falls $\langle \phi_j, \phi_k \rangle = \delta_{jk}$. Für jedes $\psi \in \mathcal{H}$ und beliebiges $n \in \mathbb{N}$ gilt bezüglich einer beliebigen Orthonormalfolge (ϕ_j) , dass

$$\psi = \underbrace{\sum_{j=1}^n \langle \phi_j, \psi \rangle \phi_j}_{:=\psi_n} + \underbrace{\psi - \sum_{j=1}^n \langle \phi_j, \psi \rangle \phi_j}_{:=\psi_n^\perp}.$$

Dabei ist ψ_n orthogonal zu ψ_n^\perp , denn $\langle \psi_n, \psi_n^\perp \rangle = 0$.

2.1.2.4 Satz. Pythagoras

Der SATZ VON PYTHAGORAS lautet in der allgemeinen Form (verwende die Orthogonalität von ψ_n und ψ_n^\perp)

$$\|\psi\|^2 = \langle \psi_n + \psi_n^\perp, \psi_n + \psi_n^\perp \rangle = \|\psi_n\|^2 + \|\psi_n^\perp\|^2 = \sum_{j=1}^n |\langle \phi_j, \psi \rangle|^2 + \|\psi_n^\perp\|^2.$$

2.1.2.5 Korollar.

Sei (ϕ_j) ONF. Dann gilt $\forall \psi, \phi \in \mathcal{H}$ und $n \in \mathbb{N}$, dass

$$\|\psi\|^2 \geq \sum_{j=1}^n |\langle \phi_j, \psi \rangle|^2 \quad \text{und} \quad |\langle \phi, \psi \rangle| \leq \|\phi\| \|\psi\|$$

(Bessel-Ungleichung) und (Schwarz-Ungleichung)

Proof. Die Bessel-Ungleichung folgt direkt aus dem Satz von Pythagoras. Die Schwarz-Ungleichung erhält man aus dem Satz des Pythagoras als Spezialfall für $n = 1$: Sei $\phi_1 = \frac{\phi}{\|\phi\|}$ und $\phi \neq 0$, dann ist

$$\|\psi\|^2 \geq \left| \left\langle \frac{\phi}{\|\phi\|}, \psi \right\rangle \right|^2 = \frac{1}{\|\phi\|^2} |\langle \phi, \psi \rangle|^2$$

und es folgt direkt die Schwarz-Ungleichung. □

2.1.2.6 Korollar.

Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf einem komplexen Vektorraum \mathcal{H} . Dann definiert $\|\psi\| := \sqrt{\langle \psi, \psi \rangle}$ eine Norm auf \mathcal{H} und das Skalarprodukt ist stetig bezüglich dieser Norm.

Proof. Die ersten beiden Eigenschaften der Norm (pos. Definitheit und Homogenität) sind sofort bewiesen und werden daher ausgelassen. Die Dreiecks-Ungleichung folgt aus der Schwarz-Ungleichung und aus $\operatorname{Re}(z) \leq |z| \forall z \in \mathbb{C}$:

$$\begin{aligned}\|\phi + \psi\|^2 &= \langle \phi + \psi, \phi + \psi \rangle \\ &= \|\phi\|^2 + \|\psi\|^2 + 2 \operatorname{Re} \langle \phi, \psi \rangle \\ &\leq (\|\phi\| + \|\psi\|)^2.\end{aligned}$$

Für den Beweis der Stetigkeit definiere konvergente Folgen $\phi_n \rightarrow \phi$ und $\psi_n \rightarrow \psi$ in \mathcal{H} . Da in metrischen Räumen Folgenstetigkeit identisch zur Stetigkeit ist, folgt die Stetigkeit aus

$$\begin{aligned}\lim_{n \rightarrow \infty} |\langle \phi_n, \psi_n \rangle - \langle \phi, \psi \rangle| &\leq \lim_{n \rightarrow \infty} (|\langle \phi_n - \phi, \psi_n \rangle| + |\langle \phi, \psi_n - \psi \rangle|) \\ &\leq \lim_{n \rightarrow \infty} (\|\phi_n - \phi\| \|\psi_n\| + \|\phi\| \|\psi_n - \psi\|) \\ &= 0.\end{aligned}$$

□

2.1.2.7 Definition. Hilbertraum

Ein HILBERTRAUM ist ein vollständiger, normierter Vektorraum, dessen Norm $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$ durch das Skalarprodukt gegeben ist. ERINNERUNG: Vollständig heißt, dass jede Cauchy-Folge gegen ein Element des Raumes konvergiert.

2.1.3 L^2 -Konvergenz von Fourier-Reihen

Für zwei Riemann-integrierbare Funktionen f und g ist das innere Produkt gegeben durch

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx$$

$(\mathcal{R}, \langle \cdot, \cdot \rangle)$ bildet einen Prä-Hilbertraum mit Norm $\|f\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$. $(L^2, \langle \cdot, \cdot \rangle)$ bildet einen Hilbertraum.

2.1.3.1 Theorem.

Sei f Riemann-integrierbar am Kreis. Dann gilt

$$\|S_N - f\|^2 = \frac{1}{2\pi} \int_0^{2\pi} |S_N(f)(x) - f(x)|^2 dx \xrightarrow{N \rightarrow \infty} 0.$$

Proof. Für alle ganzen Zahlen n sei $e_n(\theta) = e^{in\theta}$. Somit ist die Familie $\{e_n\}_{n \in \mathbb{Z}}$ orthogonal, also $\langle e_n, e_m \rangle = \delta_{nm}$. Sei nun f integrierbar auf dem Kreis und seien a_n die Fourier-Koeffizienten von f ; diese lassen sich über

$$\langle f, e_n \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta = a_n$$

darstellen. Insbesondere gilt

$$S_N(f) = \sum_{|n| \leq N} a_n e_n.$$

Dann folgt aus der Orthogonalitätseigenschaft der Familie $\{e_n\}_{n \in \mathbb{Z}}$ und der Darstellung der a_n , dass $f - \sum_{|n| \leq N} a_n e_n$ orthogonal zu allen e_n mit $|n| \leq N$ ist. Daher muss

$$f - \sum_{|n| \leq N} a_n e_n \perp \sum_{|n| \leq N} b_n e_n \quad \forall b_n \in \mathbb{C}$$

gelten. Daraus folgt nun zweierlei:

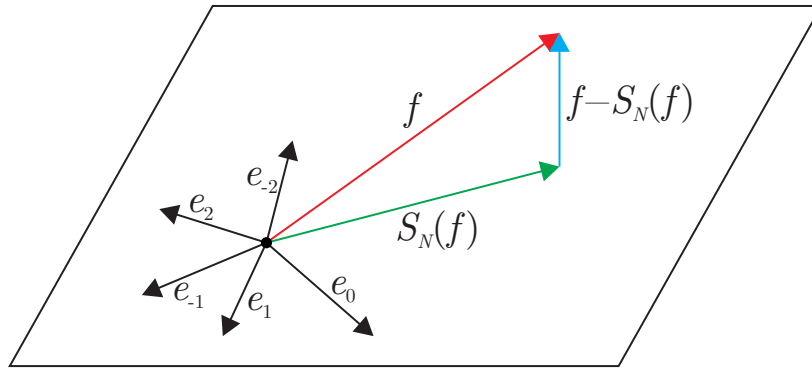


Abbildung 2.4: Die Differenz von f und $S_N(f)$ steht senkrecht auf allen Basisvektoren e_m , aus denen sich $S_N(f)$ zusammensetzt.

1. Wendet man den allgemeinen Satz von Pythagoras (Satz 2.1.2.4) auf

$$f = f - \sum_{|n| \leq N} a_n e_n + \sum_{|n| \leq N} a_n e_n$$

an (wähle also $b_n = a_n$), so ergibt sich

$$\|f\|^2 = \left\| f - \sum_{|n| \leq N} a_n e_n \right\|^2 + \left\| \sum_{|n| \leq N} a_n e_n \right\|^2.$$

Wegen Orthogonalität gilt aber auch

$$\left\| \sum_{|n| \leq N} a_n e_n \right\|^2 = \sum_{|n| \leq N} |a_n|^2$$

und damit folgt schließlich

$$\|f\|^2 = \|f - S_N(f)\|^2 + \sum_{|n| \leq N} |a_n|^2$$

2. Ferner nehmen wir im Moment das Lemma 2.1.3.2 an, welches aussagt, dass $\|f - S_N(f)\|^2$ nach oben mit beliebig anderen Fourier-Koeffizienten c_{-N}, \dots, c_N abgeschätzt werden kann. Wir können also auch die Koeffizienten des Cesaro-Mittels nehmen und wissen mit Theorem 2.1.1.8 dass $\|S_N(f) - f\| \rightarrow 0$ für $N \rightarrow \infty$.

Falls f aber nur integrierbar ist, so kann man f nicht notwendigerweise durch trigonometrische Polynome punktweise approximieren; vielmehr verwenden wir, dass wenn f auf dem Kreis integrierbar und durch B beschränkt ist, eine Folge $(f_k)_{k=1}^\infty$ stetiger Funktionen auf dem Kreis existiert, sodass

$$\sup_{x \in [0, 2\pi]} |f_k(x)| \leq B \quad \forall k \in \mathbb{N}$$

und

$$\int_0^{2\pi} |f(x) - f_k(x)| dx \xrightarrow{k \rightarrow \infty} 0.$$

Wähle also nun eine stetige Funktion g auf dem Kreis, welche den Einschränkungen

$$\sup_{\theta \in [0, 2\pi]} |g(\theta)| \leq \sup_{\theta \in [0, 2\pi]} |f(\theta)| = B$$

und

$$\int_0^{2\pi} |f(\theta) - g(\theta)| d\theta < \epsilon^2$$

genügt. Dann erhält man

$$\begin{aligned}\|f - g\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)|^2 d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)| |f(\theta) - g(\theta)| d\theta \\ &\leq \frac{2B}{2\pi} \int_0^{2\pi} |f(\theta) - g(\theta)| d\theta \\ &\leq C\epsilon^2\end{aligned}$$

für eine Konstante C . Nun lässt sich g aber wegen Theorem 2.1.1.8 durch ein trigonometrisches Polynom p approximieren, sodass $\|g - p\| < \epsilon$ für ein vorgegebenes $\epsilon > 0$. Dann gilt aber auch mit einer Konstanten C' die Abschätzung $\|f - p\| < C'\epsilon$. Daraus folgt dann mit Lemma 2.1.3.2, dass die Partialsummen der Fourier-Reihe von f im Sinne der Behauptung gegen f konvergieren. \square

2.1.3.2 Lemma. Bestmögliche Approximation

Sei f integrierbar am Kreis mit Fourier-Koeffizienten a_n , dann gilt $\|f - S_N\| \leq \|f - \sum_{|n| \leq N} c_n e_n\| \forall c_n$ mit $|n| \leq N$. Es gilt Gleichheit genau dann, wenn $c_n = a_n \forall |n| \leq N$.

Proof. Verwende die Orthogonalität von $S_N(f) - f$ und e_m mit $|m| \leq N$:

$$\begin{aligned}f - \sum_{|n| \leq N} c_n e_n &= f - S_N + \sum_{|n| \leq N} b_n e_n \quad \text{mit } b_n := a_n - c_n \\ \left\| f - \sum_{|n| \leq N} c_n e_n \right\|^2 &= \left\langle f - S_N + \sum_{|n| \leq N} b_n e_n, f - S_N + \sum_{|n| \leq N} b_n e_n \right\rangle \\ &= \langle f - S_N, f - S_N \rangle + \left\langle \sum_{|n| \leq N} b_n e_n, \sum_{|n| \leq N} b_n e_n \right\rangle \\ &= \|f - S_N\|^2 + \left\| \sum_{|n| \leq N} b_n e_n \right\|^2.\end{aligned}$$

Da der hintere Term größer oder gleich Null ist, folgt sofort die zu zeigende Ungleichung:

$$\|f - S_N\|^2 \leq \left\| f - \sum_{|n| \leq N} c_n e_n \right\|^2.$$

\square

2.1.3.3 Theorem. Parseval's Identität

Es sei f Riemann-integrierbar auf $[0, 2\pi]$, dann gilt

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |a_n|^2.$$

Proof. Durch Einfügen einer Null und durch Grenzwertbildung $N \rightarrow \infty$ erhält man:

$$\begin{aligned}\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx &= \|f\|^2 = \|f - S_N(f) + S_N(f)\|^2 \\ &= \underbrace{\|f - S_N(f)\|^2}_{\rightarrow 0} + \underbrace{\|S_N(f)\|^2}_{\rightarrow \sum_{n=-\infty}^{\infty} |a_n|^2}.\end{aligned}$$

Denn:

$$\left\langle \sum_m a_m e_m, \sum_n a_n e_n \right\rangle = \sum_{m,n} \langle a_m e_m, a_n e_n \rangle = \sum_{m,n} a_n \overline{a_m} \underbrace{\langle e_m, e_n \rangle}_{=\delta_{nm}} = \sum_n |a_n|^2.$$

□

2.1.3.4 Theorem. Riemann-Lebesgue-Lemma

Sei f integrierbar am Kreis. Dann gilt

$$|a_n| = |\hat{f}(n)| \xrightarrow{n \rightarrow \infty} 0 \Leftrightarrow \int_0^{2\pi} f(x) \sin(nx) dx \xrightarrow{n \rightarrow \infty} 0 \wedge \int_0^{2\pi} f(x) \cos(nx) dx \xrightarrow{n \rightarrow \infty} 0.$$

2.1.3.5 Lemma.

Seien F, G integrierbar am Kreis, $F(x) \sim \sum_n a_n e_n$ und $G(x) \sim \sum_n b_n e_n$, dann gilt

$$\frac{1}{2\pi} \int_0^{2\pi} F(x) \overline{G(x)} dx = \sum_{n=-\infty}^{\infty} a_n \overline{b_n}.$$

Proof. Übung.

□

2.1.3.6 Beispiel. Anwendung

Sei Γ eine Kurve mit Länge l , A die eingeschlossene Fläche. Wir möchten nun bestimmen, wann A maximal wird. Sei dazu $\gamma(t), t \in [a, b]$, die Parametrisierung von $\Gamma, l = \int_a^b |\gamma'(t)| dt$. Die Bogenlängenparametrisierung lautet dann $\gamma(s), s \in [0, l]$ mit $|\gamma'(s)| = 1, s \mapsto \gamma(s) = (x(s), y(s))$, wobei $|\gamma'(s)|^2 = x'(s)^2 + y'(s)^2 = 1$. Mittels Green's Theorem erhält man für die eingeschlossene Fläche

$$A = \iint_A 1 dx dy = \frac{1}{2} \int_{\Gamma} x dy - y dx = \frac{1}{2} \int_{\Gamma} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot d\mathbf{x}.$$

Das bedeutet also man kann schreiben $A = \frac{1}{2} \int_0^l (x(s)y'(s) - y(s)x'(s)) ds$. Die Behauptung ist nun, dass

$$A \leq \frac{l^2}{4\pi}, \quad \text{Gleichheit, wenn } \Gamma \text{ ein Kreis ist.}$$

Um die Behauptung zu zeigen, wähle o.B.d.A. $l = 2\pi$ und stelle die Komponenten der Parametrisierung als Fourier-Reihen dar: $x(s) = \sum_{n=-\infty}^{\infty} a_n e^{ins}, y(s) = \sum_{n=-\infty}^{\infty} b_n e^{ins}$. Dies ist möglich, da $x(0) = x(2\pi)$ und $y(0) = y(2\pi)$ und die Kurve als zweimal stetig differenzierbar angenommen wird. Jetzt bleibt also zu zeigen, dass $A \leq \pi$. Die Ableitungen der Komponenten der Parametrisierung lauten $x'(s) = \sum_{n=-\infty}^{\infty} i n a_n e^{ins}, y'(s) = \sum_{n=-\infty}^{\infty} i n b_n e^{ins}$. Zeige für späteren Gebrauch folgende Gleichung mittels Parseval's Identität (2.1.3.3)

$$1 = \frac{1}{2\pi} \int_0^{2\pi} (x'(s)^2 + y'(s)^2) ds = \sum_{n=-\infty}^{\infty} n^2 (|a_n|^2 + |b_n|^2).$$

Da $x(s)$ und $y(s)$ reell sind, kann man Lemma 2.1.3.5 anwenden und erhält für die Fläche

$$A = \frac{1}{2} \int_0^{2\pi} (x(s)y'(s) - y(s)x'(s)) ds = i\pi \sum_{n=-\infty}^{\infty} n (\overline{a_n} b_n - a_n \overline{b_n}).$$

Da $|x| = |\overline{x}|$ kann man mittels Dreiecksungleichung und binomischer Formel weiter abschätzen:

$$\left| \sum_{n=-\infty}^{\infty} i n (\overline{a_n} b_n - a_n \overline{b_n}) \right| \leq \sum_{n=-\infty}^{\infty} |n| 2 |a_n| |b_n| \leq \sum_{n=-\infty}^{\infty} |n| (|a_n|^2 + |b_n|^2).$$

Setze dies nun in den Ausdruck für die Fläche ein, ersetze mittels des ersten hergeleiteten Ausdrucks und erhalte

$$A \leq \pi \sum_{n=-\infty}^{\infty} |n|(|a_n|^2 + |b_n|^2) \leq \pi \sum_{n=-\infty}^{\infty} n^2(|a_n|^2 + |b_n|^2) = \pi.$$

Für Gleichheit muss $n = 1, 0, -1$ sein und deswegen auch $x(s) = a_0 + a_1 e^{is} + a_{-1} e^{-is}$, $y(s) = b_0 + b_1 e^{is} + b_{-1} e^{-is}$. Da x und y beide reell sind, fordert man $\overline{a_{-1}} = a_1$, $\overline{b_{-1}} = b_1$ sowie $a_0, b_0 \in \mathbb{R}$. Bestimme nun a_1 und b_1 , setze dazu ein und führe die Summen aus:

$$1 = 2(|a_1|^2 + |b_1|^2) \quad \wedge \quad 4|a_1||b_1| = 1$$

Also gilt $|a_1| = |b_1| = \frac{1}{2}$, d.h. $a_1 = \frac{1}{2}e^{i\alpha}$, $b_1 = \frac{1}{2}e^{i\beta}$. Setzt man dies in x und y ein ergibt sich

$$x(s) = a_0 + \cos(\alpha + s), \quad y(s) = a_0 \pm \sin(\alpha + s).$$

Bei Gleichheit ist die Kurve also tatsächlich ein Kreis.

2.1.4 Fourier-Transformation

2.1.4.1 Definition. kontinuierliche Fourier-Transformierte

Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ beschränkt, $|f(x)| \leq \frac{c}{1+x^2}$. Die KONTINUIERLICHE FOURIER-TRANSFORMIERTE ist definiert durch

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi kx} dx = \lim_{N \rightarrow \infty} \int_{-N}^N f(x)e^{-i2\pi kx} dx.$$

Diese ist auch beschränkt, denn $|\hat{f}(k)| \leq \int_{-\infty}^{\infty} |f(x)| dx < \infty$.

2.1.4.2 Definition. Schwartz-Raum

Wir definieren den Schwartz-Raum als

$$S(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid \sup_{x \in \mathbb{R}} |x|^l |f^{(k)}(x)| < \infty \quad \forall l, k \in \mathbb{N} \right\}.$$

S bezeichnet also die Menge aller Funktionen, für die alle Ableitungen existieren und die schneller abfallen als jedes Polynom. Z.B. ist $e^{-x^2} \in S$, aber $e^{-|x|} \notin S$.

2.1.4.3 Satz.

Sei $f \in S$. Dann gilt

- (a) $\widehat{f(x+h)}(k) = \hat{f}(k)e^{i2\pi hk}$
- (b) $\widehat{f(x)e^{-i2\pi xh}}(k) = \hat{f}(k+h)$
- (c) $\widehat{f(\delta x)}(k) = \frac{1}{\delta} \hat{f}\left(\frac{k}{\delta}\right)$
- (d) $\widehat{f'(x)}(k) = i2\pi k \hat{f}(k)$
- (e) $\widehat{-2\pi i x f(x)}(k) = \frac{d}{dk} \hat{f}(k)$

Proof. Für (c): Substituiere $y = \delta x$ und erhalte

$$\widehat{f(\delta x)}(k) = \int_{-\infty}^{\infty} f(\delta x)e^{-i2\pi kx} dx = \frac{1}{\delta} \int_{-\infty}^{\infty} f(y)e^{i2\pi \frac{k}{\delta} y} dy = \frac{1}{\delta} \hat{f}\left(\frac{k}{\delta}\right).$$

Für (d): Integriere partiell und nutze aus, dass die Randterme verschwinden, da f hinreichend schnell abfällt:

$$\widehat{f'(x)}(k) = \int_{-\infty}^{\infty} f'(x)e^{-i2\pi kx} dx = i2\pi k \int_{-\infty}^{\infty} f(x)e^{-i2\pi kx} dx = i2\pi k \hat{f}(k).$$

□

2.1.4.4 Theorem.

Sei $f(x) = e^{-\pi x^2}$, dann ist $f(x) = \hat{f}(x)$.

Proof. Es sei zuerst festgestellt, dass

$$\int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1,$$

was man durch

$$\left(\int_{-\infty}^{\infty} e^{-\pi x^2} dx \right)^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\pi(x^2+y^2)} dx dy = \int_0^{\infty} 2\pi \int_0^{\infty} e^{-\pi r^2} r dr d\theta = 2\pi \int_0^{\infty} r e^{-\pi r^2} dr = 1 \quad (2.1)$$

Definiere

$$F(\xi) = \hat{f}(\xi) = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x \xi} dx,$$

und beobachte $F(0) = 1$, nach obiger Rechnung. Wir sehen einfach, dass $f'(x) = -2\pi x f(x)$, und deshalb gilt

$$F'(\xi) = \int_{-\infty}^{\infty} f(x) i(-2\pi i x) e^{-\pi x^2} e^{-2\pi i x \xi} dx.$$

Mittels unserer Beziehung zwischen Fourier-transformierter und seiner Ableitung, Satz 2.1.4.3 Gl. (d), wissen wir weiters, dass

$$F'(\xi) = i(2\pi i \xi) \hat{f}(\xi) = -2\pi \xi F(\xi).$$

Definiere nun

$$G(\xi) = F(\xi) e^{\pi \xi^2},$$

dann schliessen wir dass $G'(\xi) = 0$, und dass $G(\xi) = 1$ sein muss, da $F(0) = 1$. Deshalb muss $F(\xi) = e^{-\pi \xi^2}$ sein.

Also ist $f(x) = \hat{f}(x)$. □

2.1.4.5 Definition. Familie guter Kerne (auf \mathbb{R})

Es sei $\{K_\delta(x)\}$ eine Familie von Integralkernen auf \mathbb{R} . $\{K_\delta(x)\}$ heißt FAMILIE GUTER KERNE, wenn

- (a) $\forall \delta > 0 : \int_{-\infty}^{\infty} K_\delta(x) dx = 1$
- (b) $\exists M > 0 : \forall \delta > 0 : \int_{-\infty}^{\infty} |K_\delta(x)| dx \leq M$
- (c) $\forall \eta > 0 : \int_{|x| \geq \eta} |K_\delta(x)| dx \rightarrow 0$ für $\delta \rightarrow 0$.

2.1.4.6 Theorem.

Definiere $K_\delta(x) := \frac{1}{\delta^{1/2}} e^{-\pi \frac{x^2}{\delta}}$. Dann ist $\{K_\delta(x)\}$ eine Familie guter Kerne.

Proof. Die ersten beiden Eigenschaften sind leicht zu zeigen, für die dritte wähle $\eta > 0$ beliebig, aber fest. Substituiere $y = \frac{x}{\delta^{1/2}}$

$$\frac{1}{\delta^{1/2}} \int_{|x| \geq \eta} e^{-\pi \frac{x^2}{\delta}} dx = \int_{|y| \geq \frac{\eta}{\delta^{1/2}}} e^{-\pi y^2} dy \xrightarrow{\delta \rightarrow 0} 0.$$

□

2.1.4.7 Definition. Faltung (auf \mathbb{R})

Seien $f, g \in S(\mathbb{R})$. Die Faltung ist dann definiert als

$$f * g(x) = \int_{-\infty}^{\infty} f(y) g(x - y) dy.$$

2.1.4.8 Satz.

Seien $f, g \in S(\mathbb{R})$. Dann gilt

- (a) $f * g = g * f$
 (b) $f * g \in S(\mathbb{R})$
 (c) $\widehat{f * g}(k) = \hat{f}(k)\hat{g}(k)$

Proof. Für (c): Durch die Substitution $z = x - y$ erhält man

$$\begin{aligned}\widehat{f * g}(k) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y)g(x-y)e^{-i2\pi xk} dy dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i2\pi yk} f(y)g(x-y)e^{-i2\pi(x-y)k} dy dx \\ &= \int_{-\infty}^{\infty} f(y)e^{-i2\pi yk} \int_{-\infty}^{\infty} g(z)e^{-i2\pi zk} dz dy \\ &= \hat{f}(k)\hat{g}(k).\end{aligned}$$

□

2.1.4.9 Korollar.

Sei $f \in S(\mathbb{R})$. Dann konvergiert $f * K_{\delta}(x) \rightarrow f(x)$ gleichmäßig für $\delta \rightarrow 0$.

Proof. Da $f \in S(\mathbb{R})$ gilt für ein $\epsilon > 0$, $\exists R : |f(x)| < \epsilon$ für $|x| \geq R$. f ist auf dem Intervall $[-R, R]$ gleichmäßig stetig, das bedeutet $\exists \eta > 0 : \forall x \in [-R, R] : |f(x-y) - f(x)| < \epsilon$ für $|y| < \eta$. Da K_{δ} eine Familie guter Kerne ist, kann man schreiben

$$f * K_{\delta}(x) - f(x) = \int_{-\infty}^{\infty} K_{\delta}(y)(f(x-y) - f(x)) dy.$$

Schätze nun den Betrag ab:

$$\begin{aligned}|f * K_{\delta}(x) - f(x)| &\leq \int_{|y| \leq \eta} |K_{\delta}(y)(f(x-y) - f(x))| dy + \int_{|y| \geq \eta} |K_{\delta}(y)(f(x-y) - f(x))| dy \\ &\leq \underbrace{\epsilon \int_{-\infty}^{\infty} |K_{\delta}(y)| dy}_{=1} + M \underbrace{\int_{|y| \geq \eta} |K_{\delta}(y)| dy}_{\rightarrow 0 \text{ für } \delta \rightarrow 0}.\end{aligned}$$

□

2.1.4.10 Satz.

Seien $f, g \in S(\mathbb{R})$, dann gilt

$$\int_{-\infty}^{\infty} f(x)\hat{g}(x) dx = \int_{-\infty}^{\infty} \hat{f}(y)g(y) dy.$$

Proof. Definiere $F(x, y) = f(x)g(y)e^{-i2\pi yx}$. Wenn nun F über x und y integriert wird, kann nach Fubini die Integrationsreihenfolge vertauscht werden:

$$\iint_{\mathbb{R}^2} F(x, y) d(x, y) = \underbrace{\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} F(x, y) dy \right) dx}_{=\int_{-\infty}^{\infty} f(x)\hat{g}(x) dx} = \underbrace{\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} F(x, y) dx \right) dy}_{=\int_{-\infty}^{\infty} \hat{f}(y)g(y) dy}.$$

□

2.1.4.11 Theorem.

Sei $f \in S(\mathbb{R})$, dann ist die Umkehrung der Fourier-Transformation gegeben durch

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k)e^{i2\pi xk} dk.$$

Proof. Definiere $G_\delta(x) := e^{-\pi x^2 \delta}$, die Fourier-Transformierte von G_δ ist $\hat{G}_\delta(k) = \frac{1}{\sqrt{\delta}} e^{-\pi \frac{k^2}{\delta}}$, denn definiert man $f(x) = e^{-\pi x^2}$ und verwendet, dass $\hat{f}(x) = f(x)$ und $\widehat{f(\delta x)}(k) = \frac{1}{\delta} \hat{f}\left(\frac{k}{\delta}\right)$ erhält man

$$\hat{G}_\delta(k) = \widehat{f(\sqrt{\delta}x)}(k) = \frac{1}{\sqrt{\delta}} \hat{f}\left(\frac{k}{\sqrt{\delta}}\right) = \frac{1}{\sqrt{\delta}} e^{-\pi \frac{k^2}{\delta}} = K_\delta(k).$$

Zeige zuerst, dass $f(0) = \int_{-\infty}^{\infty} \hat{f}(k) dk$ gilt. Verwende Satz 2.1.4.10 und schreibe

$$\int_{-\infty}^{\infty} f(x) \hat{G}_\delta(x) dx = \int_{-\infty}^{\infty} \hat{f}(k) G_\delta(k) dk \xrightarrow{\delta \rightarrow 0} \int_{-\infty}^{\infty} \hat{f}(k) dk.$$

Es gilt aber ebenso

$$\int_{-\infty}^{\infty} f(x) \hat{G}_\delta(x) dx = \int_{-\infty}^{\infty} f(x) K_\delta(x) dx = f * K_\delta(0) \xrightarrow{\delta \rightarrow 0} f(0),$$

nach Korollar 2.1.4.9. Also gilt insgesamt $f(0) = \int_{-\infty}^{\infty} \hat{f}(k) dk$. Definiere nun eine Funktion $F(y) := f(x+y)$ und schreibe

$$f(x) = F(0) = \int_{-\infty}^{\infty} \hat{F}(k) dk = \int_{-\infty}^{\infty} \hat{f}(k) e^{i2\pi x k} dk.$$

Beim letzten Gleichheitszeichen wurde verwendet, dass

$$\begin{aligned} \hat{F}(k) &= \int_{-\infty}^{\infty} F(y) e^{-i2\pi y k} dy = \int_{-\infty}^{\infty} f(x+y) e^{-i2\pi y k} dy \\ &= e^{i2\pi x k} \int_{-\infty}^{\infty} f(x+y) e^{-i2\pi(x+y)k} dy = e^{i2\pi x k} \int_{-\infty}^{\infty} f(z) e^{-i2\pi z k} dz \\ &= \hat{f}(k) e^{i2\pi x k}. \end{aligned}$$

□

2.1.4.12 Bemerkung.

Eine andere Schreibweise für die Fourier-Transformation ist $\hat{f}(k) = \mathcal{F}(f)(k) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi x k} dx$. Die Rücktransformation wird dann als $\mathcal{F}^{-1}(g)(x) = \mathcal{F}^*(g)(x) = \int_{-\infty}^{\infty} g(k) e^{i2\pi x k} dk$ geschrieben. Dabei gilt auf $S(\mathbb{R})$ (auch auf $L^2(\mathbb{R})$), dass $\mathcal{F}^* \mathcal{F} = \text{Id}$ und $\mathcal{F} \mathcal{F}^* = \text{Id}$.

2.1.4.13 Theorem. Plancherel/Parseval

Sei $f \in S(\mathbb{R})$. Dann gilt

$$\|f\|^2 = \|\hat{f}\|^2 \Leftrightarrow \int_{-\infty}^{\infty} |f(x)|^2 dx = \int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk.$$

Proof. Definiere $f_-(x) := f(-x)$, wie weiter oben schon bewiesen gilt außerdem $g(0) = \int_{-\infty}^{\infty} \hat{g}(k) dk$. Schreibe dann

$$\begin{aligned} \int_{-\infty}^{\infty} |f(x)|^2 dx &= \int_{-\infty}^{\infty} f(x) \overline{f(x)} dx = f * \overline{f_-}(0) \\ &= \int_{-\infty}^{\infty} \widehat{f * \overline{f_-}}(k) dk = \int_{-\infty}^{\infty} \hat{f}(k) \widehat{\overline{f_-}}(k) dk \\ &= \int_{-\infty}^{\infty} |\hat{f}(k)|^2 dk, \end{aligned}$$

denn es ist

$$\widehat{\overline{f_-}}(k) = \int_{-\infty}^{\infty} \overline{f(-x)} e^{-i2\pi x k} dx = \int_{-\infty}^{\infty} \overline{f(x)} e^{i2\pi x k} dx = \overline{\int_{-\infty}^{\infty} f(x) e^{-i2\pi x k} dx} = \overline{\hat{f}(k)}.$$

□

2.1.5 Anwendungsbeispiele von Fourier-Transformationen

2.1.5.1 Beispiel. Wärmeleitungsgleichung auf \mathbb{R}

Die Wärmeleitungsgleichung hat im Reellen in einer Dimension folgende Form:

$$\frac{\partial}{\partial t} u(x, t) = \frac{\partial^2}{\partial x^2} u(x, t).$$

Dabei bezeichnet $u(x, t)$ die Temperatur am Ort x zur Zeit t . Die Anfangstemperatur ist vorgegeben durch $u(x, 0) = f(x)$. Wendet man nun eine Fourier-Transformation auf beide Seiten an, nimmt die Gleichung folgende Form an:

$$\frac{\partial}{\partial t} \hat{u}(k, t) = -4\pi^2 k^2 \hat{u}(k, t)$$

Eine Separation der Variablen führt auf

$$\frac{\partial}{\partial t} \log \hat{u}(k, t) = -4\pi^2 k^2 \quad \Rightarrow \quad \hat{u}(k, t) = \hat{A}(k) e^{-4\pi^2 k^2 t},$$

wobei $\hat{u}(k, 0) = \hat{A}(k) = \hat{f}(k)$. Also lautet die Lösung im k -Raum $\hat{u}(k, t) = \hat{f}(k) e^{-4\pi^2 k^2 t}$. Im x -Raum erhält man $u(x, t) = f * H_t(x)$ wobei

$$H_t(x) = \int_{-\infty}^{\infty} e^{-4\pi^2 k^2 t} e^{i2\pi k x} dk = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}$$

der Wärmeleitungskern (Heat Kernel) ist. Die Lösung ist dann

$$\Rightarrow u(x, t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{\infty} f(y) e^{-\frac{(x-y)^2}{4t}} dy = f * H_t(x).$$

Da H_t eine Familie guter Kerne ist, gilt $f * H_t(x) \rightarrow f(x)$ für $t \rightarrow \infty$.

2.1.5.2 Theorem. Poisson'sche Summenformel

Sei $f \in S(\mathbb{R})$. Dann gilt $\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{i2\pi n x}$, für den Spezialfall $x = 0$ erhält man

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

Proof. Definiere $F_1(x) := \sum_{n=-\infty}^{\infty} f(x+n)$, $F_2(x) := \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{i2\pi n x}$. Man bemerkt zunächst, dass $F_2(x+1) = F_2(x)$, also F_2 periodisch mit Periode 1 ist. Weiterhin ist F_2 stetig, denn die Reihe konvergiert absolut, da die Fourier-Koeffizienten schnell genug abfallen ($f \in S(\mathbb{R})$). Die Fourier-Transformierte $\hat{f}(n)$ liefert also die Fourier-Koeffizienten der Reihenentwicklung von F_2 . F_1 ist ebenfalls periodisch mit Periode 1, denn

$$F_1(x+1) = \sum_{n=-\infty}^{\infty} f(x+n+1) = \sum_{m=-\infty}^{\infty} f(x+m) = F_1(x).$$

Da $f \in S(\mathbb{R})$ ist auch F_1 stetig. Nun bleibt also zu zeigen, dass F_1 und F_2 dieselben Fourier-Koeffizienten haben, denn dann sind sie identisch auf $[0, 1]$. Substituiere dazu $y = x+n$ im Integral:

$$\begin{aligned} \hat{F}_1(m) &= \int_0^1 F_1(x) e^{-i2\pi m x} dx = \int_0^1 \sum_{n=-\infty}^{\infty} f(x+n) e^{-i2\pi m x} dx \\ &= \sum_{n=-\infty}^{\infty} \int_0^1 f(x+n) e^{-i2\pi m x} dx = \sum_{n=-\infty}^{\infty} \int_n^{n+1} f(y) e^{-i2\pi m y} \underbrace{e^{i2\pi m n}}_{=1} dy \\ &= \int_{-\infty}^{\infty} f(y) e^{-i2\pi m y} dy = \hat{f}(m). \end{aligned}$$

□

2.1.5.3 Theorem.

Sei $\psi \in S(\mathbb{R})$ mit $\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1$. Dann gilt

$$\int_{-\infty}^{\infty} x^2 |\psi(x)|^2 dx \int_{-\infty}^{\infty} 4\pi^2 k^2 |\hat{\psi}(k)|^2 dk \geq \frac{1}{4}.$$

Die Ungleichung gilt auch für $x \mapsto (x - x_0)$ und $k \mapsto (k - k_0)$.

Proof. Aus der Normierungsbedingung erhält man mittels partieller Integration

$$1 = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = - \int_{-\infty}^{\infty} x \frac{d}{dx} |\psi(x)|^2 dx = - \int_{-\infty}^{\infty} x (\psi'(x) \overline{\psi(x)} + \psi(x) \overline{\psi'(x)}) dx.$$

Bildet man den Betrag dieses Ausdrucks und wendet die Dreiecksungleichung sowie die Cauchy-Schwarz-Ungleichung an, erhält man

$$1 \leq 2 \int_{-\infty}^{\infty} |x| |\psi'(x)| |\psi(x)| dx \leq 2 \left(\int_{-\infty}^{\infty} |x|^2 |\psi(x)|^2 dx \right)^{1/2} \underbrace{\left(\int_{-\infty}^{\infty} |\psi'(x)|^2 dx \right)^{1/2}}_{= \left(\int_{-\infty}^{\infty} 4\pi^2 k^2 |\hat{\psi}(k)|^2 dk \right)^{1/2}}.$$

□

3 Ausflug in Lineare Algebra

3.1 Vektorräume mit Skalarprodukt

3.1 Definition. Skalarprodukt

Sei V ein Vektorraum über \mathbb{K} . Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$$

heißt **Skalarprodukt** oder **inneres Produkt**, falls

- (a) $\langle u, u \rangle \geq 0 \quad \forall u \in V$ und $\langle u, u \rangle = 0 \Leftrightarrow u = 0$. (positive Definitheit)
- (b) $\langle \cdot, \cdot \rangle$ ist linear im zweiten Argument, also

$$\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$$

für alle $u, v, w \in V$ und $\alpha, \beta \in \mathbb{K}$.

- (c) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ für alle $u, v \in V$. (Symmetrie)

Einen Vektorraum V mit Skalarprodukt nennt man Skalarproduktraum oder Prä-Hilbertraum. Im Falle $\mathbb{K} = \mathbb{R}$ spricht man auch von euklidischen Räumen.

- 3.2 Bemerkung.**
- Im Fall $\mathbb{K} = \mathbb{R}$ besagt (c), dass $\langle u, v \rangle = \langle v, u \rangle$. Zusammen mit (b) impliziert das die Linearität auch im ersten Argument. Für $\mathbb{K} = \mathbb{R}$ ist ein Skalarprodukt also eine Bilinearform.
 - Im Fall $\mathbb{K} = \mathbb{C}$ ist $\langle \cdot, \cdot \rangle$ antilinear im ersten Argument, denn

$$\langle \alpha u + \beta w, v \rangle = \overline{\langle v, \alpha u + \beta w \rangle} = \overline{\alpha \langle v, u \rangle + \beta \langle v, w \rangle} = \overline{\alpha} \langle u, v \rangle + \overline{\beta} \langle w, v \rangle.$$

Man spricht dann von einer Sesquilinearform (sesqui = $1\frac{1}{2}$).

- 3.3 Beispiele.** (a) \mathbb{R}^n mit dem natürlichen Skalarprodukt

$$\langle x, y \rangle_{\mathbb{R}^n} = \sum_{k=1}^n x_k y_k$$

- (b) \mathbb{C}^n mit dem natürlichen Skalarprodukt

$$\langle x, y \rangle_{\mathbb{C}^n} = \sum_{k=1}^n \overline{x_k} y_k$$

3.4 Definition. und Satz. Die quadratsummierbaren Folgen

Sei ℓ^2 der Vektorraum der komplexen Folgen $a = (a_1, a_2, \dots)$ für welche die Reihe $\sum_{k=1}^{\infty} |a_k|^2$ konvergiert. Für $a = (a_1, a_2, \dots), b = (b_1, b_2, \dots) \in \ell^2$ konvergiert die Reihe

$$\langle a, b \rangle_{\ell^2} := \sum_{k=1}^{\infty} \overline{a_k} b_k$$

und definiert ein Skalarprodukt auf ℓ^2 .

3 Ausflug in Lineare Algebra

Beweis. Für $a = (a_1, a_2, \dots) \in \ell^2$ und $\lambda \in \mathbb{C}$ konvergiert auch $\sum_{k=1}^{\infty} |\lambda a_k|^2 = |\lambda|^2 \sum_{k=1}^{\infty} |a_k|^2$. Für $a, b \in \ell^2$ gilt

$$|a_k + b_k|^2 \leq 2(|a_k|^2 + |b_k|^2)$$

und

$$|\bar{a}_k b_k| \leq \frac{1}{2}(|a_k|^2 + |b_k|^2)$$

und somit (nach dem Majorantenkriterium)

$$\sum |a_k + b_k|^2 \leq 2 \sum |a_k|^2 + 2 \sum |b_k|^2 < \infty$$

und

$$\sum |\bar{a}_k b_k| \leq \frac{1}{2} \sum |a_k|^2 + \frac{1}{2} \sum |b_k|^2 < \infty.$$

Also sind auch λa und $a + b$ quadratsummierbar und der Ausdruck $\langle a, b \rangle_{\ell^2}$ ist endlich für $a, b \in \ell^2$.

Dass $\langle \cdot, \cdot \rangle$ ein Skalarprodukt ist, rechnet man einfach nach:

- (a) $\langle a, a \rangle = \sum \bar{a}_k a_k = \sum |a_k|^2 \geq 0$ und $\sum |a_k|^2 = 0 \Leftrightarrow a_k = 0 \forall k \Leftrightarrow a = 0$.
- (b) $\langle a, \lambda b + c \rangle = \sum \bar{a}_k (\lambda b_k + c_k) = \lambda \sum \bar{a}_k b_k + \sum \bar{a}_k c_k = \lambda \langle a, b \rangle + \langle a, c \rangle$.
- (c) $\langle a, b \rangle = \sum \bar{a}_k b_k = \overline{\sum a_k \bar{b}_k} = \overline{\langle b, a \rangle}$.

□

3.5 Bemerkung. ℓ^2 ist in der Tat ein Hilbertraum. Dazu muss man zeigen, dass jede Cauchy-folge konvergiert. Also falls $\|x_n - x_m\|_2 \rightarrow 0$ für n, m gegen unendlich, gilt zu zeigen dass ein Grenzwert $\in \ell^2$ existiert. Übung.

3.6 Bemerkung. Für zwei Riemann-integrierbare Funktionen f und g ist das innere Produkt gegeben durch

$$\langle f, g \rangle = \int_{\mathbb{R}^n} \overline{f(x)} g(x) d^n x$$

$(\mathcal{R}, \langle \cdot, \cdot \rangle)$ bildet einen Prä-Hilbertraum mit Norm $\|f\|^2 = \int_{\mathbb{R}^n} |f(x)|^2 d^n x$. $(L^2, \langle \cdot, \cdot \rangle)$ bildet einen Hilbertraum.

3.7 Definition. Orthogonalität

- (a) Zwei Vektoren u, v eines Skalarproduktraumes V heißen **orthogonal**, falls

$$\langle u, v \rangle = 0.$$

Man schreibt dann auch $u \perp v$.

- (b) Eine endliche oder abzählbare Menge von Vektoren $\{v_1, v_2, \dots\}$ in V heißt Orthonormalsystem (ONS) oder Orthonormalfolge (ONF), falls

$$\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j. \end{cases}$$

Mit Hilfe eines Orthonormalsystems kann man einen Vektor orthogonal zerlegen. Sei v_1, v_2, \dots, v_N ein ONS in V , dann gilt für beliebiges $u \in V$ und $n \leq N$, dass

$$u = \underbrace{\sum_{j=1}^n \langle v_j, u \rangle v_j}_{=: u_n} + u - \underbrace{\sum_{j=1}^n \langle v_j, u \rangle v_j}_{=: u_n^\perp} = u_n + u_n^\perp.$$

Man rechnet leicht nach, dass $\langle u_n, u_n^\perp \rangle = 0$ gilt.

Diese simple Zerlegung hat zwei wichtige Konsequenzen:

3.8 Satz. Bessel und Cauchy-Schwarz

Sei V ein Skalarproduktraum und v_1, \dots, v_N ein ONS ($N = \infty$, d.h. eine Orthonormalfolge, ist erlaubt). Für $u \in V$ definieren wir

$$\|u\| := \sqrt{\langle u, u \rangle}.$$

Dann gilt für alle $u, v \in V$ und $n \leq N$ die

$$\text{Bessel'sche Ungleichung} \quad \|u\|^2 \geq \sum_{j=1}^n |\langle v_j, u \rangle|^2,$$

und die

$$\text{Cauchy-Schwarz'sche Ungleichung} \quad |\langle u, v \rangle| \leq \|u\| \cdot \|v\|.$$

Beweis. Wegen $\langle u_n, u_n^\perp \rangle = 0$ gilt der „Satz des Pythagoras“:

$$\begin{aligned} \|u\|^2 &= \langle u, u \rangle = \langle u_n + u_n^\perp, u_n + u_n^\perp \rangle = \langle u_n, u_n \rangle + \langle u_n^\perp, u_n \rangle + \langle u_n, u_n^\perp \rangle + \langle u_n^\perp, u_n^\perp \rangle = \|u_n\|^2 + \|u_n^\perp\|^2 \\ &= \left\langle \sum_{j=1}^n \langle v_j, u \rangle v_j, \sum_{k=1}^n \langle v_k, u \rangle v_k \right\rangle + \|u_n^\perp\|^2 \\ &= \sum_{j=1}^n \sum_{k=1}^n \overline{\langle v_j, u \rangle} \langle v_k, u \rangle \underbrace{\langle v_j, v_k \rangle}_{=\delta_{jk}} + \|u_n^\perp\|^2 = \sum_{j=1}^n |\langle v_j, u \rangle|^2 + \|u_n^\perp\|^2. \end{aligned}$$

Daraus folgt mit der Positivität des Skalarprodukts sofort die Bessel'sche Ungleichung. Die Cauchy-Schwarz'sche Ungleichung ist trivial für $u = 0$, sei also $u \neq 0$. Dann ist $u_1 = \frac{u}{\|u\|}$ ein ONS (zugegeben ein sehr kleines) und Bessel für $n = 1$ liefert Cauchy-Schwarz:

$$\|v\|^2 \geq \left| \left\langle \frac{u}{\|u\|}, v \right\rangle \right|^2 = \frac{1}{\|u\|^2} |\langle u, v \rangle|^2.$$

□

3.9 Definition. Normierter Raum

Eine Abbildung

$$\|\cdot\| : V \rightarrow [0, \infty)$$

auf einem \mathbb{K} -Vektorraum heißt Norm, falls

- (a) $\|u\| = 0 \Leftrightarrow u = 0$ (Definitheit)
- (b) $\|\lambda u\| = |\lambda| \|u\|$ für alle $\lambda \in \mathbb{K}, u \in V$ (Homogenität)
- (c) $\|u + v\| \leq \|u\| + \|v\|$ für alle $u, v \in V$ (Dreiecksungleichung)

Eine Norm weist also jedem Vektor u seine Länge $\|u\|$ zu. Das Tupel $(V, \|\cdot\|)$ heißt normierter Raum.

3.10 Satz. In einem Skalarproduktraum $(V, \langle \cdot, \cdot \rangle)$ wird durch

$$\|u\| := \sqrt{\langle u, u \rangle} \quad \text{für } u \in V$$

eine Norm auf V definiert.

Beweis. $\|u\| \geq 0$ und $\|u\| = 0 \Leftrightarrow u = 0$ folgen aus der positiven Definitheit (Definition 3.1 (a)) des Skalarproduktes. Die Homogenität folgt aus der Sesquilinearität des Skalarproduktes:

$$\|\lambda u\| = \sqrt{\langle \lambda u, \lambda u \rangle} = \sqrt{\lambda \bar{\lambda} \langle u, u \rangle} = |\lambda| \sqrt{\langle u, u \rangle} = |\lambda| \|u\|.$$

Die Dreiecksungleichung folgt aus Cauchy-Schwarz und $\text{Re}(z) \leq |z|$ für $z \in \mathbb{C}$.

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle = \langle u, u \rangle + \langle u, v \rangle + \overline{\langle u, v \rangle} + \langle v, v \rangle \\ &= \|u\|^2 + 2 \text{Re} \langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\| \|v\| + \|v\|^2 = (\|u\| + \|v\|)^2. \end{aligned}$$

□

3.11 Bemerkung. Sei ℓ^p der Vektorraum der komplexen Folgen $a = (a_1, a_2, \dots)$ mit der Norm

$$\|a\|_p = \left(\sum_{k=1}^{\infty} |a_k|^p \right)^{1/p}.$$

Dieser ist auch ein Banachraum.

ℓ^∞ ist dann der Banachraum der beschränkten Folgen mit Norm

$$\|a\|_\infty = \sup_n |a_n|.$$

3.12 Definition. Äquivalenz von Normen

Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf einem Vektorraum V heißen **äquivalent**, wenn es Konstanten $c, C > 0$ gibt, so dass für alle $x \in V$ gilt:

$$c \|x\|_1 \leq \|x\|_2 \leq C \|x\|_1.$$

(Umgekehrt gilt dann offensichtlich auch

$$\frac{1}{C} \|x\|_2 \leq \|x\|_1 \leq \frac{1}{c} \|x\|_2. \quad)$$

3.13 Bemerkung. Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ äquivalente Normen auf V . Dann gilt:

- (a) $\lim x_n = a$ in $(V, \|\cdot\|_1)$ \Leftrightarrow $\lim x_n = a$ in $(V, \|\cdot\|_2)$.
- (b) (x_n) ist Cauchy in $(V, \|\cdot\|_1)$ \Leftrightarrow (x_n) ist Cauchy in $(V, \|\cdot\|_2)$.
- (c) $(V, \|\cdot\|_1)$ ist vollständig \Leftrightarrow $(V, \|\cdot\|_2)$ ist vollständig.

Beweis. (a) Sei $\lim x_n = a$ in $(V, \|\cdot\|_1)$, also $\lim_{n \rightarrow \infty} \|x_n - a\|_1 = 0$. Dann ist auch $\lim_{n \rightarrow \infty} \|x_n - a\|_2 \leq \lim_{n \rightarrow \infty} C \|x_n - a\|_1 = 0$. (b) Analog. (c) folgt aus (a) und (b). □

3.14 Satz. Alle Normen im \mathbb{C}^n sind äquivalent.

Beweis. Es genügt zu zeigen, dass jede Norm $\|\cdot\|$ zu $\|\cdot\|_1$ mit $\|x\|_1 = \sum_{j=1}^n |x_j|$ äquivalent ist. Sei $x = \sum_{j=1}^n x_j e_j \in \mathbb{R}^n$. Dann gilt

$$\|x\| = \left\| \sum x_j e_j \right\| \leq \sum |x_j| \|e_j\| \leq n \max\{\|e_j\| \mid j = 1, \dots, n\} \sum |x_j| =: C \|x\|_1.$$

Angenommen es gibt kein c mit $c \|x\|_1 \leq \|x\|$ für alle $x \in \mathbb{R}^n$, dann existiert eine Folge (x_k) in \mathbb{R}^n mit

$$\frac{1}{k} \|x_k\|_1 \geq \|x_k\| \quad \text{und o.B.d.A.} \quad \|x_k\|_1 = 1. \quad (3.1)$$

Diese enthält eine bzgl. $\|\cdot\|_1$ konvergente Teilfolge, wieder mit (x_k) bezeichnet, da die Komponentenfolgen $(x_{k,j})$ beschränkte Folgen in \mathbb{R} sind und somit nach dem Satz von Bolzano-Weierstraß konvergente Teilfolgen enthalten. Wegen (3.1) konvergiert die Folge (x_k) auch in der $\|\cdot\|$ -Norm gegen den eindeutigen Grenzwert $x := \lim_{k \rightarrow \infty} x_k$. Aufgrund der Stetigkeit der Normen, welche eine Konsequenz der umgekehrten Dreiecksungleichung ist, ist dann einerseits $1 = \lim \|x_k\|_1 = \|\lim x_k\|_1 = \|x\|_1$, also $x \neq 0$. Andererseits ist $\|x\| = \lim \|x_k\| \leq \lim \frac{1}{k} \|x_k\|_1 = 0$, also $x = 0$. □

3.15 Bemerkung. Der Raum der Funktionen, in dem $|f|^p$ eine Riemann-integrierbare Funktion ist bildet mit der Norm

$$\|f\|_p = \left(\int_{\mathbb{R}^n} |f(x)|^p d^n x \right)^{1/p}$$

einen normierten Raum.

3.16 Lemma. Jedes Orthonormalsystem v_1, v_2, \dots ist linear unabhängig.

Beweis. Aus $\sum_{j=1}^n \alpha_j v_j = 0$ folgt durch Projektion auf v_k , $k \in \{1, \dots, n\}$,

$$0 = \left\langle v_k, \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{j=1}^n \alpha_j \underbrace{\langle v_k, v_j \rangle}_{\delta_{kj}} = \alpha_k.$$

□

3.17 Korollar. Orthonormalbasen

- (a) Ist v_1, \dots, v_n ein ONS in einem n -dimensionalen Skalarproduktraum V , so ist (v_1, \dots, v_n) eine Basis für V und heißt dann Orthonormalbasis (ONB).
 (b) Sei v_1, \dots, v_n eine ONB von V , dann ist die Basisdarstellung von $u \in V$ durch

$$u = \langle v_1, u \rangle v_1 + \dots + \langle v_n, u \rangle v_n = \sum_{k=1}^n \langle v_k, u \rangle v_k$$

gegeben.

- (c) Für $u, w \in V$ und v_1, \dots, v_n eine ONB ist

$$\langle u, w \rangle = \sum_{k=1}^n \overline{\langle v_k, u \rangle} \langle v_k, w \rangle.$$

Insbesondere gilt die Parseval'sche Gleichung

$$\|u\|^2 = \sum_{k=1}^n |\langle v_k, u \rangle|^2.$$

Beweis. (a) In einem n -dimensionalen Raum bilden n linear unabhängige Vektoren eine Basis.

- (b) Sei $u = \sum_{k=1}^n \alpha_j v_j$ die Basisdarstellung von u . Dann folgt durch Projektion auf v_k , $k \in \{1, \dots, n\}$, dass

$$\langle v_k, u \rangle = \langle v_k, \sum_{j=1}^n \alpha_j v_j \rangle = \sum_{j=1}^n \alpha_j \langle v_k, v_j \rangle = \alpha_k.$$

- (c) Für u, w gilt nach (b),

$$u = \sum_{j=1}^n \langle v_j, u \rangle v_j \quad \text{und} \quad w = \sum_{k=1}^n \langle v_k, w \rangle v_k,$$

also

$$\begin{aligned} \langle u, w \rangle &= \left\langle \sum_{j=1}^n \langle v_j, u \rangle v_j, \sum_{k=1}^n \langle v_k, w \rangle v_k \right\rangle = \sum_{j=1}^n \sum_{k=1}^n \overline{\langle v_j, u \rangle} \langle v_k, w \rangle \langle v_j, v_k \rangle \\ &= \sum_{j=1}^n \overline{\langle v_j, u \rangle} \langle v_j, w \rangle. \end{aligned}$$

□

3.18 Definition. Projektion

Sei V ein Vektorraum und $P \in \mathcal{L}(V)$. Dann heißt P **Projektion** falls,

- (a) $P^2 = P$,

und **orthogonale Projektion** bezüglich eines Skalarprodukts $\langle \cdot, \cdot \rangle$, falls zusätzlich

- (b) $\langle u, Pv \rangle = \langle Pu, v \rangle$ für alle $u, v \in V$.

3.19 Bemerkung. In der Quantenmechanik werden Projektionen auf einen normierten Vektor $|u\rangle$ als $|u\rangle\langle u|$ bezeichnet. Mit der Bedeutung, dass es auf einen Vektor $|v\rangle$ so wirkt

$$|u\rangle\langle u||v\rangle = u\langle u, v\rangle.$$

Als Beispiel, sei $|u\rangle = \frac{1}{\sqrt{14}}(1, 2, 3) \in \mathbb{C}^3$, dann ist aus mathematischer Sicht $|u\rangle\langle u|$ das Tensorprodukt

$$\frac{1}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} (1, 2, 3) = \frac{1}{14} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

Also allgemein, bei normiertem $|u\rangle = (u_1, u_2, u_3)$ ist

$$|u\rangle\langle u| = (a_{ij}) = (u_i u_j).$$

3.20 Definition. Orthogonales Komplement

Für nichtleere Teilmengen $A \subset V$ eines Skalarproduktraumes heißt der Unterraum

$$A^\perp = \{u \in V \mid \langle u, v \rangle = 0 \ \forall v \in A\}$$

das **orthogonale Komplement** von A .

3.21 Satz. Sei v_1, \dots, v_n ein ONS in V und sei $P \in \mathcal{L}(V)$ definiert durch

$$Pu := \sum_{k=1}^n \langle v_k, u \rangle v_k. \quad (= u_n)$$

Dann ist P eine orthogonale Projektion, genannt Projektion auf $W = \text{Span}\{v_1, \dots, v_n\}$, und es gilt für alle $u \in V$

(a) $Pu \in W$ und $(1 - P)u \in W^\perp$.

(b) $\|u - Pu\| = \min\{\|u - w\| \mid w \in W\}$

Beweis. (a) haben wir bereits gezeigt ($\langle u_n, u_n^\perp \rangle = 0$), siehe Übungsaufgabe.

(b) Sei $w \in W$, dann ist $w = \sum_{j=1}^n \lambda_j v_j$ für geeignete $\lambda_j \in \mathbb{K}$ und es gilt

$$\begin{aligned} \|u - w\|^2 &= \left\| u - \sum_{j=1}^n \lambda_j v_j \right\|^2 = \left\langle u - \sum_{j=1}^n \lambda_j v_j, u - \sum_{k=1}^n \lambda_k v_k \right\rangle \\ &= \langle u, u \rangle - \sum_{j=1}^n \bar{\lambda}_j \langle v_j, u \rangle - \sum_{k=1}^n \lambda_k \langle u, v_k \rangle + \sum_{j=1}^n |\lambda_j|^2 \\ &= \|u\|^2 + \underbrace{\sum_{j=1}^n |\lambda_j - \langle v_j, u \rangle|^2}_{\geq 0} - \sum_{j=1}^n |\langle v_j, u \rangle|^2 \end{aligned}$$

Das Minimum wird angenommen für $\lambda_j = \langle v_j, u \rangle$.

Wir zeigen noch, dass P eine orthogonale Projektion ist. Linearität folgt aus der Linearität des Skalarprodukts. Sei $w \in W$, dann gilt nach (b) $Pw = w$ und mit (a)

$$P^2u = \underbrace{P(Pu)}_{\in W} = Pu \quad \forall u \in V.$$

P ist also eine Projektion. Wegen

$$\langle Pu, v \rangle - \langle u, Pv \rangle = \underbrace{\langle Pu, v - Pv \rangle}_{\in W} - \underbrace{\langle u - Pu, Pv \rangle}_{\in W} = 0$$

ist sie auch orthogonal. □

Wie findet man Orthonormalsysteme bzw. Orthonormalbasen, und gibt es überhaupt immer welche?

Das Orthonormalisierungsverfahren von Gram-Schmidt

Sei V Skalarproduktraum und u_1, u_2, \dots eine endliche oder abzählbare linear unabhängige Menge von Vektoren in V . Wir konstruieren nun ein ONS v_1, v_2, \dots mit

$$\text{Span}\{v_1, \dots, v_n\} = \text{Span}\{u_1, \dots, u_n\} \quad \text{für alle } n \in \mathbb{N}.$$

1. Schritt: $v_1 = \frac{u_1}{\|u_1\|}$ erfüllt $\|v_1\| = \|\frac{u_1}{\|u_1\|}\| = 1$ und ist somit ein ONS mit $\text{Span}\{v_1\} = \text{Span}\{u_1\}$.

2. Schritt: Sei P_1 die orthogonale Projektion auf $W_1 = \text{Span}\{u_1\} = \text{Span}\{v_1\}$. Da $u_2 \notin W_1$, gilt $u_2 - P_1 u_2 \neq 0$ und wir können $v_2 = \frac{u_2 - P_1 u_2}{\|u_2 - P_1 u_2\|}$ definieren. Es gilt wieder $\|v_2\| = 1$ und $v_2 \in W_1^\perp$ da $u_2 - P_1 u_2 \in W_1^\perp$ und W_1^\perp Unterraum ist. Da v_1, v_2 beide in $W_2 = \text{Span}\{u_1, u_2\}$ liegen und linear unabhängig sind, folgt $\text{Span}\{v_1, v_2\} = W_2$.

n -ter Schritt: Seien v_1, \dots, v_{n-1} schon konstruiert mit

$$\text{Span}\{v_1, \dots, v_{n-1}\} = \text{Span}\{u_1, \dots, u_{n-1}\} = W_{n-1}$$

und sei P_{n-1} die orthogonale Projektion auf W_{n-1} .

Dann folgt wieder aus $u_n \in W_{n-1}$, dass $u_n - P_{n-1} u_n \neq 0$ und

$$v_n = \frac{u_n - P_{n-1} u_n}{\|u_n - P_{n-1} u_n\|}$$

erfüllt $\|v_n\| = 1$ und $v_n \in W_{n-1}^\perp$. Wiederum folgt aus $v_1, \dots, v_n \in \text{Span}\{u_1, \dots, u_n\}$ und v_1, \dots, v_n ONS, dass $\text{Span}\{v_1, \dots, v_n\} = \text{Span}\{u_1, \dots, u_n\}$. Bei abzählbar vielen u_1, u_2 folgt die Existenz des behaupteten ONS per Induktion.

3.22 Korollar. Für jeden n -dimensionalen Skalarproduktraum V gilt:

- (a) V besitzt eine ONB.
- (b) Jedes ONS v_1, \dots, v_m lässt sich zu einer ONB (v_1, \dots, v_n) für V ergänzen.
- (c) Für $U = \text{Span}\{v_1, \dots, v_m\}$ gilt dann $U^\perp = \text{Span}\{v_{m+1}, \dots, v_n\}$ also $\dim U + \dim U^\perp = n$.

3.23 Bemerkung. Wir haben bereits gesehen, dass $V = \{e^{inx}\}$ eine ONS in $L^2([-\pi, \pi])$ bildet.

Die Norm in Skalarprodukträumen ist durch das Skalarprodukt gegeben, $\|u\| = \sqrt{\langle u, u \rangle}$. Man kann aber auch umgekehrt das Skalarprodukt aus der Norm rekonstruieren.

3.24 Satz. Polarisation und Parallelogrammgleichung

- (a) In jedem Skalarproduktraum über \mathbb{R} gilt

$$\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$$

- (b) In jedem Skalarproduktraum über \mathbb{C} gilt

$$\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2 + i\|u - iv\|^2 - i\|u + iv\|^2)$$

- (c) In jedem Skalarproduktraum gilt die **Parallelogrammgleichung**.

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Umgekehrt ist ein Vektorraum mit Norm $\|\cdot\|$ genau dann ein Skalarproduktraum, wenn die Parallelogrammgleichung gilt. Das zugehörige Skalarprodukt ist dann durch die Polarisationsgleichung gegeben.

Beweis. (a) und (b) rechnet man einfach nach, indem man auf der rechten Seite jeweils $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$ einsetzt und zusammenfasst. Ebenso zeigt man die Parallelogrammgleichung in Skalarprodukträumen. Es bleibt zu zeigen, dass durch die Polarisationsgleichung ein Skalarprodukt definiert wird, wann immer eine Norm die Parallelogrammgleichung erfüllt. Dazu prüft man direkt die Eigenschaften (a), (b) und (c) aus Definition 3.1. \square

3.25 Definition. Isometrie, orthogonale und unitäre Abbildung

- (a) Eine lineare Abbildung $L \in \mathcal{L}(V, W)$ zwischen normierten Räumen V und W heißt **Isometrie**, wenn $\|Lu\|_W = \|u\|_V$ für alle $u \in V$.
- (b) Eine surjektive Isometrie L zwischen Skalarprodukträumen heißt **unitär**, bzw. im Fall $\mathbb{K} = \mathbb{R}$ und $V = W$ auch **orthogonal**.

3.26 Lemma. Sei $L \in \mathcal{L}(V, W)$ eine Isometrie. Dann gilt

- (a) L ist injektiv.
- (b) Falls $\dim W = \dim V < \infty$, so ist L surjektiv.
- (c) Sind V und W Skalarprodukträume, so gilt auch $\langle Lu, Lv \rangle_W = \langle u, v \rangle_V \quad \forall u, v \in V$.

Für $W = V = \mathbb{R}^n$ mit dem natürlichen Skalarprodukt sind Isometrien also auch winkeltreu.

Beweis. (a) $Lu = 0 \Rightarrow \|u\|_V = \|Lu\|_W = 0 \Rightarrow u = 0$.

- (b) Für lineare Abbildungen zwischen endlichdimensionalen Räumen gilt nach Satz ??, dass L injektiv $\Leftrightarrow L$ surjektiv.
- (c) folgt aus Polarisierungsidentität. \square

3.27 Bemerkung. Für $\dim V = \infty$ gilt (b) im Allgemeinen nicht. Ein Beispiel ist der Rechtsshift R auf $\ell^2 := \{(a_n)_{n \in \mathbb{N}} \mid \sum_{n=1}^{\infty} |a_n|^2 < \infty\}$. Es ist

$$R : \ell^2 \rightarrow \ell^2, (a_1, a_2, \dots) \mapsto (0, a_1, a_2, \dots)$$

offensichtlich eine Isometrie, aber nicht surjektiv.

3.28 Lemma. Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine ONB von V . Dann ist die Koordinatenabbildung

$$\Phi_{\mathcal{B}} : V \rightarrow \mathbb{K}^n, u \mapsto (u)_{\mathcal{B}} = \begin{pmatrix} \langle v_1, u \rangle \\ \vdots \\ \langle v_n, u \rangle \end{pmatrix}$$

unitär, wobei \mathbb{K}^n mit dem natürlichen Skalarprodukt versehen ist.

Beweis. Parseval'sche Gleichung. \square

Also ist jeder endlichdimensionale Skalarproduktraum V über \mathbb{K} **unitär äquivalent** (\equiv **isometrisch isomorph**) zu \mathbb{K}^n mit dem natürlichen Skalarprodukt.

3.29 Bemerkung. In der Mathematik betrachtet man zwei Objekte vom gleichen Typ (z.B. Vektorräume, Skalarprodukträume, Gruppen usw.) als im wesentlichen identisch, wenn es strukturerhaltende Abbildungen zwischen ihnen gibt: für Vektorräume sind das die Isomorphismen (lineare Bijektionen), für Skalarprodukträume die isometrischen Isomorphismen (also die unitäre Abbildungen), für Gruppen die Gruppen-Isomorphismen (also Bijektionen, welche die Gruppenstruktur erhalten).

Wir betrachten im folgenden endlichdimensionale Skalarprodukträume. Da diese alle isometrisch isomorph zu \mathbb{K}^n sind, betrachten wir zunächst nur den \mathbb{K}^n . Man sollte aber im Hinterkopf behalten, dass alle Resultate und Definitionen sich direkt auf beliebige endlichdimensionale Skalarprodukträume mittels des isometrischen Isomorphismus aus Lemma 3.28 übertragen lassen.

Außerdem finden sich viele der folgenden Konzepte in sehr ähnlicher Form auch für ∞ -dimensionale Skalarprodukträume und insbesondere für Hilberträume wieder. Sie sind für die Physik von großer Bedeutung.

3.30 Definition. Die adjungierte Matrix

Die zu einer $m \times n$ -Matrix A adjungierte Matrix A^* ist $A^* = \overline{A^T}$, also $a_{ij}^* = \overline{a_{ji}}$. Für reelle Matrizen ist demnach $A^* = A^T$.

3.31 Lemma. Sei $A \in M(n \times m, \mathbb{K})$ dann gilt:

- (a) $\langle x, Ay \rangle_{\mathbb{K}^n} = \langle A^*x, y \rangle_{\mathbb{K}^m}$ für alle $x \in \mathbb{K}^n, y \in \mathbb{K}^m$
- (b) $\text{Rang } A^* = \text{Rang } A$.

Falls $A \in M(n \times n, \mathbb{K})$ dann gelten weiterhin

- (c) $\det A^* = \overline{\det A}$.
- (d) λ ist genau dann Eigenwert von A , wenn $\overline{\lambda}$ Eigenwert von A^* ist. Die geometrische und die algebraische Vielfachheit stimmen überein.

Beweis. (a)
$$\begin{aligned} \langle x, Ay \rangle_{\mathbb{K}^n} &= \sum_{j=1}^n \overline{x_j} (Ay)_j = \sum_{j=1}^n \overline{x_j} \sum_{k=1}^n a_{jk} y_k = \sum_{k=1}^m \sum_{j=1}^n a_{jk} \overline{x_j} y_k \\ &= \sum_{k=1}^m \sum_{j=1}^n \overline{a_{kj}^* x_j} y_k = \sum_{k=1}^m \overline{(A^*x)_k} y_k = \langle A^*x, y \rangle_{\mathbb{K}^m}. \end{aligned}$$

- (b) Da Spaltenrang = Zeilenrang gilt, folgt die Aussage daraus, dass Vektoren $a_1, \dots, a_\ell \in \mathbb{K}^n$ genau dann linear unabhängig sind, wenn $\overline{a_1}, \dots, \overline{a_\ell} \in \mathbb{K}^n$ linear unabhängig sind.
- (c) Wir wissen bereits, dass $\det A = \det A^T$. Aus der Leibnizformel ergibt sich $\det A^* = \overline{\det A^T}$.
- (d) Übungsaufgabe.

□

3.32 Lemma. (a) Sei $A \in M(m \times k, \mathbb{K})$ und $B \in M(k \times n, \mathbb{K})$, dann gilt

$$(AB)^* = B^*A^*.$$

- (b) Sei $A \in M(m \times n, \mathbb{K})$, dann ist $A^*A \in M(n \times n, \mathbb{K})$ und es gilt für $x \in \mathbb{K}^n$

$$\langle x, A^*Ax \rangle_{\mathbb{K}^n} = \|Ax\|_{\mathbb{K}^m}^2.$$

3.33 Bemerkung. Wir haben bisher nur die adjungierte Matrix aber nicht die adjungierte Abbildung definiert. Nun könnte man einfach mittels des Basisisomorphismus

$$\Phi_{\mathcal{B}} : V \rightarrow \mathbb{K}^n, \quad u \mapsto (u)_{\mathcal{B}}$$

die Definition der adjungierten Matrix auf lineare Abbildungen $L \in \mathcal{L}(V, W)$ hochziehen: Seien $\dim V = n$, $\dim W = m$ und \mathcal{A} und \mathcal{B} Orthonormalbasen, dann bildet $M_{\mathcal{A}}^{\mathcal{B}}(L)$ den \mathbb{K}^n auf \mathbb{K}^m ab und $M_{\mathcal{A}}^{\mathcal{B}}(L)^*$ den \mathbb{K}^m auf den \mathbb{K}^n . Also wäre

$$L^* := \Phi_{\mathcal{A}}^{-1} (M_{\mathcal{A}}^{\mathcal{B}}(L))^* \Phi_{\mathcal{B}}$$

eine lineare Abbildung von W nach V . Diese Definition ist zwar naheliegend und gewissermassen auch korrekt, man übersieht dabei aber einen wichtigen strukturellen Aspekt.

Exkurs: Der Dualraum**3.34 Definition. Dualraum**

Sei V ein endlichdimensionaler Vektorraum. Dann heißt der Vektorraum $\mathcal{L}(V, \mathbb{K})$, also der Raum der linearen Abbildungen von V nach \mathbb{K} , der **Dualraum** von V und wird mit V' bezeichnet.

3.35 Bemerkung. In unendlichdimensionalen Vektorräumen ist V' der Raum der stetigen linearen Abbildungen von V nach \mathbb{K} . In endlichdimensionalen Räumen sind aber alle linearen Abbildungen stetig.

3.36 Beispiel. Der Dualraum $(\mathbb{K}^n)'$ von \mathbb{K}^n (=Spaltenvektoren) ist der Raum der $1 \times n$ -Matrizen, $\mathcal{L}(\mathbb{K}^n, \mathbb{K}) = M(1 \times n, \mathbb{K})$, also der Raum der Zeilenvektoren: Für $x \in (\mathbb{K}^n)'$ und $y \in \mathbb{K}^n$ ist

$$x(y) = (x_1, \dots, x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n.$$

Es sind also \mathbb{K}^n und $(\mathbb{K}^n)'$ in natürlicher Weise isomorph.

3.37 Bemerkung. Für beliebige Skalarprodukträume V gibt es eine natürliche Identifizierung I zwischen V und V' : Zu $u \in V$ ist $I(u) \in V' = \mathcal{L}(V, \mathbb{K})$ gegeben durch

$$I(u) : V \rightarrow \mathbb{K}, v \mapsto \langle u, v \rangle.$$

Tatsächlich läßt sich jede lineare Abbildung $L : V \rightarrow \mathbb{K}$ in der Form $L(v) = \langle u, v \rangle$ für ein geeignetes $u \in V$ schreiben. Für \mathbb{K}^n ist das klar und mit Lemma 3.28 folgt das für beliebige endlichdimensionale Skalarprodukträume. Somit ist $I : V \rightarrow V'$ eine antilineare Bijektion. Für unendlichdimensionale Hilberträume gilt ein entsprechendes Resultat.

Wir formulieren das ganze noch etwas allgemeiner.

3.38 Satz. Sei also $H, \langle \cdot, \cdot \rangle$ ein Hilbertraum, und $L \in H'$. Dann gibt es genau ein $u \in H$, mit

$$L(v) = \langle u, v \rangle \quad \forall v \in H.$$

Beweis. Wir skizzieren den Beweis. Zuerst betrachten wir

$$M = \text{kern} L.$$

Es gilt $H = M \oplus M^\perp$. Nun ist es entscheidend zu erkennen dass $\dim M^\perp = 1$. Angenommen der Raum hätte zwei linear unabhängige Vektoren, $\phi, \psi \in M^\perp$ dann gelte

$$L(\lambda\phi + \mu\psi) = \lambda L(\phi) + \mu L(\psi) = 0$$

mit $\lambda = -\mu \frac{L(\psi)}{L(\phi)}$. Damit muss $\lambda\phi + \mu\psi = 0$ sein, und die beiden Vektoren können nicht linear unabhängig sein. Sei nun $\phi \in M^\perp$ mit $\|\phi\| = 1$. Dann gilt für $v \in H$, dass

$$v = \phi \langle \phi, v \rangle + z,$$

with $z \in M$. Also

$$L(v) = L(\phi) \langle \phi, v \rangle = \langle u, v \rangle,$$

with $u = L(\phi)\phi$. □

3.39 Bemerkung. Es ist nun relativ einfach zu zeigen, dass für $1 < p, q < \infty$ gilt, dass

$$(\ell^p)' = \ell^q.$$

3.40 Definition. Adjungierte Abbildung

Für $L \in \mathcal{L}(V, W)$ ist die adjungierte Abbildung L^\dagger definiert durch

$$L^\dagger : W' \rightarrow V', \underbrace{(L^\dagger w')}_{\in V'}(v) := w'(Lv) \quad \forall w' \in W', v \in V.$$

Diese Definition nimmt keinen Bezug auf ein Skalarprodukt. In Skalarprodukträumen kann man nun aber die Dualräume V' und W' mit den Räumen selbst identifizieren und bekommt somit eine Abbildung $L^* : W \rightarrow V$ gemäß

$$L^* : W \rightarrow V, \quad L^*w = I_V^{-1} L^\dagger I_W w.$$

Dann folgt die Eigenschaft

$$\langle w, Lv \rangle_W = \langle L^*w, v \rangle_V \quad \text{für alle } w \in W, v \in V,$$

direkt aus der Definition:

$$\langle w, Lv \rangle_W = \underbrace{Iw}_{\in W'}(\underbrace{Lv}_{\in W}) = \underbrace{(L^\dagger(Iw))}_{\in V'}(\underbrace{v}_{\in V}) = \langle I^{-1}L^\dagger Iw, v \rangle_V.$$

Sei nun H Hilbertraum und $L \in \mathcal{L}(H)$. Dann gilt

$$\langle L^*u, v \rangle = \langle u, Lv \rangle.$$

Wir merken uns also: Der adjungierte Operator zu $L \in \mathcal{L}(V, W)$ ist eigentlich eine Abbildung zwischen den Dualräumen W' und V' . Für Skalarprodukträume kann man aber die Dualräume mit den Räumen selbst identifizieren und somit die adjungierte Abbildung als Abbildung von W nach V auffassen.

3.41 Bemerkung. In der Physik nennt man die Elemente eines Vektorraums V oft die kontravarianten Vektoren und die Elemente des Dualraums V' die kovarianten Vektoren. Hat man ein Skalarprodukt auf V , so kann man ko- und kontravariante Vektoren ineinander umwandeln. Die Bezeichnungen gehen auf das unterschiedliche Transformationsverhalten unter Basiswechseln zurück, siehe Übungen. Angenommen wir haben Vektorraum mit Basis X_1, X_2 und innerem Produkt $g(Z, Y)$, mit $g_{ij} = g(X_i, X_j)$. Dann sind für einen Vektor $Y = y^i X_i$, die y^i die kontravarianten Vektoren und $y_i = g_{i,j} y^j$ die kovarianten Vektoren. $g(Y, \cdot)$ ist lineares Funktional, $L_Y(\cdot)$. Diesem Funktional entspricht der kovariante Vektor $g_{i,j} y^j$.

3.42 Satz. Eine Matrix $A \in M(n \times n, \mathbb{K})$ definiert genau dann eine unitäre Abbildung $A : \mathbb{K}^n \rightarrow \mathbb{K}^n$, wenn $A^* = A^{-1}$ gilt.

Beweis. Mit Lemma 3.26 ist A genau dann unitär, wenn A isometrisch ist. Sei also $A^* = A^{-1}$ und $x \in \mathbb{K}^n$, dann gilt

$$\|Ax\|^2 = \langle Ax, Ax \rangle = \langle A^*Ax, x \rangle = \langle A^{-1}Ax, x \rangle = \langle x, x \rangle = \|x\|^2.$$

Also ist A eine Isometrie. Sei umgekehrt A eine Isometrie, dann ist für jedes $x \in \mathbb{K}^n$

$$\langle A^*Ax, y \rangle = \langle Ax, Ay \rangle = \langle x, y \rangle$$

für alle $y \in \mathbb{K}^n$ und somit $A^*Ax = x$, also $A^* = A^{-1}$. □

3.43 Korollar. Seien V und W endlichdimensionale Skalarprodukträume und $L \in \mathcal{L}(V, W)$. Es ist L genau dann unitär, wenn $L^* = L^{-1}$ gilt.

3.44 Bemerkung. Eine $n \times n$ -Matrix ist genau dann unitär bzw. orthogonal, wenn ihre Spalten (oder analog ihre Zeilen) eine ONB des \mathbb{K}^n bilden, denn für die Spalten von $A = (a_1, \dots, a_n)$ gilt dann

$$\langle a_j, a_i \rangle = (A^*A)_{ji} = (E_n)_{ji} = \delta_{ji}.$$

Die Transformationsmatrix S eines Basiswechsels zwischen ONBen des \mathbb{K}^n ist also unitär, da die Spalten als Bilder der Basisvektoren wieder eine ONB bilden. Umgekehrt kann man jede unitäre Matrix als Transformationsmatrix eines Basiswechsels zwischen ONBen interpretieren.

3.45 Satz. Sei S eine unitäre $n \times n$ -Matrix.

- (a) Jeder Eigenwert λ von S liegt auf dem Einheitskreis, d.h. $|\lambda| = 1$.
- (b) $|\det S| = 1$

Beweis. (a) Sei $\lambda \in \mathbb{C}$ Eigenwert von S und u ein zugehöriger Eigenvektor. Dann gilt

$$\|u\| = \|Su\| = \|\lambda u\| = |\lambda| \|u\|.$$

Da $u \neq 0$, muss $|\lambda| = 1$ gelten.

(b) $1 = \det E_n = \det S^* S = \det S^* \det S = \overline{\det S} \det S = |\det S|^2$. □

3.46 Beispiel. Die unitären Abbildungen auf dem \mathbb{R}^2 , also in diesem Fall die orthogonalen, sind die Spiegelungen und Drehungen. Nur diese lassen alle Winkel und Längen intakt. Formal folgt das so: da die Spalten einer orthogonalen Matrix eine ONB bilden, kann man für die erste Spalte

$$s_1 = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \quad \text{für ein } \varphi \in [0, 2\pi)$$

ansetzen. Es gibt nun genau zwei zu s_1 orthogonale Vektoren der Länge eins, nämlich

$$s_2 = \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \quad \text{und} \quad s_3 = \begin{pmatrix} \sin \varphi \\ -\cos \varphi \end{pmatrix}.$$

Es ist $D_\varphi = (s_1, s_2)$ die Drehung um den Winkel φ und $S_{\varphi/2} = (s_1, s_3)$ die Spiegelung an der Ursprungsgeraden mit Winkel $\varphi/2$ zur x -Achse.

Es gilt $\det D_\varphi = 1$, da Drehungen die Orientierung einer Basis nicht ändern und $\det S_{\varphi/2} = -1$, da Spiegelungen die Orientierung einer Basis ändern.

Exkurs: Gruppen

3.47 Definition. Gruppe

Eine Gruppe ist ein Paar (G, \circ) bestehend aus einer Menge G und einer Abbildung

$$\begin{aligned} \circ : G \times G &\rightarrow G \\ (a, b) &\mapsto a \circ b \end{aligned}$$

mit den folgenden Eigenschaften:

- (a) $(a \circ b) \circ c = a \circ (b \circ c)$ für alle $a, b, c \in G$ (Assoziativität)
- (b) Es gibt ein $e \in G$ mit $a \circ e = e \circ a = a$ für alle $a \in G$ (Existenz des neutralen Elements).
- (c) Zu jedem $a \in G$ gibt es ein $a^{-1} \in G$ mit $a \circ a^{-1} = a^{-1} \circ a = e$. (Existenz des Inversen).

Eine Teilmenge $H \subset G$, welche selbst wieder eine Gruppe ist, heißt Untergruppe.

3.48 Bemerkung. Es genügt in (b) und (c) jeweils nur $e \circ a = a$ und $a^{-1} \circ a = e$ zu fordern. Die jeweils andere Gleichung folgt dann:

- Jedes linksinverse Element ist auch rechtsinvers: sei $a^{-1} \circ a = e$, dann ist aufgrund der Assoziativität

$$\begin{aligned} a \circ a^{-1} &= e \circ (a \circ a^{-1}) = ((a^{-1})^{-1} \circ a^{-1}) \circ (a \circ a^{-1}) = \\ &= (a^{-1})^{-1} \circ (a^{-1} \circ a) \circ a^{-1} = (a^{-1})^{-1} \circ a^{-1} = e. \end{aligned}$$

- Jedes linksneutrale Element ist auch rechtsneutral: sei $e \circ a = a$, dann ist

$$a \circ e = a \circ (a^{-1} \circ a) = (a \circ a^{-1}) \circ a = a.$$

3.49 Beispiel. Matrixgruppen:

Die Verknüpfung ist im Folgenden durch das Matrixprodukt gegeben, das neutrale Element ist jeweils die Einheitsmatrix.

- (a) Alle invertierbaren $n \times n$ -Matrizen bilden die lineare Gruppe $GL(n, \mathbb{K})$:
für $A, B \in GL(n, \mathbb{K})$ ist auch AB wieder invertierbar, da $(AB)^{-1} = B^{-1}A^{-1}$.

- (b) Die unitären komplexen Matrizen bilden die unitäre Gruppe $U(n)$:
für $A, B \in U(n)$ ist auch AB wieder unitär, da $(AB)^* = B^*A^* = B^{-1}A^{-1} = (AB)^{-1}$.
 $U(n)$ ist Untergruppe von $GL(n)$.
- (c) Die orthogonalen $n \times n$ -Matrizen bilden die orthogonale Gruppe $O(n)$. $O(n)$ ist Untergruppe von $U(n)$.
- (d) Die spezielle lineare Gruppe ist $SL(n, \mathbb{K}) = \{A \in GL(n, \mathbb{K}) \mid \det A = 1\}$. Wegen $\det AB = \det A \cdot \det B$ ist $SL(n, \mathbb{K})$ wieder eine Gruppe.
- (e) Die spezielle unitäre Gruppe $SU(n) = \{A \in U(n) \mid \det A = 1\}$.
- (f) Die spezielle orthogonale Gruppe $SO(n) = \{A \in O(n) \mid \det A = 1\}$ besteht genau aus den Drehmatrizen. Analog zu (a)–(f) definiert man für Vektorräume V die Gruppen $GL(V)$, $SL(V)$ und für Skalarprodukträume zusätzlich $U(V)$ und $SU(V)$.
- (g) Sei Allgemein M eine Menge und $\text{Bij}(M)$ die Menge der bijektiven Abbildung $f : M \rightarrow M$. Dann ist $(\text{Bij}(M), \circ)$ eine Gruppe. Hier steht \circ für die Komposition von Abbildungen und das neutrale Element ist die Identität.

3.50 Definition. Abelsche Gruppe

Ist eine Gruppe (G, \circ) kommutativ, d.h.

$$a \circ b = b \circ a \quad \text{für alle } a, b \in G,$$

dann heißt G eine **abelsche** Gruppe.

3.51 Beispiel. • $(\mathbb{Z}, +)$ ist eine abelsche Gruppe

- $(\mathbb{R}, +)$ ist eine abelsche Gruppe
- Ist $(V, +, \cdot)$ ein Vektorraum über \mathbb{K} , so ist $(V, +)$ eine abelsche Gruppe
- $SO(2)$ ist eine abelsche Gruppe
- $SO(3)$ ist nicht abelsch

3.52 Beispiel. Transformation des Raumes und der Raumzeit

- Die **Bewegungsgruppe** B ist die Gruppe der längentreuen Abbildungen $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit $\|f(x) - f(y)\| = \|x - y\|$ für alle $x, y \in \mathbb{R}^3$. Man kann zeigen, dass sich jedes $f \in B$ schreiben läßt als Verknüpfung einer Translation und einer Drehung, also für $f \in B$ gibt es $a \in \mathbb{R}^3$ und $D \in SO(3)$ so, dass $f(x) = Dx + a$.
- **Die Galilei-Gruppe** (nichtrelativistische Raumzeitsymmetrie)

Die nichtrelativistischen Gesetze der Physik (Newton'sche Mechanik, Quantenmechanik) sind invariant unter

- Drehungen und Translationen des Raumes

$$x \mapsto x' = Dx + a, \quad D \in SO(3), \quad a \in \mathbb{R}^3.$$

- Verschiebung des Zeitnullpunktes

$$t \mapsto t' = t + \tau, \quad \tau \in \mathbb{R}.$$

- gleichförmige Bewegung

$$x \mapsto x' = x + tv, \quad v \in \mathbb{R}^3. \quad (\text{Galilei-Boosts})$$

Zusammen bilden diese Transformationen die Galilei-Gruppe. Fassen wir Zeit und Ort zu einem Vektor

$\begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R}^4$ (=Raumzeit) zusammen, so sind alle Galilei-Transformationen von der Form

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto T_{(D,a,\tau,v)} \begin{pmatrix} t \\ x \end{pmatrix} := \begin{pmatrix} t + \tau \\ Dx + a + tv \end{pmatrix}.$$

Zusammen bilden die Abbildungen $T_{(D,a,\tau,v)} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ die inhomogene Galilei-Gruppe. Die Gruppeneigenschaft rechnet man direkt nach:

$$\begin{aligned} T_{(D_2,a_2,\tau_2,v_2)} \circ T_{(D_1,a_1,\tau_1,v_1)} \begin{pmatrix} t \\ x \end{pmatrix} &= T_{(D_2,a_2,\tau_2,v_2)} \begin{pmatrix} t + \tau_1 \\ D_1x + a_1 + tv_1 \end{pmatrix} = \\ &= \begin{pmatrix} t + \tau_1 + \tau_2 \\ D_2(D_1x + a_1 + tv_1) + a_2 + (t + \tau_1)v_2 \end{pmatrix} \\ &= T_{(D_2D_1, D_2a_1+a_2+\tau_1v_2, \tau_1+\tau_2, D_2v_1+v_2)} \begin{pmatrix} t \\ x \end{pmatrix}. \end{aligned}$$

Die Galilei Transformationen erhalten die Gleichzeitigkeit. Für $\begin{pmatrix} t \\ x \end{pmatrix}$ und $\begin{pmatrix} s \\ y \end{pmatrix}$ gilt: Falls $t = s$, dann stimmen auch nach jeder Galilei Transformation die ersten Komponenten wieder überein. Setzt man $a = 0$ und $\tau = 0$, so bilden die $T_{(D,v)}$ die homogene Galileigruppe, eine Untergruppe von $GL(\mathbb{R}^4)$:

$$T_{(D,v)} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ v & D \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} t \\ Dx + vt \end{pmatrix}.$$

Was bedeutet es nun, dass ein physikalisches Gesetz oder eine Theorie Galilei-invariant ist? Dazu sollten die Lösungen durch Objekte auf der Raumzeit \mathbb{R}^4 gegeben sein, z. B. Weltlinien von Teilchen, d.h. Kurven

$$\gamma_j : \mathbb{R} \rightarrow \mathbb{R}^4, \quad \xi \mapsto \gamma_j(\xi) = \begin{pmatrix} t_j(\xi) \\ x_j(\xi) \end{pmatrix}, \quad j = 1, \dots, N,$$

wobei γ_j die Weltlinie des j -ten Teilchens ist und normalerweise angenommen wird, dass $t_j : \mathbb{R} \rightarrow \mathbb{R}$, $\xi \mapsto t_j(\xi)$ eine Bijektion ist. Eine Theorie deren Lösungen solche Weltlinien sind, z.B. Newton'sche Mechanik, heißt nun Galilei-invariant, wenn mit jeder Lösung $(\gamma_j)_{j=1, \dots, N}$ und für jede Galileitransformation $T_{(D,a,\tau,v)}$ auch $(T_{(D,a,\tau,v)} \circ \gamma_j)_{j=1, \dots, N}$ eine Lösung ist. Die Parametrisierung der Kurven γ_j spielt hier übrigens keine Rolle.

• **Die Lorentz-Gruppe** (relativistische Raumzeitsymmetrie)

Ersetzt man die Galilei-Boosts $T_{(1,0,0,v)}$ durch Lorentz-Boosts,

$$\begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} t' \\ x' \end{pmatrix} = L_v \begin{pmatrix} t \\ x \end{pmatrix} := \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} t + \frac{v \cdot x}{c^2} \\ x + vt \end{pmatrix} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} 1 & \frac{v^T}{c^2} \\ v & \mathbb{1} \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}$$

erhält man die Lorentzgruppe (oder auch Poincarégruppe). Die Lorentztransformationen erhalten die Gleichzeitigkeit nicht: für $\begin{pmatrix} t \\ x \end{pmatrix}$ und $\begin{pmatrix} s \\ y \end{pmatrix}$ mit $t = s$ stimmen die ersten Komponenten von $L_v \begin{pmatrix} t \\ x \end{pmatrix}$ und $L_v \begin{pmatrix} t \\ y \end{pmatrix}$ im Allgemeinen nicht überein. Andererseits ist L_v auf den letzten drei Komponenten auch nicht mehr längentreu:

$$L_v \begin{pmatrix} t \\ x \end{pmatrix} - L_v \begin{pmatrix} t \\ y \end{pmatrix} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} \frac{v \cdot (x-y)}{c^2} \\ (x-y) \end{pmatrix}.$$

Was bedeutet das für die Lösungen Lorentz-invarianter Theorien? Nehmen wir an, die Endpunkte eines ruhenden Stabes seien durch die Weltlinien

$$\gamma_1(\xi) = \begin{pmatrix} \xi \\ 0 \end{pmatrix} \quad \text{und} \quad \gamma_2(\zeta) = \begin{pmatrix} \zeta \\ 1 \end{pmatrix}$$

gegeben. (Der einfacheren Notation wegen nehmen wir einen eindimensionalen Raum also eine zweidimensionale Raumzeit an). In einer Lorentz-invarianten Theorie ist dann auch

$$\gamma'_1(\xi) = (L_v \circ \gamma_1)(\xi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} \xi \\ v \xi \end{pmatrix}, \quad \gamma'_2(\zeta) = (L_v \circ \gamma_2)(\zeta) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \begin{pmatrix} \zeta + \frac{v}{c^2} \\ 1 + v \zeta \end{pmatrix}$$

eine Lösung. Wie lang ist nun aber dieser bewegte Stab? Die Zeitkomponenten der Lösungen stimmen überein, wenn wir $\xi = \zeta + \frac{v}{c^2}$ wählen. Die Differenz der Ortskomponenten ist dann

$$\left\| \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(v \left(\zeta + \frac{v}{c^2} \right) - 1 - v \zeta \right) \right\| = \sqrt{1 - \frac{v^2}{c^2}},$$

also um den Faktor $\sqrt{1 - \frac{v^2}{c^2}}$ kleiner als die ursprüngliche Differenz. Das ist die sogenannte Längen- oder Lorentzkontraktion. Bewegte Körper erscheinen uns (in unserem ruhenden Koordinatensystem) in der Bewegungsrichtung verkürzt. Natürlich könnten wir die Lorentztransformation auch als einen Koordinatenwechsel interpretieren, d.h. $\gamma'_1(\xi)$ und $\gamma'_2(\zeta)$ wären die Koordinaten des Stabes in einem gleichförmig mit der Geschwindigkeit $-v$ bewegten Koordinatensystem. Jemand der dieses Koordinatensystem verwendet (man spricht dann von einem bewegten Beobachter), würde dann ebenfalls die Länge des Stabes mit $\sqrt{1 - \frac{v^2}{c^2}}$ angeben.

3.2 Symmetrische Operatoren

3.53 Definition. Symmetrie und Selbstadjungiertheit

- (a) Eine $n \times n$ -Matrix A heißt **symmetrisch** oder **selbstadjungiert**, wenn $A = A^*$ gilt.
- (b) Eine lineare Abbildung $T \in \mathcal{L}(V)$ auf einem Skalarproduktraum V heißt **symmetrisch** oder **selbstadjungiert**, wenn

$$\langle Tu, v \rangle = \langle u, Tv \rangle \quad \text{für alle } u, v \in V.$$

3.54 Bemerkung. Zumindest auf einem endlichdimensionalen Raum ist also T symmetrisch genau dann wenn $T = T^*$ gilt. (Vgl. Lemma 3.31 und Exkurs Dualraum). Dies ist wiederum genau dann der Fall, wenn die Matrix $M_{\mathcal{B}}(T)$ bzgl. einer (und damit bzgl. jeder) ONB symmetrisch ist.

3.55 Satz. Seien S und T Endomorphismen eines Skalarproduktraumes V . Dann gilt

- (a) Sind S und T selbstadjungiert, so auch $\alpha S + \beta T$ für $\alpha, \beta \in \mathbb{R}$.
- (b) Ist T selbstadjungiert und invertierbar, so ist auch T^{-1} selbstadjungiert.
- (c) Ist V ein komplexer Skalarproduktraum, so ist T genau dann selbstadjungiert, wenn

$$\langle u, Tu \rangle \in \mathbb{R} \quad \text{für alle } u \in V.$$

Beweis. (a) Klar, Sesquilinearität des Skalarprodukts.

(b) Seien $u, v \in V$, dann ist

$$\langle T^{-1}u, v \rangle = \langle T^{-1}u, TT^{-1}v \rangle = \langle TT^{-1}u, T^{-1}v \rangle = \langle u, T^{-1}v \rangle.$$

(c) Übungsaufgabe. □

Wir werden zeigen, dass selbstadjungierte Operatoren auf endlichdimensionalen Skalarprodukträumen immer diagonalisierbar sind. Das folgt im Wesentlichen schon aus den folgenden sehr simplen Beobachtungen:

3.56 Satz. Sei V ein Skalarproduktraum und $T \in \mathcal{L}(V)$ selbstadjungiert. Dann gilt:

- (a) Alle Eigenwerte von T sind reell.
- (b) Eigenvektoren von T zu verschiedenen Eigenwerten sind orthogonal.
- (c) Ist u ein Eigenvektor von T , so ist der zu u orthogonale Teilraum

$$W = \{u\}^\perp = \{v \in V \mid \langle u, v \rangle = 0\}$$

T -invariant, d.h. $T(W) \subset W$ (also $Tw \in W \forall w \in W$).

Beweis. (a) Sei $Tu = \lambda u$ mit $u \neq 0$, dann ist

$$\lambda \langle u, u \rangle = \langle u, \lambda u \rangle = \langle u, Tu \rangle = \langle Tu, u \rangle = \langle \lambda u, u \rangle = \bar{\lambda} \langle u, u \rangle,$$

also folgt mit $\langle u, u \rangle \neq 0$, dass $\lambda = \bar{\lambda}$.

(b) Sei $Tu = \lambda u$ und $Tv = \mu v$ für $\lambda \neq \mu$ und $u, v \neq 0$, dann ist

$$\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, \mu v \rangle = \mu \langle u, v \rangle,$$

also folgt mit $\lambda \neq \mu$, dass $\langle u, v \rangle = 0$.

(c) Sei $Tu = \lambda u$ und $w \in W$, also $\langle w, u \rangle = 0$. Dann ist auch

$$\langle Tw, u \rangle = \langle w, Tu \rangle = \langle w, \lambda u \rangle = \lambda \langle w, u \rangle = 0,$$

also $Tw \in W$. □

3.57 Satz. Jeder selbstadjungierte Endomorphismus T auf einem endlichdimensionalen Skalarproduktraum über \mathbb{R} oder \mathbb{C} besitzt wenigstens einen Eigenwert (welcher dann gemäß Satz 3.56 reell ist.)

Beweis. (a) Im Fall $\mathbb{K} = \mathbb{C}$ besitzt das charakteristische Polynom P_T von T nach dem Fundamentalsatz der Algebra eine Nullstelle $\lambda \in \mathbb{C}$. Diese ist ein Eigenwert von T .

(b) Auch im Fall $\mathbb{K} = \mathbb{R}$ hat das charakteristische Polynom P_T von T sicherlich eine Nullstelle $\lambda \in \mathbb{C}$. Diese ist Eigenwert von $A = M_{\mathcal{B}}(T)$ als Element von $\mathcal{L}(\mathbb{C}^n)$, wobei \mathcal{B} eine ONB ist. Nach Satz 3.56 ist aber λ reell und somit auch Eigenwert von T . □

3.58 Satz. Diagonalisierbarkeit selbstadjungierter Endomorphismen

Jeder selbstadjungierte Endomorphismus auf einem endlichdimensionalen Skalarproduktraum ist diagonalisierbar, besitzt also eine Orthonormalbasis aus Eigenvektoren.

Beweis. Induktion über $\dim V = n$:

Im Fall $\dim V = 1$ ist jede lineare Abbildung $T : V \rightarrow V$ diagonalisierbar, denn ist $V = \text{Span}\{v\}$ mit $v \neq 0$, so gibt es wegen $Tv \in V$ ein $\lambda \in \mathbb{K}$ mit $Tv = \lambda v$.

Sei also die Behauptung für alle Skalarprodukträume der Dimension n richtig, V ein $(n+1)$ -dimensionaler Skalarproduktraum und $T \in \mathcal{L}(V)$ selbstadjungiert.

Gemäß Satz 3.57 hat T einen reellen Eigenwert λ_1 . Sei $Tv_1 = \lambda_1 v_1$ mit $\|v_1\| = 1$ und $W_1 = \{v_1\}^\perp$.

Nach Korollar 3.22 ist $\dim W_1 = \dim V - 1 = n$. Gemäß Satz 3.56 ist $T_1 = T|_{W_1}$ eine selbstadjungierte Abbildung auf dem Skalarproduktraum W_1 . Nach Induktionsannahme gibt es eine ONB v_2, \dots, v_{n+1} von W_1 und Zahlen $\lambda_2, \dots, \lambda_{n+1} \in \mathbb{K}$ mit

$$T_1 v_k = T v_k = \lambda_k v_k \quad \text{für } k = 2, \dots, n+1.$$

Also bilden (v_1, \dots, v_{n+1}) eine ONB für V mit

$$T v_k = \lambda_k v_k \quad \text{für } k = 1, \dots, n+1. \quad \square$$

3.59 Korollar. Hauptachsentransformation selbstadjungierter Operatoren

Zu einem selbstadjungierten Endomorphismus $T : V \rightarrow V$ eines n -dimensionalen Skalarproduktraumes V über \mathbb{K} läßt sich stets eine unitäre Abbildung

$$S : \mathbb{K}^n \rightarrow V$$

finden, welche T diagonalisiert, also

$$S^{-1}TS = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \lambda_m & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_m \end{pmatrix}$$

erfüllt. Hier sind $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte von T , wobei jeder Eigenwert entsprechend seiner geometrischen Vielfachheit oft wiederholt wird.

3.60 Korollar. Spektraldarstellung selbstadjungierter Operatoren

Ist $T : V \rightarrow V$ selbstadjungiert auf dem endlichdimensionalen Skalarproduktraum V , dann gilt

$$T = \sum_{k=1}^m \lambda_k P_k,$$

wobei $\lambda_1, \dots, \lambda_m$ die verschiedenen Eigenwerte sind und $P_k : V \rightarrow V$ jeweils die orthogonale Projektion auf den Eigenraum E_{λ_k} bezeichnet.

Beweis. Es reicht aus zu zeigen, dass beide Seiten auf Eigenvektoren von T dieselbe Wirkung haben, denn es gibt ja eine Basis aus Eigenvektoren. Sei also $Tv = \lambda_j v$, dann gilt

$$P_k v = \delta_{kj} v,$$

da Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind. Also ist

$$\sum_{k=1}^m \lambda_k P_k v = \lambda_j v.$$

□

3.61 Bemerkung. Diese Spektraldarstellung läßt sich auf geeignete selbstadjungierte Operatoren auf ∞ -dimensionalen Räumen verallgemeinern.

Zusammenfassend geben wir an dieser Stelle noch das Rezept zur Hauptachsentransformation von selbstadjungierten Matrizen an (bzw. Hauptachsentransformation von Operatoren: dann bestimme im nullten Schritt die Matrix bzgl. irgendeiner ONB).

Rezept zur Hauptachsentransformation von selbstadjungierten Matrizen

Sei $A \in M(n \times n, \mathbb{K})$ selbstadjungiert. Zur Bestimmung einer Hauptachsentransformation $S \in U(n)$ für A führe man die folgenden Schritte durch.

1. Schritt. Man bestimme die verschiedenen Eigenwerte $\lambda_1, \dots, \lambda_m$ von A als die verschiedenen Nullstellen des charakteristischen Polynoms

$$P_A(\lambda) = \det(A - \lambda E_n).$$

2. Schritt. Für jedes $k = 1, \dots, m$ bestimme man eine Basis $(w_1^{(k)}, \dots, w_{n_k}^{(k)})$ des Eigenraums E_{λ_k} , indem man das Gauß'sche Verfahren zur Lösung des Gleichungssystems $(A - \lambda_k E_n)x = 0$ anwendet.

3. Schritt. Man überführe die Basis $w_1^{(k)}, \dots, w_{n_k}^{(k)}$ mittels des Gram-Schmidt'schen Verfahrens in eine Orthonormalbasis $(v_1^{(k)}, \dots, v_{n_k}^{(k)})$.

4. Schritt. Durch Aneinanderreihung dieser Basen entsteht dann die ONB

$$(v_1, \dots, v_n) = (v_1^{(1)}, \dots, v_{n_1}^{(1)}, \dots, v_1^{(m)}, \dots, v_{n_m}^{(m)})$$

von V aus Eigenvektoren und die Transformationsmatrix S enthält die Basisvektoren (v_1, \dots, v_n) als Spalten.

3.62 Definition. Quadratische Form

Eine symmetrische Sesquilinearform auf einem \mathbb{K} -Vektorraum heißt quadratische Form. D.h.

$$Q : V \times V \rightarrow \mathbb{K}$$

ist eine quadratische Form, falls

- (a) $Q(u, \alpha v + w) = \alpha Q(u, v) + Q(u, w)$ für alle $u, v, w \in V$ und $\alpha \in \mathbb{K}$.
- (b) $Q(u, v) = \overline{Q(v, u)}$ für alle $u, v \in V$.

Für eine quadratische Form Q gilt zwar aufgrund der Symmetrie $Q(v, v) \in \mathbb{R}$, aber nicht wie beim Skalarprodukt $Q(v, v) \geq 0$ bzw. $Q(v, v) = 0 \Leftrightarrow v = 0$.

3.63 Satz. Sei V ein Skalarproduktraum.

- (a) Zu jedem symmetrischen Operator T auf V wird durch

$$Q_T(u, v) := \langle u, Tv \rangle$$

eine quadratische Form definiert.

- (b) Falls $\dim V < \infty$, dann gibt es zu jeder quadratischen Form Q auf V einen eindeutigen symmetrischen Operator $T \in \mathcal{L}(V)$ mit

$$Q(u, v) = \langle u, Tv \rangle \quad \forall u, v \in V.$$

Beweis. (a) Linearität von Q_T im zweiten Argument ist offensichtlich und die Symmetrie ebenfalls:

$$Q_T(u, v) = \langle Tu, v \rangle = \overline{\langle v, Tu \rangle} = \overline{Q_T(u, v)}.$$

- (b) Sei eine ONB $\mathcal{A} = (v_1, \dots, v_n)$ von V gegeben. Dann hat der Operator definiert durch

$$Tv_k = \sum_{j=1}^n Q(v_j, v_k)v_j =: \sum_{j=1}^n a_{jk}v_j$$

die gewünschten Eigenschaften: T ist symmetrisch, denn die Matrix $M_{\mathcal{A}}(T) = (a_{jk})$ ist selbstadjungiert,

$$a_{jk} = Q(v_j, v_k) = \overline{Q(v_k, v_j)} = \overline{a_{kj}},$$

und es gilt

$$\langle v_m, Tv_k \rangle = Q(v_m, v_k)$$

nach Definition. Wegen Sesquilinearität auf beiden Seiten folgt

$$\langle v, Tu \rangle = Q(v, u) \quad \forall u, v \in V.$$

Die Eindeutigkeit folgt wieder aus der Definitheit des Skalarprodukts:

$$\langle u, Tv \rangle = \langle u, \tilde{T}v \rangle \quad \forall u \in V \quad \Rightarrow \quad Tv = \tilde{T}v.$$

□

3.64 Satz. und Definition. Hauptachsentransformation quadratischer Formen

Sei V ein n -dimensionaler Skalarproduktraum und Q eine quadratische Form, also

$$Q(u, v) = \langle u, Tv \rangle$$

mit einem symmetrischen Operator T . Nach Satz 3.57 gibt es eine ONB $\mathcal{A} = (v_1, \dots, v_n)$ aus Eigenvektoren von T :

$$Tv_1 = \lambda_1 v_1, \dots, Tv_n = \lambda_n v_n.$$

Dann hat Q bezüglich \mathcal{A} die Hauptachsendarstellung

$$Q(u) := Q(u, u) = \sum_{j=1}^n \lambda_j |\langle v_j, u \rangle|^2.$$

Die Geraden $\text{Span}\{v_j\}$ heißen die Hauptachsen von Q .

3.65 Korollar. Rayleigh-Ritz-Prinzip

Seien die Eigenwerte $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ des symmetrischen Operators T der Größe nach geordnet, so gilt

$$\lambda_1 = \min\{Q_T(u) \mid \|u\| = 1\} \quad \text{und} \quad \lambda_n = \max\{Q_T(u) \mid \|u\| = 1\}.$$

Beweis. Übungsaufgabe. □

3.66 Definition. Positive Operatoren und Matrizen

Ein symmetrischer Operator T heißt

- **positiv** und wir schreiben $T \geq 0$, falls $\langle u, Tu \rangle \geq 0 \quad \forall u \in V$.
- **positiv definit** und wir schreiben $T > 0$, falls $\langle u, Tu \rangle > 0 \quad \forall u \neq 0$.

Entsprechend definiert man positiv (definite) quadratische Formen und Matrizen, bzw. negative und negativ definite Operatoren.

3.67 Satz. In einem endlichdimensionalen Skalarproduktraum gilt

- $T \geq 0 \Leftrightarrow$ alle Eigenwerte von T sind nichtnegativ.
- $T > 0 \Leftrightarrow$ alle Eigenwerte von T sind positiv.

Beweis. Folgt sofort aus der Darstellung

$$\langle u, Tu \rangle = \sum_{k=1}^n \lambda_k |\langle v_k, u \rangle|^2$$

aus Satz 3.64. □

3.68 Beispiel. Die Matrix der Gaußschen Normalengleichung (vgl. Ende Kapitel 3)

Zu lösen ist

$$Ax = y$$

mit $A \in M(m \times n, \mathbb{R})$ und $y \in \mathbb{R}^m$ bekannt und $x \in \mathbb{R}^n$ gesucht, wobei $\text{Rang} A = n < m$. Im Allgemeinen ist $y \notin \text{Rang} A$ und somit hat die Gleichung keine Lösung. Deshalb sucht man die beste Approximation, d.h. ein u das $\|Au - y\|$ minimiert. Am Ende von Kapitel 3 haben wir gezeigt, dass ein solches U die Gleichung $A^T A u = A^T y$ lösen muss. D.h., $B = A^T A$ ist zu invertieren. Es ist $B \in M(n \times n, \mathbb{R})$ positiv definit, denn

$$\langle x, Bx \rangle_{\mathbb{R}^n} = \langle x, A^T A x \rangle_{\mathbb{R}^n} = \langle x, A^* A x \rangle_{\mathbb{R}^n} = \langle Ax, Ax \rangle_{\mathbb{R}^m} = \|Ax\|_{\mathbb{R}^m}^2 \geq 0.$$

Da A vollen Rang hat, also $\text{Rang} A = n$, ist $\text{Kern} A = \{0\}$ und somit

$$\langle x, Bx \rangle_{\mathbb{R}^n} = \|Ax\|_{\mathbb{R}^m}^2 > 0 \quad \forall x \neq 0.$$

Nun sind aber positiv definite Matrizen insbesondere injektiv und somit invertierbar. Also haben die Gaußschen Normalengleichungen immer eine eindeutige Lösung. Dieselbe Lösung erhält man, wenn man $Au = P_{\text{Bild}A} y$ löst, wobei $P_{\text{Bild}A}$ die orthogonale Projektion auf das Bild von A ist.

3.69 Bemerkung. Quadratische Formen und Kegelschnitte

Die Lösungsmenge der Gleichung

$$\langle x, Ax \rangle + \langle b, x \rangle = c \quad x, b \in \mathbb{R}^2, c \in \mathbb{R}, A \in M(2 \times 2, \mathbb{R})$$

beschreibt einen Kegelschnitt, denn nach Hauptachsentransformation $S \in O(2)$ von A (Drehung des Koordinatensystems) gilt mit

$$S^{-1} A S = D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad y = S^{-1} x \quad \text{und} \quad \tilde{b} = S^{-1} b,$$

dass

$$\langle x, Ax \rangle + \langle b, x \rangle = c \Leftrightarrow \langle y, D y \rangle + \langle \tilde{b}, y \rangle = c = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \tilde{b}_1 y_1 + \tilde{b}_2 y_2.$$

Als Lösungsmenge hat man beispielsweise für $c > 0$ und

- (a) $A > 0$ und $b = 0$, also $\lambda_1, \lambda_2 > 0$, eine Ellipse mit Hauptachsen $\lambda_1^{-1/2}Se_1$ und $\lambda_2^{-1/2}Se_2$,
- (b) $\lambda_1 \cdot \lambda_2 < 0$ und $b = 0$ eine Hyperbel,
- (c) $\lambda_1 > 0, \lambda_2 = 0$ und $b \neq 0$ eine Parabel, bzw. für $b = 0$ ein Geradenpaar,
- (d) $A < 0$ und $b = 0$ die leere Menge.

3.70 Definition. Funktionalkalkül für selbstadjungierte Operatoren

Sei $T \in \mathcal{L}(V)$ selbstadjungiert, $\dim V = n < \infty$ mit Eigenwerten $\lambda_1, \dots, \lambda_n$. Für jede Funktion $f : \mathbb{R} \supset U \rightarrow \mathbb{K}$ mit $\{\lambda_1, \dots, \lambda_n\} \subset U$ definiert man

$$f(T) : V \rightarrow V, v \mapsto f(T)v := S \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix} S^{-1}v.$$

3.71 Beispiel. Wurzel eines positiven Operators

Sei $A \in \mathcal{L}(V)$ positiv und selbstadjungiert. Dann ist

$$\sqrt{A} := S \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix} S^{-1}$$

selbst wieder positiv und es gilt $\sqrt{A} \cdot \sqrt{A} = A$.

3.3 Klassifikation von Matrizen

Wir führen nun die Begriffe der Äquivalenzrelation und der zugehörigen Äquivalenzklassen ein. Ziel des Folgenden ist eine mathematische Präzisierung der Idee, dass verschiedenen Objekte im Bezug auf bestimmte Eigenschaften gleich sind. Beispielsweise sind alle Vektorräume der endlichen Dimension n zueinander isomorph.

3.72 Definition. Äquivalenzrelation

Sei M eine Menge. Unter einer **Äquivalenzrelation** auf M versteht man eine Relation (d.h. eine Teilmenge $R \subset M \times M$, wobei man statt $(x, y) \in R$ auch $x \sim y$ schreibt), welche die folgenden Eigenschaften hat.

- (a) Reflexivität: $x \sim x$ für alle $x \in M$
- (b) Symmetrie: $x \sim y \Leftrightarrow y \sim x$
- (c) Transitivität: $x \sim y$ und $y \sim z \Rightarrow x \sim z$

3.73 Beispiel. Sei V ein Vektorraum über \mathbb{K} und $U \subset V$ ein Unterraum. Wir definieren \sim_U auf V durch

$$x \sim_U y \Leftrightarrow x - y \in U.$$

Dann ist \sim_U eine Äquivalenzrelation.

Beweis. (a) $x \sim_U x$, denn $x - x = 0 \in U$ für alle $x \in V$.

(b) $x \sim_U y \Leftrightarrow x - y \in U \Leftrightarrow y - x \in U \Leftrightarrow y \sim_U x$

(c) $x \sim_U y$ und $y \sim_U z \Rightarrow x - y \in U$ und $y - z \in U \Rightarrow x - z = (x - y) + (y - z) \in U \Rightarrow x - z \in U$.

□

3.74 Definition. Äquivalenzklassen

Sei \sim eine Äquivalenzrelation auf M . Für $x \in M$ heißt die Teilmenge

$$[x] := \{y \in M \mid x \sim y\} \subset M$$

die **Äquivalenzklasse** von x bezüglich \sim .

3.75 Satz. Sei \sim eine Äquivalenzrelation auf M . Dann gilt

- (a) $M = \cup_{x \in M} [x]$
- (b) $[x] \cap [y] \neq \emptyset \Leftrightarrow x \sim y \Leftrightarrow [x] = [y]$.

Beweis. (a) ist trivial, da $x \in [x]$. (b) ist es auch, aber das muss man sich zumindest einmal klar machen. Dazu findet man

$$[x] \cap [y] \neq \emptyset \Leftrightarrow \exists z \in [x] \cap [y] \Leftrightarrow \exists z \in M \text{ mit } z \sim x \text{ und } z \sim y \Leftrightarrow x \sim y.$$

Sei nun $x \sim y$ und $z \in [x]$, also $z \sim x$. Dann ist auch $z \sim y$ und somit $z \in [y]$, also $[x] \subset [y]$, und analog findet man $[y] \subset [x]$. Ist umgekehrt $[x] = [y]$, so ist insbesondere $y \in [x]$, also $x \sim y$. \square

3.76 Definition. Eine Menge von nichtleeren Teilmengen von M die paarweise disjunkt sind und deren Vereinigung ganz M ergibt, nennt man Zerlegung von M .

Die Menge der Äquivalenzklassen $\{[x] \mid x \in M\}$ ist also eine Zerlegung von M .

3.77 Definition. Quotient bzgl. einer Äquivalenzrelation

Ist \sim eine Äquivalenzrelation auf M , so nennt man die Menge

$$M/\sim := \{[x] \mid x \in M\}$$

der Äquivalenzklassen den **Quotienten** von M nach \sim . Die Abbildung

$$\pi_{\sim} : M \rightarrow M/\sim, \quad x \mapsto [x]$$

heißt die kanonische Projektion der Äquivalenzrelation.

Man kann also die Elemente der Menge M dadurch klassifizieren, dass man eine Äquivalenzrelation \sim auf M angibt. Geht man nun zum Quotienten M/\sim über, so betrachtet man diejenigen Elemente der Menge M als im wesentlichen gleich, welche unter \sim äquivalent sind.

3.78 Beispiel. Auf dem \mathbb{R}^3 wird durch

$$x \sim y \Leftrightarrow x_3 = y_3$$

eine Äquivalenzrelation definiert. Es ist \mathbb{R}^3/\sim isomorph zu $\text{Span}\{e_3\}$ wenn wir zu jedem $[x] \in \mathbb{R}^3/\sim$ den Vertreter $y \in [x]$ mit $y_1 = y_2 = 0$ wählen. Die kanonische Projektion π_{\sim} ist dann einfach die orthogonale Projektion auf $\text{Span}\{e_3\}$. Diese Charakterisierung von \mathbb{R}^3/\sim ist eine Klassifikation durch Repräsentanten.

Sei \sim eine Äquivalenzrelation auf einer Menge M . Die Menge M “nach \sim ” oder “bis auf \sim ” zu klassifizieren heißt, M/\sim und möglichst auch π_{\sim} zu “verstehen”. Zwei häufige Varianten dieses zunächst etwas vagen Konzepts sind

- (a) **Klassifikation durch Repräsentanten.** Die Klassifikation durch Repräsentanten besteht darin, eine “überschaubare” Teilmenge $M_0 \subset M$ anzugeben, so dass

$$\pi|_{M_0} : M_0 \rightarrow M/\sim$$

bijektiv ist, es also zu jedem $x \in M$ genau ein $x_0 \in M_0$ mit $x \sim x_0$ gibt.

Beispiel 3.78: $M = \mathbb{R}^3$, $M_0 = \text{Span}\{e_3\}$.

- (b) **Klassifikation durch charakteristische Daten.** Die Klassifikation durch charakteristische Daten besteht darin, eine “wohlbekannte” Menge D und eine Abbildung $c : M \rightarrow D$ zu finden, so dass

$$x \sim y \Leftrightarrow c(x) = c(y)$$

gilt. Die Abbildung $c : M/\sim \rightarrow D$, $[x] \mapsto c(x)$ ist dann eine Bijektion. Typische Mengen D die als charakteristische Daten Verwendung finden sind \mathbb{Z} , \mathbb{N} oder Polynome.

3.79 Beispiel. Sei V eine unendliche Menge und M die Menge aller endlichen Teilmengen von V . Für $X, Y \in M$ sei

$$X \sim Y \iff \text{es gibt eine Bijektion } f : X \rightarrow Y.$$

Dann erhält man eine Klassifikation von M nach \sim durch charakteristische Daten, indem man setzt: $D := \mathbb{N}$ und $c(X) := |X| = \text{Anzahl der Elemente von } X$.

3.80 Definition. Zwei Matrizen $A, B \in M(m \times n, \mathbb{K})$ heißen äquivalent, geschrieben $A \sim B$, wenn es invertierbare Matrizen P und Q gibt, sodass

$$B = Q^{-1}AP$$

gilt. Offensichtlich ist hierdurch eine Äquivalenzrelation auf $M(m \times n, \mathbb{K})$ erklärt:

- (a) $A = E_m A E_n$
- (b) $B = Q^{-1}AP \iff A = QBP^{-1}$
- (c) $B = Q^{-1}AP, C = R^{-1}BS \implies C = R^{-1}Q^{-1}APS = (QR)^{-1}A(PS)$.

3.81 Satz. Zwei $m \times n$ -Matrizen sind genau dann äquivalent, wenn sie denselben Rang haben.

Beweis. " \implies ": Sei $A \sim B$, dann ist $\dim(\text{Bild}B) = \dim(\text{Bild}Q^{-1}AP) = \dim(\text{Bild}A)$, da Q^{-1} und P Isomorphismen sind.

" \impliedby ": Sei $\text{Rang}A = \text{Rang}B = r$. Sei (v_1, \dots, v_{n-r}) Basis von $\text{Kern}B$ und (v_1, \dots, v_n) Basis von ganz \mathbb{K}^n . Dann ist $(w_1, \dots, w_r) := (Bv_{n-r+1}, \dots, Bv_n)$ eine Basis von $\text{Bild}B$, die wir zu einer Basis (w_1, \dots, w_m) von ganz \mathbb{K}^m ergänzen. Analog verfahren wir für A und erhalten Basen (v'_1, \dots, v'_n) für \mathbb{K}^n und (w'_1, \dots, w'_m) für \mathbb{K}^m . Seien nun P und Q die Isomorphismen, welche die ungestrichenen in die gestrichenen Basen überführen. Dann gilt

$$QB = AP$$

für die Basisvektoren v_1, \dots, v_n und somit für all $v \in \mathbb{K}^n$. □

Der Rang ist also ein charakteristisches Datum für die Klassifikation der $m \times n$ -Matrizen bis auf Äquivalenz, d.h.

$$M(m \times n, \mathbb{K}) / \sim \cong \{0, 1, 2, \dots, \min(m, n)\}.$$

Man erhält aber auch leicht eine Klassifikation durch Repräsentanten:

$$E_r^{m \times n} = \left(\begin{array}{cc|cc} \hline & & & \\ \hline 1 & 0 & & \\ & \ddots & & 0 \\ 0 & 1 & & \\ \hline & & 0 & 0 \\ \hline \end{array} \right) \begin{array}{l} \\ \\ \\ \\ \\ \end{array} \left. \vphantom{\begin{array}{c} \\ \\ \\ \\ \\ \end{array}} \right\} m$$

$\underbrace{\hspace{10em}}_n$

Man nennt die einfachsten Repräsentanten einer Äquivalenzklasse oft auch Normalformen. Äquivalenz von Matrizen ist offenbar eine sehr grobe Klassifikation. Eine feinere ist Ähnlichkeit.

3.82 Definition. Zwei Matrizen $A, B \in M(n \times n, \mathbb{K})$ heißen **ähnlich**, wenn es eine reguläre $n \times n$ -Matrix P gibt so, dass

$$B = P^{-1}AP$$

gilt.

Auch Ähnlichkeit ist wieder eine Äquivalenzrelation und ähnliche Matrizen sind sicherlich äquivalent im Sinne von Definition 4.8. Umgekehrt sind sicherlich nicht alle äquivalenten Matrizen ähnlich, da z.B. ähnliche Matrizen jeweils das gleiche charakteristische Polynom haben. Wir führen nun die Normalformenklassifikation ähnlicher Matrizen ein, die so genannte Jordansche Normalform.

3.83 Definition. Sei $z \in \mathbb{C}$ und $m \geq 1$. Die $m \times m$ -Matrix

$$J_m(z) := \begin{pmatrix} z & 1 & 0 & \cdots & 0 \\ 0 & z & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & z & 1 \\ 0 & \cdots & \cdots & 0 & z \end{pmatrix}$$

heißt das Jordankästchen der Größe m zum Eigenwert z .

$J_m(z)$ hat den einzigen Eigenwert z mit algebraischer Vielfachheit m und geometrischer Vielfachheit 1: Das charakteristische Polynom von $J_m(z)$ ist

$$P_{J_m(z)}(\lambda) = (z - \lambda)^m$$

und der einzige Eigenvektor ist $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, denn

$$J_m(z) - zE_m = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}.$$

Für $m \geq 2$ ist $J_m(z)$ also so nichtdiagonalisierbar, wie eine komplexe $m \times m$ -Matrix nur sein kann.

Die Jordansche Normalform für eine komplexe $n \times n$ -Matrix A hat nun folgende Form. Zu jedem Eigenwert λ_k gehört eine Blockmatrix B_k aus n_k einzelnen Jordankästchen:

$$B_k = \begin{array}{c} \left. \begin{array}{c|ccc|} \lambda_k & 1 & 0 & & \\ & \ddots & \ddots & & \\ & & \ddots & & \\ 0 & & & 1 & \\ \hline & & & & \lambda_k \end{array} \right\} m_1^{(k)} \\ \left. \begin{array}{c|ccc|} & & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \ddots \end{array} \right\} m_2^{(k)} \\ \left. \begin{array}{c|ccc|} & & & & \\ & & & & \\ & & & & \\ & & & & \\ \hline & & & & \lambda_k & 1 & 0 \\ & & & & \ddots & \ddots & \\ & & & & & \ddots & \\ & & & & & & 1 \\ & & & & & & \lambda_k \\ \hline & & & & & & 0 \end{array} \right\} m_{n_k}^{(k)} \end{array}$$

Die algebraische Vielfachheit von λ_k ist dabei $\sum_{j=1}^{n_k} m_j^{(k)}$ und die geometrische Vielfachheit ist n_k . Die gesamte Jordansche Normalform

$$\begin{array}{c|ccc|} B_1 & & & 0 \\ \hline & B_2 & & \\ & & \ddots & \\ & & & B_r \\ \hline 0 & & & \end{array}$$

3.84 Satz. Jordansche Normalform

Sei $A \in M(n \times n, \mathbb{C})$ mit verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r \in \mathbb{C}$ mit geometrischen Vielfachheiten n_1, \dots, n_r und algebraischen Vielfachheiten ℓ_1, \dots, ℓ_r .

Dann gibt es zu jedem $k = 1, \dots, r$ eindeutig bestimmte Zahlen

$$m_1^{(k)} \leq m_2^{(k)} \leq \dots \leq m_{n_k}^{(k)}, \quad \sum_{j=1}^{n_k} m_j^{(k)} = \ell_k$$

mit der Eigenschaft, dass es eine invertierbare Matrix $P \in M(n \times n, \mathbb{C})$ gibt, für die $P^{-1}AP$ die Blockmatrix ist, welche durch Aneinanderreihung der Jordan Kästchen

$$J_{m_1^{(1)}}(\lambda_1), \dots, J_{m_{n_1}^{(1)}}(\lambda_1), J_{m_1^{(2)}}(\lambda_2), \dots, J_{m_{n_2}^{(2)}}(\lambda_2), \dots, J_{m_1^{(r)}}(\lambda_r), \dots, J_{m_{n_r}^{(r)}}(\lambda_r)$$

längs der Diagonalen entsteht.

Bis auf die Reihenfolge der Eigenwerte liefert die Jordansche Normalform also eine Klassifikation ähnlicher Matrizen durch Repräsentanten.

3.85 Bemerkung. Eine Matrix ist also genau dann diagonalisierbar, wenn alle Jordankästchen die Größe 1 haben.

3.86 Beispiel. Bei der Lösung des gedämpften harmonischen Oszillators, vgl. Beispiel ??, standen wir vor dem Problem eine nichtdiagonalisierbare 2×2 -Matrix zu exponentieren. Gemäß Satz 4.18 können wir diese auf die Form

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

bringen. Dann gilt aber wegen

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix},$$

dass

$$\exp \left[\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} t \right] := \sum_{n=0}^{\infty} \frac{t^n}{n!} \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

3.87 Bemerkung. Für Diagonalmatrizen

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

und Funktionen $f : \mathbb{K} \rightarrow \mathbb{K}$ hatten wir

$$f(D) := \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_n) \end{pmatrix}$$

definiert. In Beispiel 3.86 hatten wir gesehen, dass man auch Jordankästchen exponentieren kann. Wir wollen nochetwas allgemeiner untersuchen, inwiefern man Funktionen von Jordankästchen bilden kann. Dazu setzen wir

$$J_m(\lambda) =: \lambda E_m + N_m$$

mit der nilpotenten Matrix

$$N_m = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}.$$

Eine Matrix oder ein Operator A heißt nilpotent, falls $A^n = 0$ für ein $n \in \mathbb{N}$. Im Fall von N_m gilt

$$((N_m)^n)_{ij} = \begin{cases} 1 & \text{falls } j - i = n \\ 0 & \text{sonst.} \end{cases} \quad \text{für } n \in \mathbb{N}_0,$$

also insbesondere $(N_m)^n = 0$ für $n \geq m$. Damit ist

$$J_m(\lambda)^n = (\lambda E_m + N_m)^n = \sum_{k=0}^n \binom{n}{k} \lambda^{n-k} N_m^k = \sum_{k=0}^{m-1} \binom{n}{k} \lambda^{n-k} N_m^k,$$

also beispielsweise

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}^{23} = \begin{pmatrix} \lambda^{23} & 23 \cdot \lambda^{22} & \frac{23 \cdot 22}{2!} \lambda^{21} \\ 0 & \lambda^{23} & 23 \cdot \lambda^{22} \\ 0 & 0 & \lambda^{23} \end{pmatrix}.$$

Für Polynome bzw. bei 0 analytische Funktionen f (d.h. Funktionen die sich als konvergente Potenzreihe um 0 schreiben lassen) findet man allgemeiner

$$f(J_m(\lambda)) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \dots & \frac{f^{(m-1)}(\lambda)}{(m-1)!} \\ 0 & f(\lambda) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\ \vdots & & \ddots & f(\lambda) & f'(\lambda) \\ 0 & \dots & \dots & 0 & f(\lambda) \end{pmatrix}.$$

Schließlich können wir nun auch noch den Beweis des Satzes von Cayley-Hamilton nachliefern:

3.88 Satz. von Cayley-Hamilton

Sei $L \in \mathcal{L}(V)$ ein Endomorphismus des endlichdimensionalen Vektorraums V und P_L sein charakteristisches Polynom. Dann ist

$$P_L(L) = 0.$$

Hier steht $P_L(L) \in \mathcal{L}(V)$ für diejenige Abbildung, die man erhält, wenn man L in das charakteristische Polynom einsetzt, also

$$P_L(L) = a_n L^n + a_{n-1} L^{n-1} + \dots + a_1 L + a_0 E_n.$$

Beweis. Sei \mathcal{B} eine Basis in der $M_{\mathcal{B}}(L)$ Jordansche Normalform hat. Es genügt nun, jeden Jordanblock für sich zu betrachten. Ist $J_m(\lambda_j)$ ein Block in der Jordanschen Normalform, so ist λ_j sicherlich Eigenwert der algebraischen Vielfachheit mindestens m . Das charakteristische Polynom enthält also einen Faktor $(\lambda - \lambda_j)^m$. Da

$$(J_m(\lambda_j) - \lambda_j E_m)^m = N_m^m = 0,$$

folgt $P_L(L) = 0$. □

Auch die Hauptachsentransformation selbstadjungierter Matrizen liefert eine Klassifikation durch Repräsentanten.

3.89 Definition. Es bezeichne $\text{Sym}(n, \mathbb{K})$ den Vektorraum der selbstadjungierten $n \times n$ -Matrizen. Zwei Matrizen heißen **orthogonal ähnlich**, wenn es eine orthogonale Matrix $P \in U(n)$ gibt, so dass

$$B = P^{-1}AP$$

gilt.

3.90 Satz. Jede selbstadjungierte $n \times n$ -Matrix ist bis auf die Reihenfolge der Eigenwerte zu genau einer Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

orthogonal ähnlich.

Der Weg zur Jordanschen Normalform: eine Beweisskizze von Satz 4.18

1. Schritt: Jede komplexe $n \times n$ -Matrix A lässt sich durch eine unitäre Ähnlichkeitstransformation Q auf obere Dreiecksform

$$R = \begin{pmatrix} \lambda_1 & r_{12} & \cdots & r_{1n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_{n-1,n} \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = Q^{-1}AQ$$

bringen, wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A , wiederholt gemäß ihrer algebraischen Vielfachheit, sind.

Beweis. Induktion über n , wobei $n = 1$ klar ist. Es gelte die Aussage also für $n - 1$ und es sei $A \in M(n, \mathbb{C})$. Dann hat A mindestens einen Eigenvektor v und es gilt $Av = \lambda v$. Sei T ein Basiswechsel zu einer ONB (v, w_1, \dots, w_{n-1}) , dann ist

$$\tilde{A} = T^{-1}AT = \begin{pmatrix} \lambda & x^T \\ 0 & B \end{pmatrix}$$

mit $x \in \mathbb{C}^{n-1}$. Nach Induktionsannahme gibt es für $B \in M(n - 1, \mathbb{C})$ eine unitäre Abbildung \tilde{Q} mit

$$\tilde{Q}^{-1}B\tilde{Q} = R_{n-1}.$$

Sei $Q := T \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{pmatrix}$, dann ist

$$\begin{aligned} Q^{-1}AQ &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q}^{-1} \end{pmatrix} T^{-1}AT \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{Q}^{-1} \end{pmatrix} \begin{pmatrix} \lambda & x^T \tilde{Q} \\ 0 & B \tilde{Q} \end{pmatrix} \\ &= \begin{pmatrix} \lambda & x^T \tilde{Q} \\ 0 & \tilde{Q}^{-1}B\tilde{Q} \end{pmatrix} = \begin{pmatrix} \lambda & x^T \tilde{Q} \\ 0 & R_{n-1} \end{pmatrix} = R_n. \end{aligned}$$

□

2. Schritt: Hauptvektoren

3.91 Definition. Sei λ Eigenwert von $A \in M(n, \mathbb{C})$ mit algebraischer Vielfachheit m , dann heißt jeder Vektor v mit

$$(A - \lambda E_n)^m v = 0, \quad v \neq 0,$$

Hauptvektor zum Eigenwert λ .

3.92 Satz. Alle Hauptvektoren zu einem Eigenwert λ von $A \in M(n, \mathbb{C})$ mit algebraischer Vielfachheit m bilden einen m -dimensionalen Unterraum H_λ , den so genannten Hauptraum des Eigenwertes. Dieser ist invariant unter A . Der Gesamtraum ist die direkte Summe der Haupträume,

$$\mathbb{C}^n = \bigoplus_{\lambda} H_{\lambda}.$$

3.93 Definition. Man sagt eine Summe $\sum_{\lambda} H_{\lambda} = \text{Span}\{\cup_{\lambda} H_{\lambda}\}$ ist direkt und schreibt $\bigoplus_{\lambda} H_{\lambda}$, falls sich jeder Vektor $v \in \sum_{\lambda} H_{\lambda}$ in eindeutiger Weise als Linearkombination $v = \sum_{\lambda} v_{\lambda}$ mit $v_{\lambda} \in H_{\lambda}$ schreiben lässt.

Beweis. (von Satz 3.92) Sei $v \in H_{\lambda}$ also $(A - \lambda)^m v = 0$. Dann ist auch $Av \in H_{\lambda}$, denn

$$(A_{\lambda})^m Av = (A_{\lambda})^m (A - \lambda)v = (A_{\lambda})(A - \lambda)^m v = 0.$$

Um $\dim H_{\lambda}$ zu bestimmen, transformiert man A auf obere Dreiecksform:

$$A \rightarrow B = Q^{-1}AQ = \begin{pmatrix} \lambda & & & * \\ & \ddots & & \\ & & \lambda & \\ 0 & & & R \end{pmatrix}$$

Es ist klar, dass $(A - \lambda)^m v = 0$ genau dann, wenn $(B - \lambda)^m Q^{-1}v = 0$, die Haupträume von A und B zu λ also die gleiche Dimension haben. Nun sind die Vektoren $e_i, i = 1, \dots, m$, genau die Hauptvektoren von B , also $\dim H_\lambda = m$. Da $\sum_\lambda \dim H_\lambda = \sum_\lambda m_\lambda = n$ ist, genügt es, die lineare Unabhängigkeit der Haupträume zu zeigen. Das ist eine Übungsaufgabe. \square

3. Schritt: Zyklische Teilräume

Jeder Hauptraum H_λ einer Matrix $A \in M(n, \mathbb{C})$ läßt sich in eine direkte Summe zyklischer Teilräume zerlegen:

$$H_\lambda = V_1 \oplus V_2 \oplus \dots \oplus V_\ell. \quad (3.2)$$

Ein Teilraum V heißt hier zyklisch, falls es einen Vektor $v \in V$ gibt (einen zyklischen Vektor), so dass die Vektoren

$$v, (A - \lambda E)v, (A - \lambda E)^2 v, \dots, (A - \lambda E)^k v,$$

eine Basis von V bilden und $w := (A - \lambda E)^k v$ Eigenvektor von A zum Eigenwert λ ist. D.h., jeder der Teilräume V_i in (3.2) besitzt eine Basis der Form

$$B^{k_i} v_i, B^{k_i-1} v_i, \dots, v_i,$$

wobei wir $B = (A - \lambda E)$ abkürzen und $w_i := B^{k_i} v_i$ Eigenvektor von A zum Eigenwert λ ist. Es ist dann $\dim V_i = k_i$ und $\sum_{i=1}^\ell k_i = \dim H_\lambda$.

Es läßt A die zyklischen Teilräume V_i invariant und hat dort die Matrixdarstellung

$$J_{k_i}(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

Beweis. (Skizze) Um die geeigneten Eigenvektoren $w_i \in \text{Kern} B$ zu finden, betrachtet man die Teilräume

$$U_j := \text{Kern} B \cap \text{Bild} B^j \subset \text{Kern} B = E_\lambda.$$

Offensichtlich ist

$$\{0\} \subset U_m \subset U_{m-1} \subset \dots \subset U_0 = \text{Kern} B.$$

Man bestimmt nun die w_i induktiv, ausgehend von der leeren Menge. Sind w_1, \dots, w_{i-1} festgelegt, so nimmt man das größte k_i , für welches U_{k_i} nicht in $\text{Span}\{w_1, \dots, w_{i-1}\}$ enthalten ist, es also ein $v_i \neq 0$ gibt mit

$$B^{k_i} v_i = w_i \notin \text{Span}\{w_1, \dots, w_{i-1}\}, \quad B w_i = 0.$$

So entsteht eine Basis $\{w_1, \dots, w_\ell\}$ von E_λ und zyklische Basen der zugehörigen V_i . Jetzt muss man noch

$$H_\lambda = V_1 \oplus \dots \oplus V_\ell$$

zeigen.

(i) Die lineare Unabhängigkeit zeigen wir nur am Beispiel zweier zyklischer Ketten

$$B v_1, v_1, B v_2, v_2.$$

Sei

$$\alpha_1 B v_1 + \alpha_2 v_1 + \beta_1 B v_2 + \beta_2 v_2 = 0,$$

dann liefert Multiplikation mit B

$$\alpha_2 \underbrace{B v_1}_{w_1} + \beta_2 \underbrace{B v_2}_{w_2} = 0.$$

Es sind aber w_1 und w_2 nach Konstruktion linear unabhängig, also $\alpha_2 = \beta_2 = 0$ und somit auch $\alpha_1 = \beta_1 = 0$.

- (ii) Es bleibt zu zeigen, dass jeder Vektor $u \in H_\lambda$ durch die Vektoren $B^j v_i$ darstellbar ist. Sei also $B^k u = 0$ für ein $k \leq m$. Für $k = 1$ bilden die w_i nach Konstruktion eine Basis für den Eigenraum $\text{Kern} B$. Seien also die Vektoren in $\text{Kern} B^{k-1}$ darstellbar und $u \in \text{Kern} B^k$. Es ist

$$\underbrace{B^{k-1} u}_{\in \text{Kern} B} = \sum_i \alpha_i w_i = \sum_i \alpha_i B^{k_i} v_i,$$

wobei $\alpha_i = 0$ falls $k_i < k - 1$. Damit ergibt sich

$$B^{k-1} \underbrace{\left(u - \sum_i \alpha_i B^{k_i - k + 1} v_i \right)}_{u'} = 0.$$

Also ist u' darstellbar und somit auch u . □

3.94 Beispiel. Für eine komplexe 3×3 -Matrix gibt es die folgenden qualitativ verschiedenen Jordanschen Normalformen $J = Q^{-1} A Q$.

- (i) Paarweise verschiedenen Eigenwerte λ, ρ, σ :

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & \sigma \end{pmatrix}.$$

- (ii) Ein doppelter Eigenwert λ :

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \rho \end{pmatrix} \quad \text{oder} \quad J = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \rho \end{pmatrix}.$$

Im zweiten Fall ist $Q = (w_\lambda, v_\lambda, w_\rho)$ mit Eigenvektoren w_λ, w_ρ welche man aus

$$(A - \lambda)w_\lambda \quad \text{und} \quad (A - \rho)w_\rho = 0$$

erhält. Den Hauptvektor v_λ bekommt man als Lösung von

$$(A - \lambda)v_\lambda = w_\lambda.$$

- (iii) Dreifacher Eigenwert λ :

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

Im zweiten Fall bestimmt man zunächst einen Hauptvektor v durch Lösen von

$$(A - \lambda)v \neq 0.$$

Der zugehörige Eigenvektor ist $w = (A - \lambda)v$. Schließlich bestimmt man noch einen weiteren linear unabhängigen Eigenvektor aus

$$(A - \lambda)\tilde{w} = 0 \quad \text{mit} \quad \langle w, \tilde{w} \rangle = 0.$$

Dann ist $Q = (w, v, \tilde{w})$.

Im dritten Fall ist $Q = (w, v, u)$ mit

$$(A - \lambda)w = 0, \quad (A - \lambda)v = w, \quad (A - \lambda)u = v.$$

4 Lineare Differentialgleichungen

4.1 Definition. Lineare Systeme

Sei $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \rightarrow M(n, \mathbb{R})$ eine stetige Abbildung.

(a) Man nennt dann die Differentialgleichung

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

ein **nicht-autonomes, homogenes, lineares System**.

(b) Ist $b : I \rightarrow \mathbb{R}^n$ stetig, so heißt

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

ein **nicht-autonomes, inhomogenes, lineares System**.

4.2 Motivation. Aus der linearen Algebra wissen wir, dass im autonom homogenen Fall

$$\dot{x} = Ax$$

die globale Lösung durch $x(t) = e^{At}x(0)$ gegeben ist. Es liegt nahe, dass auch im nicht-autonomen linearen Fall die Lösungen höchstens exponentiell wachsen können, falls $\|A(t)\|$ beschränkt bleibt. Da sie dann aber nicht in endlicher Zeit nach Unendlich laufen können, existieren sie gemäß Satz 1.39 für alle Zeiten in I .

4.3 Satz. Globale Existenz für lineare Systeme

Sei $I \subset \mathbb{R}$ offenes Intervall, $A : I \rightarrow M(n, \mathbb{R})$ und $b : I \rightarrow \mathbb{R}^n$ seien stetig. Dann existiert zu jedem $t_0 \in I$ und $x_0 \in \mathbb{R}^n$ eine eindeutige maximale Lösung $x : I \rightarrow \mathbb{R}^n$ von

$$\dot{x} = A(t)x + b(t) \quad \text{mit} \quad x(t_0) = x_0.$$

Beweis. Wir führen den Beweis nur für lokal Lipschitz-stetiges $A(t)$ und $b(t)$ aus. Sonst müssten wir im Beweis nochmals die Picard-Iteration für dieses spezielle Problem durchgehen, worauf wir aus zeitgründen verzichten. Gemäß Bemerkung 1.22 liefert die Anwendung von Satz 1.36 auf das lokal Lipschitz-stetige Vektorfeld (hier verwenden wir die Zusatzannahme, dass $A(t)$ und $b(t)$ auch lokal Lipschitz-stetig sind!)

$$v(t, x) = \begin{pmatrix} 1 \\ A(t)x + b(t) \end{pmatrix}$$

die Existenz einer eindeutigen maximalen Lösung auf dem Zeitintervall $J := (s_-, s_+) \subset I =: (t_-, t_+)$. Wir zeigen nun $s_+ = t_+$ (analog sieht man dann $s_- = t_-$, also $I = J$). Angenommen $s_+ < t_+$, dann muss nach Satz 1.39 die maximale Lösung $(s_-, s_+) \rightarrow I \times \mathbb{R}^n$, $t \mapsto (t, x(t))$ jedes Kompaktum von $I \times \mathbb{R}^n$ verlassen, also $\|x(t)\| \rightarrow \infty$ für $t \rightarrow s_+$ gelten. Wir zeigen nun, dass $\|x(t)\|$ beschränkt bleibt für $t \rightarrow s_+$ falls $s_+ < t_+$. Insgesamt folgt dann $s_+ = t_+$.

Setze dazu $L := \max\{\|A(t)\| \mid t_0 \leq t \leq s_+\}$ und $M := \max\{\|b(t)\| \mid t_0 \leq t \leq s_+\}$, so gilt für die stetige Funktion $u : [t_0, s_+) \rightarrow [0, \infty)$, $u(t) = \|x(t)\|$, dass

$$\begin{aligned} u(t) &= \|x(t)\| = \|x(t_0) + \int_{t_0}^t \dot{x}(s) ds\| \leq \|x(t_0)\| + \int_{t_0}^t \|\dot{x}(s)\| ds \\ &= u(t_0) + \int_{t_0}^t \|A(s)x(s) + b(s)\| ds \\ &\leq \left(u(t_0) + \int_{t_0}^{s_+} \|b(s)\| ds \right) + \int_{t_0}^t \underbrace{\|A(s)\|}_{\leq L} \cdot \underbrace{\|x(s)\|}_{=u(s)} ds \\ &\leq C + L \int_{t_0}^t u(s) ds. \end{aligned}$$

4.4 Lemma. von Gronwall

Sei $a < b$ und $u : [a, b] \rightarrow [0, \infty)$ eine stetige Funktion. Es gebe Konstanten $L, C \geq 0$ so, dass für alle $t \in [a, b]$ gilt

$$u(t) \leq C + L \int_a^t u(s) \, ds.$$

Dann ist

$$u(t) \leq Ce^{L(t-a)} \quad \text{für alle } t \in [a, b].$$

Ende des Beweises von Satz 4.3. Also ist $u(t) \leq Ce^{L(t-t_0)} \leq Ce^{L(s_+-t_0)}$ für alle $t_0 \leq t < s_+$ und somit $t \mapsto \|x(t)\|$ beschränkt für $t \rightarrow s_+$. \square

Beweis. des Gronwall Lemmas. Sei zunächst $C > 0$. Setze $U : [a, b] \rightarrow [0, \infty)$

$$U(t) := C + L \int_a^t u(s) \, ds.$$

Dann ist U stetig differenzierbar und $\dot{U}(t) = Lu(t) \geq 0$, also monoton steigend. Außerdem ist $U(a) = C > 0$, also $U(t) \geq C$ für alle $t \in [a, b]$. Es folgt

$$\frac{d}{dt} \ln(U(t)) = \frac{\dot{U}(t)}{U(t)} = \frac{Lu(t)}{U(t)} \leq L$$

da $u \leq U$. Also ist

$$\ln U(t) - \ln C = \ln U(t) - \ln U(a) = \int_a^t \frac{d}{ds} \ln U(s) \, ds \leq \int_a^t L \, ds = L(t - a).$$

und somit

$$u(t) \leq U(t) = \exp(\ln U(t)) \leq \exp(\ln C + L(t - a)) = Ce^{L(t-a)}.$$

Falls $C = 0$, so liefert unser Argument, dass $u(t) \leq \tilde{C}e^{Lt}$ für jedes $\tilde{C} > 0$, also $u(t) \equiv 0$. \square

Die wichtigste Eigenschaft homogener linearer Systeme ist, dass Linearkombinationen von Lösungen wieder Lösungen sind.

4.5 Satz. Der Lösungsraum homogener linearer Systeme

Sei $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \rightarrow M(n, \mathbb{R})$ stetig. Sei $L_h \subset C^1(I, \mathbb{R}^n)$ die Menge aller maximalen Lösungen des homogenen linearen Systems

$$\dot{x} = A(t)x \quad (*)$$

auf \mathbb{R}^n . Dann gilt:

- (a) L_h ist ein n -dimensionaler Unterraum von $C^1(I, \mathbb{R}^n)$.
- (b) Ist $t_0 \in I$, $r \in \mathbb{N}$ und $\xi_1, \dots, \xi_r \in \mathbb{R}^n$, so seien $x_1, \dots, x_r \in L_h$ die Lösungen von $\dot{x} = A(t)x$ mit $x_j(t_0) = \xi_j$, $j = 1, \dots, r$. Dann sind äquivalent:
 - (i) (x_1, \dots, x_r) ist linear unabhängig in L_h .
 - (ii) $(x_1(t), \dots, x_r(t))$ ist linear unabhängig in \mathbb{R}^n für alle $t \in I$.
 - (iii) (ξ_1, \dots, ξ_r) ist linear unabhängig in \mathbb{R}^n .

Beweis. Zu (a): Seien $x, y \in L_h$ und $\lambda \in \mathbb{R}$. Dann ist

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

und

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

Also sind auch $x + y$ und λx Lösungen von (*) und somit ist L_h ein Untervektorraum. Die Aussage zur Dimension folgt sofort aus Teil (b).

Zu (b): (i) \Rightarrow (ii): Seien $(x_1, \dots, x_r) \subseteq L_h$ linear unabhängig und sei $\tilde{t} \in I$ beliebig aber fest. Erfüllen dann $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ die Gleichung

$$\lambda_1 x_1(\tilde{t}) + \dots + \lambda_r x_r(\tilde{t}) = 0 \quad (**)$$

als Gleichung in \mathbb{R}^n , so ist zu zeigen, dass aus (**) schon $\lambda_j = 0$ für $j = 1, \dots, r$ folgt. Setze dazu $x : I \rightarrow \mathbb{R}^n$, $x(t) := \lambda_1 x_1(t) + \dots + \lambda_r x_r(t)$. Nach (a) ist $x(t)$ Lösung von (*) und wegen (**) ist $x(\tilde{t}) = 0$ und somit, wegen der Eindeutigkeit der Lösung, $x(t) = 0$ für alle $t \in I$. Also ist

$$\lambda_1 x_1 + \dots + \lambda_r x_r = 0$$

in $C^1(I, \mathbb{R}^n)$ und nach Voraussetzung $\lambda_1 = \dots = \lambda_r = 0$. Es ist also $(x_1(\tilde{t}), \dots, x_r(\tilde{t})) \subset \mathbb{R}^n$ linear unabhängig für alle $\tilde{t} \in I$.

(ii) \Rightarrow (iii): klar, weil $\xi_j = x_j(t_0)$ ist,

(iii) \Rightarrow (i): auch klar, denn ist

$$\lambda_1 x_1 + \dots + \lambda_r x_r = 0$$

so ist insbesondere

$$\lambda_1 \xi_1 + \dots + \lambda_r \xi_r = \lambda_1 x_1(t_0) + \dots + \lambda_r x_r(t_0) = 0.$$

Also muss nach Voraussetzung $\lambda_1 = \dots = \lambda_r = 0$ sein. □

4.6 Korollar. Der Propagator

Sei $A : I \rightarrow M(n, \mathbb{R})$ stetig, $t_0 \in I$ und $\varphi^t : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der zu $\dot{x} = A(t)x$ gehörige Fluss (vgl. Bemerkung 1.12 (c)) zum Anfangszeitpunkt t_0 : Sei $x : I \rightarrow \mathbb{R}^n$ die Lösung von $\dot{x} = A(t)x$ zum Anfangswert $x(t_0) = x_0$, so ist also $\varphi^t(x_0) = x(t)$.

Dann ist φ^t für jedes feste $t \in I$ ein linearer Isomorphismus des \mathbb{R}^n , welcher im Folgenden auch mit $\Phi(t)$ bezeichnet und **Propagator** genannt wird.

Die Abbildung $\Phi : \mathbb{R} \rightarrow L(\mathbb{R}^n)$ erfüllt die Matrix Differentialgleichung

$$\dot{\Phi}(t) = A(t)\Phi(t) \quad \text{mit Anfangsbedingung} \quad \Phi(t_0) = \text{Id}.$$

Beweis. Übungsaufgabe. □

4.7 Definition. Fundamentalsystem

Sei $A : I \rightarrow M(n, \mathbb{R})$ stetig. Ein n -Tupel $(x_1, \dots, x_n) \subset L_h$ von Lösungen der homogenen Gleichung $\dot{x} = A(t)x$ heißt **Lösungs-Fundamentalsystem**, wenn (x_1, \dots, x_n) eine Basis von L_h ist.

4.8 Satz. Der Lösungsraum inhomogener linearer Systeme

Sei $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \rightarrow M(n, \mathbb{R})$ und $b : I \rightarrow \mathbb{R}^n$ seien stetig. Mit $L_i \subset C^1(I, \mathbb{R}^n)$ werde die Teilmenge aller Lösungen der inhomogenen Gleichung

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

bezeichnet, also

$$L_i = \{x \in C^1(I, \mathbb{R}^n) \mid \dot{x} = A(t)x + b(t)\}$$

und

$$L_h = \{x \in C^1(I, \mathbb{R}^n) \mid \dot{x} = A(t)x\}.$$

(a) Ist $x : I \rightarrow \mathbb{R}^n$ eine Lösung der inhomogenen Gleichung, so gilt

$$L_i = x + L_h := \{x + \varphi \in C^1(I, \mathbb{R}^n) \mid \varphi \in L_h\}.$$

(b) **Variation der Konstanten**

Sei $\Phi(t) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ der Propagator des homogenen Systems zur Anfangszeit $t_0 \in I$ und sei $x_0 \in \mathbb{R}^n$. Dann ist

$$x(t) = \Phi(t) \left(x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right)$$

die Lösung der inhomogenen Gleichung $\dot{x} = A(t)x + b(t)$ zum Anfangswert $x(t_0) = x_0$.

Beweis. (a) Ist x eine Lösung von $\dot{x} = A(t)x + b(t)$, so ist \tilde{x} genau dann eine weitere Lösung von $\dot{x} = A(t)x + b(t)$, wenn $\varphi := \tilde{x} - x$ Lösung der homogenen Gleichung ist:

$$\dot{\tilde{x}} = A(t)\tilde{x} + b(t) \quad \Leftrightarrow \quad \dot{\varphi} = \dot{\tilde{x}} - \dot{x} = A(t)\tilde{x} - A(t)x = A(t)\varphi.$$

(b) Es gilt $x(t_0) = \Phi(t_0)x_0 = x_0$ und

$$\begin{aligned} \dot{x}(t) &= \dot{\Phi}(t) \left(x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + \Phi(t)\Phi(t)^{-1} b(t) \\ &= A(t)\Phi(t) \left(x_0 + \int_{t_0}^t \Phi(s)^{-1} b(s) ds \right) + b(t) \\ &= A(t)x(t) + b(t). \end{aligned}$$

□

4.9 Bemerkung. Die Bezeichnung „Variation der Konstanten“ kommt aus folgendem Ansatz für die Lösung: $x(t) = \Phi(t)c(t)$ statt $x(t) = \Phi(t)x_0$ wie für die homogene Gleichung. Dann muss nämlich $c(t)$ folgende Differentialgleichung erfüllen:

$$\dot{x} = \dot{\Phi}c + \Phi\dot{c} = A\Phi c + \Phi\dot{c} = x + \Phi\dot{c} \stackrel{!}{=} x + b(t)$$

also $\Phi\dot{c} = b$ oder $\dot{c}(t) = \Phi^{-1}(t)b(t)$. Somit folgt

$$c(t) = x_0 + \int_{t_0}^t \Phi^{-1}(s)b(s) ds.$$

4.10 Bemerkung. Die Menge aller Lösungen des homogenen Systems $\dot{x} = A(t)x$ ist ein n -dimensionaler Unterraum von $C^1(I, \mathbb{R}^n)$. Die Menge aller Lösungen des inhomogenen Systems ist ein n -dimensionaler affiner Unterraum von $C^1(I, \mathbb{R}^n)$. Kennt man die vollständige Lösung des homogenen Systems (also zumindest n linear unabhängige Lösungen), so kann man auch das inhomogene System vollständig lösen.

4.11 Beispiel. Betrachte die Differentialgleichung

$$\dot{x} = 2tx + t^3 \tag{*}$$

auf \mathbb{R} . Mit $a(t) = 2t$ und $b(t) = t^3$ wird das zu $\dot{x} = a(t)x + b(t)$. Die Lösung der homogenen Gleichung $\dot{x} = a(t)x$ mit $x(t_0) = x_0$ ist

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

also mit $t_0 = 0$

$$x(t) = e^{\int_0^t 2(s) ds} x_0 = e^{t^2} x_0.$$

Da $\dim L_h = 1$ ist, ist $x(t) = e^{t^2} x_0$ eine Basis für L_h und der Propagator ist $\Phi(t) = e^{t^2}$. Variation der Konstanten liefert nun die Lösung der inhomogenen Gleichung zum Anfangswert x_0 durch

$$x_i(t) = e^{t^2} \left(x_0 + \int_0^t e^{-s^2} s^3 ds \right) = e^{t^2} \left(x_0 + \frac{1}{2} \right) - \frac{1}{2}(1 + t^2),$$

wobei

$$\int_0^t e^{-s^2} s^3 ds = \frac{1}{2} \int_0^{t^2} e^{-\tau} \tau d\tau = -\frac{1}{2} e^{-\tau} \tau \Big|_0^{t^2} + \frac{1}{2} \int_0^{t^2} e^{-\tau} d\tau = -\frac{1}{2} e^{-t^2} t^2 - \frac{1}{2} e^{-t^2} + \frac{1}{2}.$$

4.12 Bemerkung. Dyson Reihe

Auch im Fall $n > 1$ kann man noch eine explizite Formel für die Lösung der homogenen Gleichung angeben, die sogenannte Dyson Reihe. Sei $A : I \rightarrow M(n, \mathbb{R})$ stetig, dann ist die Lösung $\Phi : I \rightarrow M(n, \mathbb{R})$ von

$$\dot{\Phi} = A(t)\Phi$$

mit $\Phi(t_0) = E_n$ gegeben durch die absolut konvergente Reihe

$$\Phi(t) = E_n + \sum_{j=1}^{\infty} \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \cdots \int_{t_0}^{\tau_{j-1}} d\tau_j A(\tau_1) \cdots A(\tau_j).$$

Beweis. Übungsaufgabe. □

4.13 Bemerkung. Für den autonomen Fall $A(t) \equiv A$ erhält man die Lösung $\Phi : \mathbb{R} \rightarrow M(n, \mathbb{R})$ von

$$\dot{\Phi} = A\Phi$$

mit $\Phi(0) = E_n$ durch exponentieren

$$\Phi(t) = e^{At}.$$

Insbesondere existieren die Lösungen für alle Zeiten. Um $e^{At} := \sum_{j=0}^{\infty} \frac{(At)^j}{j!}$ explizit auszurechnen, transformiert man A auf Jordansche Normalform (vgl. Lineare Algebra).

Sei z.B. A diagonalisierbar, also

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = D$$

eine Diagonalmatrix und $S \in GL_n(\mathbb{R})$, so löst $\Psi = S^{-1}\Phi S$ die Gleichung

$$\dot{\Psi} = S^{-1}\dot{\Phi}S = S^{-1}A\Phi S = S^{-1}AS S^{-1}\Phi S = D\Psi,$$

also

$$\Psi(t) = e^{Dt} = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix}$$

und somit

$$\Phi(t) = S \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} S^{-1}.$$

Um die Normalformtheorie von Matrizen über \mathbb{C} benutzen zu können (viele Matrizen haben eben keine reellen Eigenwerte, aber komplexe), betrachtet man die Differentialgleichung $\dot{x} = A(t)x$ statt auf \mathbb{R}^n auf \mathbb{C}^n und sucht Lösungen $z : I \rightarrow \mathbb{C}^n$ von

$$\dot{z} = A(t)z.$$

Die Zeitvariable bleibt aber reell! Die Lösungen bilden dann einen Unterraum der komplexen Dimension n in $C^1(I, \mathbb{C}^n)$ und alles in diesem Kapitel gesagte gilt analog.

Beachte, dass für $A(t) \in M(n, \mathbb{R})$ zwar S und $\lambda_1, \dots, \lambda_n$ komplex sein können, der Propagator

$$\Phi(t) = S \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} S^{-1}$$

ist aber immer noch reell.

Die Resultate zu linearen Differentialgleichungssystemen 1. Ordnung lassen sich direkt auf lineare (Systeme) n -ter Ordnung übertragen:

4.14 Definition. Homogene lineare DGL m -ter Ordnung

Sei $I \subset \mathbb{R}$ ein offenes Intervall und $a_k : I \rightarrow \mathbb{K}$, $k = 1, \dots, m - 1$, stetige Funktionen. Hier ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann heißt

$$x^{(m)}(t) + a_{m-1}(t)x^{(m-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = 0$$

homogene lineare Differentialgleichung m -ter Ordnung.

Ist $b : I \rightarrow \mathbb{K}$ stetig, so heißt

$$x^{(m)}(t) + a_{m-1}(t)x^{(m-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = b(t)$$

inhomogene lineare Differentialgleichung m -ter Ordnung.

4.15 Satz. Lösungsraum linearer DGLen m -ter Ordnung

- (a) Sei L_h die Menge aller Lösungen $x : I \rightarrow \mathbb{K}$ der homogenen Gleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = 0.$$

Dann ist L_h ein m -dimensionaler Unterraum von $C^m(I, \mathbb{K})$.

- (b) Sei L_i die Menge aller Lösungen $x : I \rightarrow \mathbb{K}$ der inhomogenen Differentialgleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = b.$$

Dann gilt für beliebiges $\tilde{x} \in L_i$

$$L_i = \tilde{x} + L_h.$$

- (c) Ein m -Tupel $(\varphi_1, \dots, \varphi_m)$ in L_h ist genau dann linear unabhängig, wenn für ein und damit für alle $t \in I$ die „Wronski-Determinante“

$$W(t) := \det \begin{pmatrix} \varphi_1(t) & \dots & \varphi_m(t) \\ \dot{\varphi}_1(t) & \dots & \dot{\varphi}_m(t) \\ \vdots & & \vdots \\ \varphi_1^{(m-1)}(t) & \dots & \varphi_m^{(m-1)}(t) \end{pmatrix}$$

von Null verschieden ist.

Beweis. Die Differentialgleichung

$$x^{(m)} + a_{m-1}x^{(m-1)} + \dots + a_1\dot{x} + a_0x = b \quad (*)$$

ist äquivalent zu dem inhomogenen linearen System 1. Ordnung

$$\begin{aligned} \dot{y}_0 &= y_1 \\ \dot{y}_1 &= y_2 \\ &\vdots \\ \dot{y}_{m-2} &= y_{m-1} \\ \dot{y}_{m-1} &= -a_0y_0 - a_1y_1 - \dots - a_{m-1}y_{m-1} + b \end{aligned} \quad (**)$$

Jeder Lösung $x : I \rightarrow \mathbb{K}$ von $(*)$ entspricht eine Lösung

$$\begin{pmatrix} x \\ \dot{x} \\ \vdots \\ x^{(m-1)} \end{pmatrix} : I \rightarrow \mathbb{K}^m$$

von $(**)$ und umgekehrt. Entsprechendes gilt für die homogenen Gleichungen ($b = 0$). Damit folgen die Behauptungen aus Satz 4.8. □

4.16 Beispiel. Die Differentialgleichung

$$\ddot{x} - \frac{1}{2t} \dot{x} + \frac{1}{2t^2} x = 0$$

auf dem Intervall $I = (0, \infty)$ besitzt die Lösungen $\varphi_1(t) := t$ und $\varphi_2(t) := \sqrt{t}$, wovon man sich durch Einsetzen überzeugt.

Die Wronski-Determinante von (φ_1, φ_2) ist

$$W(t) = \det \begin{pmatrix} \varphi_1(t) & \varphi_2(t) \\ \dot{\varphi}_1(t) & \dot{\varphi}_2(t) \end{pmatrix} = \det \begin{pmatrix} t & \sqrt{t} \\ 1 & \frac{1}{2\sqrt{t}} \end{pmatrix} = -\frac{\sqrt{t}}{2}.$$

Da $W(t) \neq 0$ für $t \in I$, bilden φ_1 und φ_2 ein Lösungs-Fundamentalsystem. Die allgemeine Lösung der Differentialgleichung ist also

$$\varphi(t) = c_1 \varphi_1(t) + c_2 \varphi_2(t) = c_1 t + c_2 \sqrt{t}$$

mit beliebige Konstanten $c_1, c_2 \in \mathbb{C}$.

4.17 Beispiele. (a) Die Legendresche Differentialgleichung

Die Legendresche Differentialgleichung auf $I = (-1, 1)$ zu $n \in \mathbb{N}$ ist

$$(1 - t^2)\ddot{x} - 2t\dot{x} + n(n+1)x = 0$$

bzw., da $(1 - t^2) \neq 0$,

$$\ddot{x} - \frac{2t}{(1-t^2)} \dot{x} + \frac{n(n+1)}{(1-t^2)} x = 0.$$

Das Legendre Polynom der Ordnung n ist definiert durch

$$P_n(t) := \frac{1}{2^n n!} \left(\frac{d}{dt} \right)^n (t^2 - 1)^n$$

und löst die Legendresche Differentialgleichung der Ordnung n .

(b) Die Hermitesche Differentialgleichung

Die Hermitesche Differentialgleichung auf $I = \mathbb{R}$ zu $n \in \mathbb{N}$ ist

$$\ddot{x} - 2t\dot{x} + 2nx = 0.$$

Das Hermite Polynom der Ordnung n ist definiert durch

$$H_n(t) = (-1)^n e^{t^2} \left(\frac{d}{dt} \right)^n e^{-t^2}$$

und löst die Hermitesche Differentialgleichung der Ordnung n .

(c) Die Laguerresche Differentialgleichung

Die Laguerresche Differentialgleichung auf $I = (0, \infty)$ zu $n \in \mathbb{N}$ ist

$$t\ddot{x} + (1-t)\dot{x} + nx = 0.$$

Das Laguerresche Polynom der Ordnung n ist definiert durch

$$L_n(t) := e^t \left(\frac{d}{dt} \right)^n (t^n e^{-t})$$

und löst die Laguerresche Differentialgleichung der Ordnung n .

In allen drei Fällen (a), (b) und (c) sind die Polynome jeweils nur eine spezielle Lösung und gemäß Satz 4.15 gibt es jeweils noch eine weitere linear unabhängige Lösung.

4.18 Satz. Sei $\varphi : I \rightarrow \mathbb{R}$ eine Lösung der homogenen linearen Differentialgleichung 2. Ordnung

$$\ddot{x} + a(t)\dot{x} + b(t)x = 0.$$

Im Intervall $J \subset I$ gelte $\varphi(t) \neq 0$. Dann erhält man über J eine zweite von φ linear unabhängige Lösung $\psi : J \rightarrow \mathbb{R}$ durch den Ansatz

$$\psi(t) = \varphi(t)u(t),$$

wobei u eine nicht-konstante Lösung der Differentialgleichung

$$\ddot{u} + \left(2 \frac{\dot{\varphi}(t)}{\varphi(t)} + a(t)\right) \dot{u} = 0 \quad (*)$$

ist.

4.19 Bemerkung. Die Gleichung $(*)$ ist eine lineare Differentialgleichung 1. Ordnung für \dot{u} , welche die Lösung

$$\begin{aligned} \dot{u}(t) &= \dot{u}(t_0) e^{-\int_{t_0}^t \left(2 \frac{\dot{\varphi}(s)}{\varphi(s)} + a(s)\right) ds} \\ &= \dot{u}(t_0) e^{-2(\ln \varphi(t) - \ln \varphi(t_0))} e^{-\int_{t_0}^t a(s) ds} \\ &= \dot{u}(t_0) \frac{\varphi(t_0)^2}{\varphi(t)^2} e^{-\int_{t_0}^t a(s) ds} \end{aligned}$$

hat. Man erhält u dann durch eine weitere Integration.

Beweis. von Satz 4.18.

$$\psi = \varphi u \quad \Rightarrow \quad \dot{\psi} = \dot{\varphi} u + \varphi \dot{u} \quad \Rightarrow \quad \ddot{\psi} = \ddot{\varphi} u + 2\dot{\varphi} \dot{u} + \varphi \ddot{u}$$

Also

$$\ddot{\psi} + a \dot{\psi} + b \psi = \ddot{\varphi} u + 2\dot{\varphi} \dot{u} + \varphi \ddot{u} + a \dot{\varphi} u + a \varphi \dot{u} + b \varphi u = 2\dot{\varphi} \dot{u} + \varphi \ddot{u} + a \varphi \dot{u},$$

da $\ddot{\varphi} + a \dot{\varphi} + b \varphi = 0$. Somit löst ψ die Differentialgleichung, wenn $\varphi \ddot{u} + 2\dot{\varphi} \dot{u} + a \varphi \dot{u} = 0$ bzw. wenn

$$\ddot{u} + \left(2 \frac{\dot{\varphi}}{\varphi} + a\right) \dot{u} = 0.$$

Ist u nicht konstant, so sind $\psi = u\varphi$ und φ auf J linear unabhängig. □

4.20 Beispiel. Für $n = 1$ ist die Legendresche Differentialgleichung auf $I = (-1, 1)$

$$\ddot{x} - \frac{2t}{1-t^2} \dot{x} + \frac{2}{1-t^2} x = 0.$$

Sie hat die Lösung $P_1(t) = t$. Also erhält man auf $(0, 1)$ eine zweite Lösung durch den Ansatz $\psi(t) = t u(t)$, wobei

$$\begin{aligned} \dot{u}(t) &= \dot{u}(t_0) \frac{t_0^2}{t^2} e^{\int_{t_0}^t \frac{2s}{1-s^2} ds} = \dot{u}(t_0) \frac{t_0^2}{t^2} e^{-\ln(1-s^2)|_{t_0}^t} = \dot{u}(t_0) \frac{t_0^2}{t^2} \frac{1-t_0^2}{1-t^2} \\ &= \dot{u}(t_0) t_0^2 (1-t_0^2) \left(\frac{1}{t^2} + \frac{1}{2} \left(\frac{1}{1+t} + \frac{1}{1-t} \right) \right) \end{aligned}$$

somit ist

$$\begin{aligned} u(t) &= u(t_0) + \int_{t_0}^t \dot{u}(s) ds = u(t_0) + \dot{u}(t_0) t_0^2 (1-t_0^2) \left(-\frac{1}{s} \Big|_{t_0}^t + \frac{1}{2} (\ln(1+s) - \ln(1-s)) \Big|_{t_0}^t \right) \\ &= u(t_0) + \dot{u}(t_0) t_0^2 (1-t_0^2) \left(\frac{1}{t_0} - \frac{1}{t} + \frac{1}{2} \ln \frac{1+t}{1-t} \right). \end{aligned}$$

Wir wählen nun für beliebiges $t_0 \in (0, 1)$

$$\dot{u}(t_0) = \frac{1}{t_0^2(1-t_0^2)} \quad \text{und} \quad u(t_0) = -\dot{u}(t_0)t_0(1-t_0^2)$$

und erhalten so die Lösung

$$u(t) = \frac{1}{2} \ln \frac{1+t}{1-t} - \frac{1}{t}.$$

Es ist also

$$\psi(t) = t u(t) = \frac{t}{2} \ln \frac{1+t}{1-t} - 1$$

eine von $\varphi(t)$ linear unabhängige Lösung auf dem Intervall $(0, 1)$. Man rechnet nun aber direkt nach, dass $\psi(t)$ die Differentialgleichung auf dem ganzen Intervall $(-1, 1)$ löst.

4.21 Bemerkung. Lineare Differentialgleichungen mit konstanten Koeffizienten

Eine homogene lineare Differentialgleichung m -ter Ordnung ist von der Form

$$\sum_{j=0}^m a_j x^{(j)}(t) = 0 \quad \text{mit} \quad a_m = 1, \quad (*)$$

für $x : \mathbb{R} \rightarrow \mathbb{R}$ oder $x : \mathbb{R} \rightarrow \mathbb{C}$. Sie ist äquivalent zu dem System erster Ordnung

$$\dot{y}(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ -a_0 & -a_1 & \cdots & -a_{m-2} & -a_{m-1} \end{pmatrix} y(t) =: A y(t)$$

für $y : \mathbb{R} \rightarrow \mathbb{R}^m$ oder $y : \mathbb{R} \rightarrow \mathbb{C}^m$. Jeder Eigenvektor v von A zum Eigenwert λ liefert eine Lösung $y(t) = e^{\lambda t} v$ des Systems erster Ordnung. Das charakteristische Polynom von A ist

$$P_A(\lambda) = \sum_{j=0}^m a_j \lambda^j,$$

läßt sich also direkt aus der Differentialgleichung $(*)$ ablesen. Damit haben wir ein Rezept zum Auffinden von Lösungen von $(*)$ gefunden: zu jeder Nullstelle λ_0 von $P_A(\lambda)$ ist

$$x_{\lambda_0}(t) = e^{\lambda_0 t}$$

eine Lösung von $(*)$. Hat $P_A(\lambda)$ tatsächlich m verschiedene Nullstellen, so bilden die zugehörigen Lösungen ein Fundamentalsystem. Aber auch beim Vorliegen einer ℓ -fachen Nullstelle kann man direkt ℓ linear unabhängige Lösungen angeben: sei λ_0 ℓ -fache Nullstelle von $P_A(\lambda)$, dann sind

$$x_{\lambda_0,0}(t) = e^{\lambda_0 t}, \quad x_{\lambda_0,1}(t) = t e^{\lambda_0 t}, \quad x_{\lambda_0,2}(t) = t^2 e^{\lambda_0 t}, \quad \dots, \quad x_{\lambda_0,\ell-1}(t) = t^{\ell-1} e^{\lambda_0 t}$$

ℓ linear unabhängige Lösungen von $(*)$. (Beweis in den Übungen).

4.22 Beispiel. Der gedämpfte harmonische Oszillator

Das charakteristische Polynom zur Differentialgleichung

$$\ddot{x} + 2\gamma\dot{x} + \omega^2 x = 0, \quad \gamma, \omega > 0,$$

hat die Form $P(\lambda) = \lambda^2 + 2\gamma\lambda + \omega^2$ und somit die Nullstellen

$$\lambda_{\pm} = -\gamma \pm \sqrt{\gamma^2 - \omega^2}.$$

Wir unterscheiden wieder drei Fälle:

Überdämpfte Bewegung: $\gamma > \omega$, also $\lambda_{p,m} \in \mathbb{R}$ und $\lambda_+ \neq \lambda_-$. Zwei linear unabhängige Lösungen sind somit

$$x_1(t) = e^{\lambda_+ t} = e^{(-\gamma + \sqrt{\gamma^2 - \omega^2})t} \quad \text{und} \quad x_2(t) = e^{\lambda_- t} = e^{(-\gamma - \sqrt{\gamma^2 - \omega^2})t}$$

Kritische Dämpfung: $\gamma = \omega$, also $\lambda_* = -\gamma$ zweifacher Eigenwert. Zwei linear unabhängige Lösungen sind diesmal

$$x_1(t) = e^{\lambda_* t} = e^{-\gamma t} \quad \text{und} \quad x_2(t) = t e^{\lambda_* t} = t e^{-\gamma t}.$$

Gedämpfte Schwingung: $\gamma < \omega$, also $\lambda_{\pm} \in \mathbb{C}$ und $\lambda_+ \neq \lambda_-$. Zwei linear unabhängige Lösungen sind

$$x_1(t) = e^{\lambda_+ t} = e^{-\gamma t} e^{i\tilde{\omega}t} \quad \text{und} \quad x_2(t) = e^{\lambda_- t} = e^{-\gamma t} e^{-i\tilde{\omega}t} \quad \text{mit} \quad \tilde{\omega} := \sqrt{\omega^2 - \gamma^2}.$$

In allen drei Fällen läßt sich jede Lösung als Linearkombination von $x_1(t)$ und $x_2(t)$ schreiben.

4.1 Qualitative Theorie autonomer Systeme

Es seien F, G stetig partiell differenzierbar.

$$\begin{aligned} \dot{x} &= F(x, y) \\ \dot{y} &= G(x, y) \end{aligned} \tag{4.1}$$

(x_0, y_0) sei stationärer Punkt (Gleichgewichtspunkt), das heißt

$$F(x_0, y_0) = G(x_0, y_0) = 0$$

Es ist dann

$$\frac{F(x, y)}{G(x, y)} = \frac{\dot{x}}{\dot{y}} = \frac{\frac{dx}{dt}}{\frac{dy}{dt}} = \frac{dx}{dy}$$

oder

$$\frac{G(x, y)}{F(x, y)} = \frac{dy}{dx}$$

falls entweder $F(x, y) \neq 0$ oder $G(x, y) \neq 0$ in Umgebungen von (x_0, y_0) .

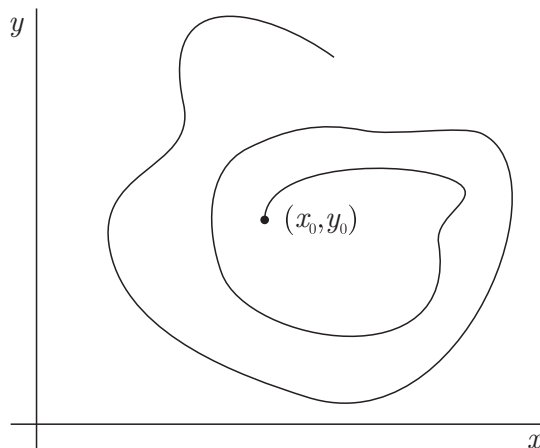


Abbildung 4.1: Trajektorie eines Teilchens

4.1.1 Definition. stabiler stationärer Punkt

Ein stationärer Punkt (x_0, y_0) heißt STABIL, wenn es zu jedem $\epsilon > 0$ ein δ gibt mit folgender Eigenschaft: Gilt für eine Lösung $(x(t), y(t))$ von (4.1) und für ein \bar{t} , dass $(x(\bar{t}), y(\bar{t})) \in B_\delta(x_0, y_0)$ so ist $(x(t), y(t)) \in B_\epsilon(x_0, y_0) \forall t > \bar{t}$.

(x_0, y_0) ist ASYMPTOTISCH STABIL, wenn er stabil ist und es ein $R > 0$ gibt mit der Eigenschaft: Gilt für ein t , dass $(x(t), y(t)) \in B_R(x_0, y_0)$ so gilt

$$(x(t), y(t)) \rightarrow (x_0, y_0) \text{ für } t \rightarrow \infty$$

Siehe zur Veranschaulichung Abbildung 4.2.

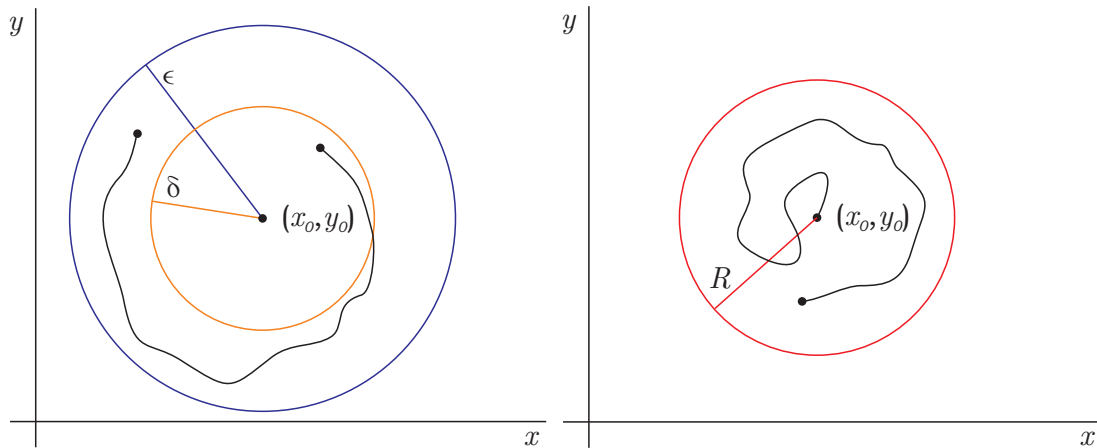


Abbildung 4.2: stabiler (links) und asymptotisch stabiler (rechts) stationärer Punkt

4.1.0.4 Beispiel. harmonischer Oszillator

Die Differentialgleichung des harmonischen Oszillators ist

$$\ddot{x} + \omega^2 x = 0$$

Transformiere diese auf ein System 1. Ordnung.

$$\dot{x} = y$$

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \cos \omega t + \frac{y_0}{\omega} \sin \omega t \\ -x_0 \omega \sin \omega t + y_0 \cos \omega t \end{pmatrix}$$

Suche Eigenwerte:

$$\det \left(\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} - \lambda \text{Id} \right) = 0$$

$$\lambda^2 + \omega^2 = 0$$

$$\Rightarrow \lambda_{1/2} = \pm i\omega$$

Suche Eigenvektoren (Eigenvektor ist immer ein Vielfaches von $(1, a)$ oder $(0, 1)$):

$$\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ a \end{pmatrix} = i\omega \begin{pmatrix} 1 \\ a \end{pmatrix}$$

$$\Leftrightarrow a = i\omega$$

$$\lambda_1 = i\omega, c_1 = \begin{pmatrix} 1 \\ i\omega \end{pmatrix}$$

Ist der erste Eigenwert komplex, ist der zweite sein konjugiert Komplexes:

$$\lambda_2 = -i\omega, c_2 = \begin{pmatrix} 1 \\ -i\omega \end{pmatrix}$$

Die Lösung des linearen Systems ist:

$$e^{\lambda_1 t} c_1 = e^{i\omega t} \begin{pmatrix} 1 \\ i\omega \end{pmatrix}$$

Die Lösung setzt sich zusammen aus Real- und Imaginärteil:

$$\varphi_1 = \operatorname{Re} \left(e^{i\omega t} \begin{pmatrix} 1 \\ i\omega \end{pmatrix} \right) = \begin{pmatrix} \cos \omega t \\ -\omega \sin \omega t \end{pmatrix}$$

$$\varphi_2 = \operatorname{Im} \left(e^{i\omega t} \begin{pmatrix} 1 \\ i\omega \end{pmatrix} \right) = \begin{pmatrix} \sin \omega t \\ \omega \cos \omega t \end{pmatrix}$$

$$\vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \mu_1 \varphi_1(t) + \mu_2 \varphi_2(t)$$

So findet man

$$x(t) = \mu_1 \cos \omega t + \mu_2 \sin \omega t$$

$$y(t) = \dot{x}(t)$$

Die Lösung lässt sich schöner schreiben, klammert man $\sqrt{\mu_1^2 + \mu_2^2}$ aus:

$$x(t) = \sqrt{\mu_1^2 + \mu_2^2} \left(\frac{\mu_1}{\sqrt{\mu_1^2 + \mu_2^2}} \cos \omega t + \frac{\mu_2}{\sqrt{\mu_1^2 + \mu_2^2}} \sin \omega t \right)$$

Wir nennen die Brüche a und b ,

$$x(t) = \sqrt{\mu_1^2 + \mu_2^2} (a \cos \omega t + b \sin \omega t)$$

Wegen $a^2 + b^2 = 1$ kann man auch hier trigonometrische Funktionen benutzen:

$$a = \cos \phi$$

$$b = \sin \phi$$

Mit Formeln für das Produkt von Winkelfunktionen ergibt sich:

$$x(t) = \sqrt{\mu_1^2 + \mu_2^2} \cos(\omega t - \phi)$$

$$y(t) = -\sqrt{\mu_1^2 + \mu_2^2} \omega \sin(\omega t - \phi)$$

Im Phasenraum ergibt das Ellipsen, vgl. Abbildung 4.3.

4.1.0.5 Beispiel. gedämpfter harmonischer Oszillator

Die Differentialgleichung des gedämpften harmonischen Oszillators (vgl. Abbildung 4.4) ist

$$\ddot{x} + 2\rho\dot{x} + \omega^2 x = 0$$

Transformiere diese auf ein System 1. Ordnung. Sei dazu $\dot{x} = y$ und $\dot{y} = \ddot{x} = -2\rho y - \omega^2 x$. Das System lautet dann

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\rho \end{pmatrix}}_{:=A} \begin{pmatrix} x \\ y \end{pmatrix}$$

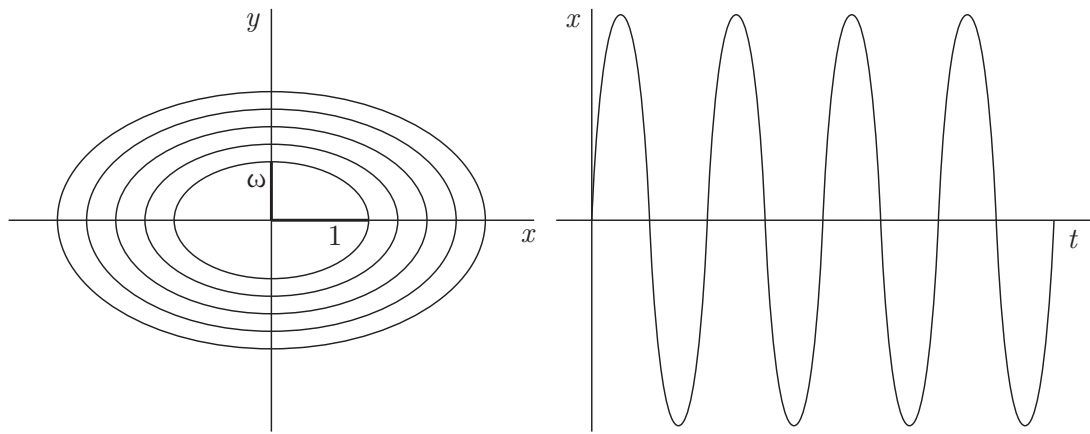
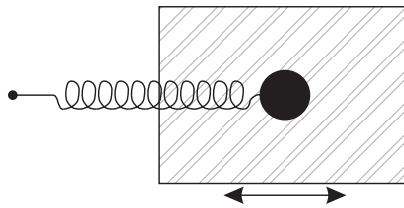
Abbildung 4.3: Phasenraum und x - t -Diagramm eines ungedämpften harmonischen Oszillators

Abbildung 4.4: Gedämpfter harmonischer Oszillator

Berechne nun die Eigenwerte des Problems mithilfe des charakteristischen Polynoms:

$$\begin{aligned} \det(A - \lambda \text{Id}) &= \det \begin{pmatrix} -\lambda & 1 \\ -\omega^2 & -2\rho - \lambda \end{pmatrix} \\ &= \lambda(2\rho + \lambda) + \omega^2 \\ &\stackrel{!}{=} 0 \end{aligned}$$

Auflösen der Gleichung nach λ ergibt

$$\lambda_{1,2} = -\rho \pm \sqrt{\rho^2 - \omega^2}$$

Betrachte nun den Fall, dass $\rho > \omega$, was einer starken Dämpfung entspricht. Die Eigenwerte $\lambda_{1,2}$ sind dann reell und negativ. Bestimme die zugehörigen Eigenvektoren $\varphi_{1,2}$:

$$\begin{aligned} A\varphi_1 &= \lambda_1\varphi_1 \\ \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\rho \end{pmatrix} \begin{pmatrix} 1 \\ a \end{pmatrix} &= \lambda_1 \begin{pmatrix} 1 \\ a \end{pmatrix} \\ \Rightarrow a &= \lambda_1 \end{aligned}$$

Also sind die Eigenvektoren

$$\varphi_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad \varphi_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$$

Die allgemeine Lösung des linearen homogenen Systems ist

$$\begin{aligned} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} &= \mu_1 e^{\lambda_1 t} \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} + \mu_2 e^{\lambda_2 t} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \\ &= \begin{pmatrix} \mu_1 e^{\lambda_1 t} + \mu_2 e^{\lambda_2 t} \\ \mu_1 \lambda_1 e^{\lambda_1 t} + \mu_2 \lambda_2 e^{\lambda_2 t} \end{pmatrix} \end{aligned}$$

Um das Verhalten im Phasenraum zu bestimmen sei nun o.B.d.A. $\lambda_1 > \lambda_2$. Dann ist $\lambda_2 - \lambda_1 < 0$. Forme die Lösung entsprechend um:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = e^{\lambda_1 t} \left(\mu_1 \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} + \mu_2 e^{(\lambda_2 - \lambda_1)t} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right)$$

Verwende nun vom Anfang des Kapitels, dass $\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}}$.

$$\begin{aligned} \frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} &= \frac{\mu_1 e^{\lambda_1 t} \lambda_1^2 + \mu_2 e^{\lambda_2 t} \lambda_2^2}{\mu_1 e^{\lambda_1 t} \lambda_1 + \mu_2 e^{\lambda_2 t} \lambda_2} \\ &= \frac{\mu_1 \lambda_1^2 + \mu_2 \lambda_2^2 e^{(\lambda_2 - \lambda_1)t}}{\mu_1 \lambda_1 + \mu_2 \lambda_2 e^{(\lambda_2 - \lambda_1)t}} \xrightarrow{t \rightarrow \infty} \lambda_1 \end{aligned}$$

Also ist für $t \rightarrow \infty$ $dy = \lambda_1 dx$ und damit $\sim y = \lambda_1 x$. Die Trajektorien im Phasenraum sind in Abbildung 4.5 dargestellt. Für den Fall, dass $\rho < \omega$ (schwache Dämpfung) sind die Eigenwerte komplex konjugiert:

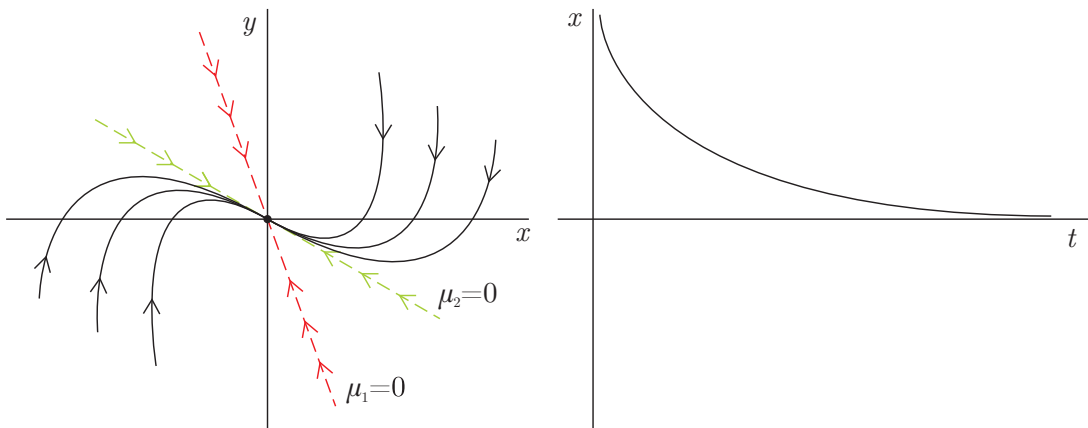


Abbildung 4.5: Phasenraum und x - t -Diagramm eines stark gedämpften harmonischen Oszillators

$$\lambda_{1,2} = -\rho \pm i\sqrt{\omega^2 - \rho^2} = -\rho \pm i\omega_1$$

Die Eigenvektoren und allgemeinen Lösungen sind (analog zu oben)

$$\text{Eigenvektor: } a_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}$$

$$a_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$$

$$\text{Lösung: } \varphi_1 = \left\{ \text{Re} \left(e^{i\lambda_1 t} \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \right), \text{Im} \left(e^{i\lambda_1 t} \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \right) \right\}$$

$$\varphi_2 = \left\{ \text{Re} \left(e^{i\lambda_2 t} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right), \text{Im} \left(e^{i\lambda_2 t} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \right) \right\}$$

Man findet also für die Lösung in $x(t)$ und $y(t)$:

$$x(t) = e^{-\rho t} (\mu_1 \cos \omega_1 t + \mu_2 \sin \omega_1 t)$$

$$y(t) = e^{-\rho t} ((\mu_2 \omega_1 - \mu_1 \rho) \cos \omega_1 t - (\mu_2 \rho + \mu_1 \omega_1) \sin \omega_1 t)$$

Schreibe $x(t)$ in Polarkoordinaten:

$$\begin{aligned} x(t) &= e^{-\rho t} \sqrt{\mu_1^2 + \mu_2^2} (\cos \phi \cos \omega_1 t + \sin \phi \sin \omega_1 t) \\ &= e^{-\rho t} \sqrt{\mu_1^2 + \mu_2^2} \cos(\phi - \omega_1 t) \end{aligned}$$

Macht man dasselbe für $y(t)$, erkennt man, dass $(x(t), y(t))$ im Phasenraum beschränkt ist:

$$|(x(t), y(t))| \leq e^{-\rho t}$$

Für eine Skizze siehe Abbildung 4.6.

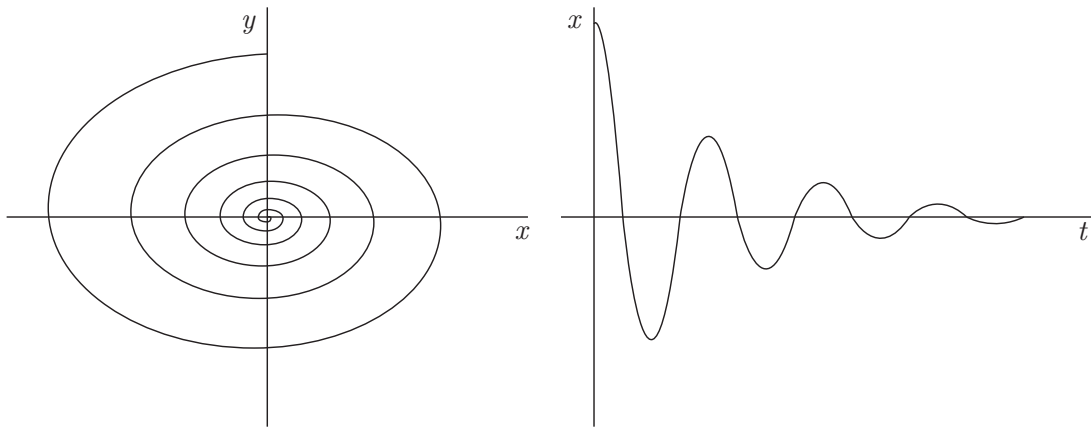


Abbildung 4.6: Phasenraum und x - t -Diagramm eines schwach gedämpften harmonischen Oszillators

4.23 Beispiel. Schließlich können wir nun den Fall behandeln, dass $\omega = \rho$ ist. Ziel war es, die lineare Differentialgleichung

$$\begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\rho \end{pmatrix}}_{=:A} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix}$$

für den Fall $\gamma = \omega_0$ zu lösen, in welchem A nicht diagonalisierbar ist. Es hat dann

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

den einzigen Eigenwert $\lambda = -\rho$ mit dem eindimensionalen Eigenraum E_λ aufgespannt von

$$w = \begin{pmatrix} 1 \\ -\rho \end{pmatrix}.$$

Den fehlenden Hauptvektor erhält man aus $(A - \lambda)v = w$, also

$$\begin{pmatrix} \rho & 1 \\ -\rho^2 & -\rho \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -\rho \end{pmatrix} \Rightarrow v = \begin{pmatrix} 1 \\ 1 - \rho \end{pmatrix}.$$

Die Transformationsmatrix auf Jordansche Normalform ist also

$$Q = (w, v) = \begin{pmatrix} 1 & 1 \\ -\rho & 1 - \rho \end{pmatrix}, \quad Q^{-1} = \frac{1}{\det Q} \begin{pmatrix} 1 - \rho & -1 \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} 1 - \rho & -1 \\ \rho & 1 \end{pmatrix}.$$

Damit finden wir

$$\begin{aligned} e^{At} &= Qe^{Jt}Q^{-1} = Q \begin{pmatrix} e^{-\rho t} & te^{-\rho t} \\ 0 & e^{-\rho t} \end{pmatrix} Q^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ -\rho & 1 - \rho \end{pmatrix} \begin{pmatrix} e^{-\rho t} & te^{-\rho t} \\ 0 & e^{-\rho t} \end{pmatrix} \begin{pmatrix} 1 - \rho & -1 \\ \rho & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{-\rho t} + \gamma t e^{-\rho t} & te^{-\rho t} \\ \dots & \dots \end{pmatrix}. \end{aligned}$$

Also ist

$$x(t) = (e^{-\rho t} + \rho t e^{-\rho t}) x(0) + te^{-\rho t} v(0).$$

Stabilität linearer Systeme**A diagonalisierbar**

Betrachte nur Probleme in \mathbb{R}^2 mit

$$\dot{x} = Ax \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

wobei A ähnlich einer Diagonalmatrix D ist, das heißt

$$A = TDT^{-1} \quad (T^{-1}AT = D)$$

dabei hat D die Form

$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

und $\lambda_{1,2}$ sind Eigenwerte. Diese berechnen sich wie folgt:

$$\begin{aligned} \lambda_{1,2} &= \frac{s}{2} \pm \sqrt{\frac{s^2}{4} - \Delta} \\ s &= \text{Spur } A = a_{11} + a_{22} \\ \Delta &= \det A = \lambda_1 \lambda_2 \end{aligned}$$

Bei Δ wurde der Multiplikationssatz von Determinanten verwendet:

$$\begin{aligned} \det A &= \det(TDT^{-1}) \\ &= \det T \det(T^{-1}) \det D \\ &= \det T (\det T)^{-1} \det D \\ &= \det D \end{aligned}$$

Unterscheide nun zwischen der transformierten Basis der Eigenwerte y und der Basis der Abbildung x , es ist hier

$$y = T^{-1}x \quad x = Ty$$

Die lineare Differentialgleichung lautet dann in der transformierten Basis

$$\begin{aligned} \dot{y} &= T^{-1}\dot{x} \\ &= T^{-1}Ax \\ &= T^{-1}ATy \\ &= Dy \end{aligned}$$

Die Eigenvektoren zu den Eigenwerten $\lambda_{1,2}$ sind $\lambda_1 e_1$ und $\lambda_2 e_2$. Die Allgemeine Lösung ist dann

$$y(t) = \mu_1 e^{\lambda_1 t} e_1 + \mu_2 e^{\lambda_2 t} e_2 = e^{tD} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

Wie sieht nun die transformierende Abbildung T aus? Betrachte zur Bestimmung die Bilder der Einheitsvektoren:

$$Te_1 = c_1 \quad Te_2 = c_2$$

Es ist dann

$$Ac_1 = ATe_1 = T \underbrace{T^{-1}AT}_{=D} e_1 = TDe_1 = \lambda_1 Te_1$$

Damit ist

$$T = (c_1, c_2)$$

Also ist die allgemeine Lösung in der Basis von A :

$$x(t) = Ty(t) = \mu_1 e^{\lambda_1 t} c_1 + \mu_2 e^{\lambda_2 t} c_2$$

Mit der Anfangsbedingung $x(0) = x_0$ lassen sich die Koeffizienten $\mu_{1,2}$ bestimmen:

$$\begin{aligned} \mu_1 c_1 + \mu_2 c_2 &= x_0 \\ T \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} &= x_0 \\ \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} &= T^{-1} x_0 \end{aligned}$$

Es folgt damit für $x(t)$:

$$\begin{aligned} x(t) &= \left(e^{\lambda_1 t} c_1, e^{\lambda_2 t} c_2 \right) T^{-1} x_0 \\ &= T e^{tD} T^{-1} x_0 \end{aligned}$$

da

$$\left(e^{\lambda_1 t} c_1, e^{\lambda_2 t} c_2 \right) = T e^{tD} = (c_1, c_2) \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}$$

4.1.0.6 Bemerkung. Lösungspropagator

Die Lösung $x(t)$ kann geschrieben werden als

$$x(t) = T e^{tD} T^{-1} x_0 = \phi(t) x_0$$

wobei $\phi(t)$ der Lösungspropagator ist. Es gilt: $\phi(0) = \text{Id}$. Es sei bemerkt, dass der Lösungspropagator auch anders gefunden werden kann

$$M(t) = \left(e^{\lambda_1 t} c_1, e^{\lambda_2 t} c_2 \right)$$

Aus $M(t)$ lässt sich der Lösungspropagator mittels $\phi(t) = M(t)M(0)^{-1}$ gewinnen.

A nicht diagonalisierbar

Wenn A nicht diagonalisierbar ist, dann wissen wir dass A einen Eigenwert λ haben muss wobei die geometrische Vielfachheit gleich eins ist. Also ist A ähnlich zu Jordanmatrix

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

also gibt es reguläre Matrix P , sodass

$$P^{-1} A P = J$$

ist. Und $P = (c_1, c_2)$ wobei

$$A c_1 = \lambda c_1, \quad (A - 1) c_2 = c_1.$$

Wir wissen nun dass eine allgemeine Lösung geschrieben werden kann als

$$x(t) = e^{tA} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = P e^{tJ} P^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = P \begin{pmatrix} e^{t\lambda} & t e^{t\lambda} \\ 0 & e^{t\lambda} \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$$

mit

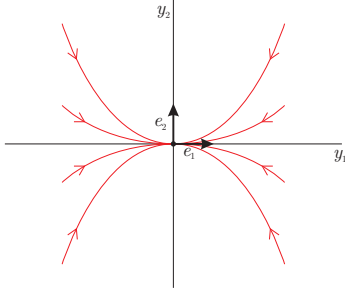
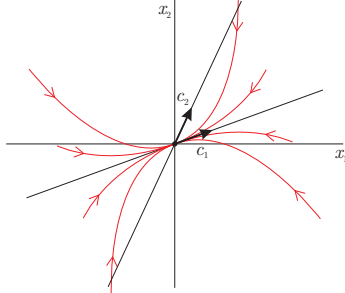
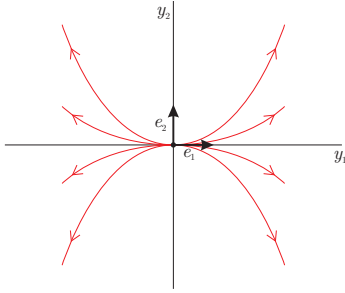
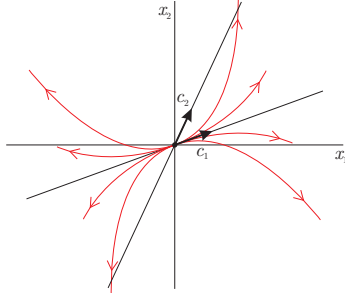
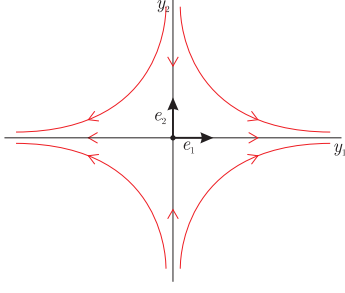
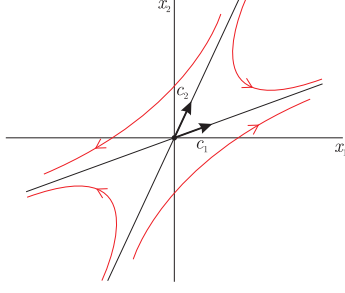
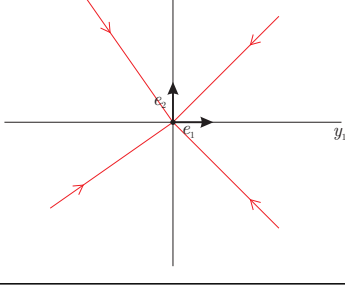
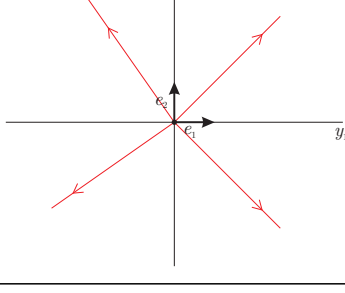
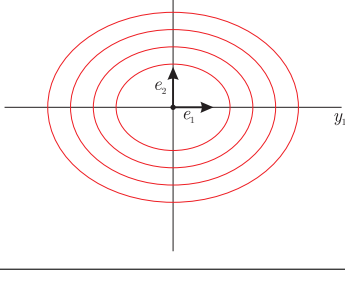
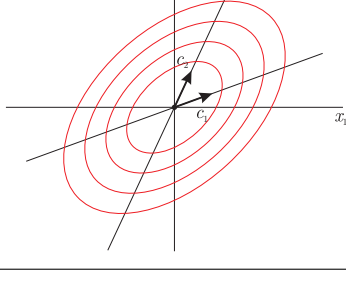
$$\begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = P^{-1} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Also

$$x(t) = c_1 (e^{t\lambda} a_0 + t e^{t\lambda} b_0) + c_2 e^{t\lambda} b_0.$$

Damit ist klar dass das System stabil ist wenn $\lambda < 0$ und instabil wenn $\lambda \geq 0$.

Betrachte die folgende Tabelle für eine Übersicht:

Eigenwerte		Verhalten um (0, 0)	
Vorzeichen	Art		
$\lambda_1, \lambda_2 < 0$	reell		
$\lambda_1, \lambda_2 > 0$	reell		
$\lambda_1 > 0, \lambda_2 < 0$	reell		
$\lambda_1 = \lambda_2 < 0$ (links) $\lambda_1 = \lambda_2 > 0$ (rechts)	reell		
$\lambda_{1,2} = \pm i\beta$	rein komplex		

Eigenwerte		Verhalten um $(0, 0)$
Vorzeichen	Art	
$\lambda_{1,2} = \alpha \pm i\beta$ $\alpha < 0$ (links) $\alpha > 0$ (rechts)	konjugiert komplex	

4.1.0.7 Satz.

Sei $\dot{x} = Ax$, $\dot{x} = F(x, y)$ $\dot{y} = G(x, y)$. Der Punkt $(0, 0)$ ist asymptotisch stabil, wenn jeder Eigenwert von A einen negativen Realteil hat (falls $A \in M(n \times n)$).

4.1.0.8 Beispiel. gedämpfter harmonischer Oszillator

$$m\ddot{x} + k\dot{x} + kx = 0$$

$$y = \dot{x}$$

$$\dot{y} = \ddot{x} = -\frac{r}{m}y - \frac{k}{m}x$$

$$\frac{d}{dt} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{k}{m} & -\frac{r}{m} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Die Energie E des Systems ist gegeben durch

$$E = \frac{m}{2}y^2 + \frac{k}{2}x^2$$

Und deren zeitliche Änderung ist

$$\begin{aligned} \frac{d}{dt}E(x(t), y(t)) &= \frac{\partial E}{\partial x}\dot{x} + \frac{\partial E}{\partial y}\dot{y} \\ &= kxy + my \left(-\frac{r}{m}y - \frac{k}{m}x \right) \\ &= -ry^2 \leq 0 \end{aligned}$$

Die Energie nimmt also entlang Lösungen ab (oder bleibt konstant). Daher ist das System stabil. Zusammengefasst muss für die Energie eines stabilen Systems gelten:

$$\begin{aligned} E(0, 0) &= 0 \\ E(x, y) &\neq 0 \quad \forall (x, y) \neq (0, 0) \\ \frac{d}{dt}E &\leq 0 \end{aligned}$$

Die Ljapunoff-Methode

Sei $(0, 0)$ ein isolierter stationärer Punkt der Gleichung (4.1). $E(x, y)$ heißt Ljapunoff-Funktion für (4.1), wenn E in einer Umgebung U von $(0, 0)$ folgende Eigenschaften besitzt:

- E ist stetig diffbar

- E verschwindet in $(0, 0)$, $E > 0$ außerhalb.
- $\frac{\partial E}{\partial x}F + \frac{\partial E}{\partial y}G$ verschwindet in $(0, 0)$ und ist ≤ 0 außerhalb.

Für < 0 im letzten Punkt ist die Funktion STRENG Ljapunoff.

4.1.0.9 Theorem.

Das System (4.1) habe $(0, 0)$ als stationären Punkt und besitze eine strenge Ljapunoff-Funktion. Dann ist $(0, 0)$ STABIL. Ist E streng Ljapunoff, so ist $(0, 0)$ ASYMPTOTISCH STABIL.

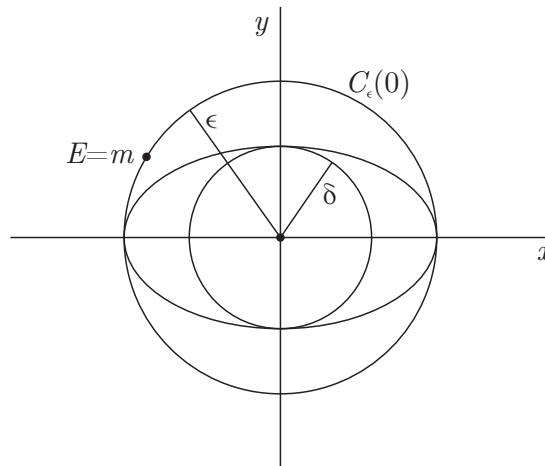


Abbildung 4.7: Stationärer Punkt

Proof. (a) Stabilität

Gebe $\epsilon > 0$ vor (vgl. Abbildung 4.7). $C_\epsilon(0)$ sei die Kreisscheibe um den Ursprung mit Radius ϵ . m sei das Minimum der Ljapunoff-Funktion auf dieser Kreisscheibe.

$$m = \min E(x, y) \quad (x, y) \in C_\epsilon(0)$$

Nun kann man ein δ finden, sodass E in einer Umgebung mit Radius δ kleiner als m bleibt.

$$\max_{B_\delta(0)} E(x, y) < m$$

Die Trajektorie $(x(t), y(t))$ sei Lösung zu (4.1), $(x(0), y(0)) \in \overline{B_\delta(0)}$ (= Abschluss von $B_\delta(0)$)

$$\begin{aligned} \frac{d}{dt} E(x(t), y(t)) &= \frac{\partial E}{\partial x} \dot{x} + \frac{\partial E}{\partial y} \dot{y} \\ &= \frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G \leq 0 \quad (\text{laut Voraussetzung}) \end{aligned}$$

Da $E(x(0), y(0)) < m$ und $\frac{d}{dt} E \leq 0 \Rightarrow E < m \quad \forall t \geq 0 \Rightarrow$ System ist stabil.

(b) asymptotische Stabilität

Sei E strenge Ljapunoff-Funktion. Dann ist $E(x(t), y(t))$ streng monoton abnehmend und positiv, es ist also

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

so konvergiert $E(x(t), y(t))$ gegen einen Grenzwert. Für $\lambda = 0$ folgt

$$(x(t), y(t)) \rightarrow (0, 0),$$

da $E \neq 0$ außerhalb $(0, 0)$. Nun ist zu zeigen, dass λ nur 0 sein kann.

Angenommen, $\lambda > 0$. Dann gibt es eine Umgebung B_ρ von $(0, 0)$ mit Radius $0 < \rho < \epsilon$, sodass

$E(x, y) < \lambda \quad \forall (x, y) \in \overline{B_\rho}$. Auf dem Kreisring zwischen B_ρ und einer Kreislinie mit Radius $\epsilon > \rho$ gilt dann

$$\frac{d}{dt}E = \frac{\partial E}{\partial x}\dot{x} + \frac{\partial E}{\partial y}\dot{y} = \frac{\partial E}{\partial x}F + \frac{\partial E}{\partial y}G \leq M \quad \forall (x, y) \in \overline{B_\rho} \setminus B_\epsilon$$

mit

$$M = \sup \left\{ \frac{\partial E}{\partial x}F + \frac{\partial E}{\partial y}G \mid (x, y) \in \overline{B_\epsilon} \setminus B_\rho \right\} < 0 \quad (\text{nach Definition})$$

So lässt sich E nach oben abschätzen:

$$E(t) = E(0) + \int_0^t \frac{d}{ds}E(s)ds \leq E(0) + tM \xrightarrow[t \rightarrow \infty]{} -\infty$$

Dies ist ein Widerspruch zu $\lambda > 0$. Aus $\lambda = 0$ folgt: $(0, 0)$ ist asymptotisch stabil. □

Linearisieren

Um die Frage nach der Stabilität eines stationären Punktes zu klären, lohnt es sich auch, (4.1) um diesen Punkt zu linearisieren. Der stationäre Punkt sei (x_0, y_0)

$$F(x_0, y_0) = 0$$

$$G(x_0, y_0) = 0$$

Die Entwicklung ist

$$\begin{aligned} \begin{pmatrix} F(x, y) \\ G(x, y) \end{pmatrix} &= \begin{pmatrix} F(x_0, y_0) \\ G(x_0, y_0) \end{pmatrix} + A \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} \\ &= A \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} \end{aligned}$$

mit

$$\frac{|f(x, y)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \rightarrow 0 \quad \text{und} \quad \frac{|g(x, y)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \rightarrow 0 \quad \text{für} \quad (x, y) \rightarrow (x_0, y_0)$$

und

$$A = D \begin{pmatrix} F \\ G \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix} \Big|_{(x_0, y_0)}$$

Gleichung (4.1) ist äquivalent zu

$$\frac{d}{dt} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix} \Big|_{(x_0, y_0)} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}$$

Denn: Setzt man $\bar{x} = x - x_0$ und $\bar{y} = y - y_0$ so ergibt sich (zusammen mit $\tilde{f}(\bar{x}, \bar{y}) = f(x - x_0, y - y_0)$ und $\tilde{g}(\bar{x}, \bar{y}) = g(x - x_0, y - y_0)$)

$$\frac{d}{dt} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = A \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} + \begin{pmatrix} \tilde{f}(\bar{x}, \bar{y}) \\ \tilde{g}(\bar{x}, \bar{y}) \end{pmatrix}$$

4.1.0.10 Satz.

Sei $(0, 0)$ stationärer Punkt von (4.1), F, G stetig differenzierbar. Sei ferner

$$A = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial y} \end{pmatrix} \Big|_{(0,0)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

4 Lineare Differentialgleichungen

- (a) Hat A ausnahmslos Eigenwerte mit negativem Realteil, so ist $(0, 0)$ isolierter stationärer Punkt und ASYMPTOTISCH STABIL.
- (b) Hat A mindestens einen Eigenwert mit positivem Realteil, so ist $(0, 0)$ INSTABIL.

Proof. Zur Isoliertheit: Angenommen es gibt eine Folge (x_n, y_n) von stationären Punkten. Schreibe diese nun in Polarkoordinaten:

$$x_n = r_n \cos \phi_n \quad y_n = r_n \sin \phi_n$$

Alle diese Punkte sind nun eine Lösung von (4.1) und zwar $(0, 0)$. Es gilt also

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= A \begin{pmatrix} r_n \cos \phi_n \\ r_n \sin \phi_n \end{pmatrix} + \begin{pmatrix} f(r_n \cos \phi_n) \\ g(r_n \sin \phi_n) \end{pmatrix} \\ &= A \begin{pmatrix} \cos \phi_n \\ \sin \phi_n \end{pmatrix} + \begin{pmatrix} \frac{f(r_n \cos \phi_n)}{r_n} \\ \frac{g(r_n \sin \phi_n)}{r_n} \end{pmatrix} \\ \underbrace{\begin{pmatrix} \cos \phi_n \\ \sin \phi_n \end{pmatrix}}_{|\cdot|=1} &= \underbrace{-A^{-1} \begin{pmatrix} \frac{f(r_n \cos \phi_n)}{r_n} \\ \frac{g(r_n \sin \phi_n)}{r_n} \end{pmatrix}}_{|\cdot| \rightarrow 0} \end{aligned}$$

Dies erzeugt einen Widerspruch. Also ist $(0, 0)$ isoliert. Zur asymptotischen Stabilität: Sei

$$E = ax^2 + bxy + cy^2$$

und $(x(t), y(t))$ eine Lösung von (4.1) Es ist dann

$$\begin{aligned} \frac{d}{dt} E(x(t), y(t)) &= \frac{\partial E}{\partial x} \dot{x} + \frac{\partial E}{\partial y} \dot{y} \\ &= \frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G \\ &= \frac{\partial E}{\partial x} (a_{11}x + a_{12}y) + \frac{\partial E}{\partial y} (a_{21}x + a_{22}y) + \frac{\partial E}{\partial x} f + \frac{\partial E}{\partial y} g \\ &= (2ax + by)(a_{11}x + a_{12}y) + (bx + 2cy)(a_{21}x + a_{22}y) + \frac{\partial E}{\partial x} f + \frac{\partial E}{\partial y} g \\ &= -(x^2 + y^2) + \frac{\partial E}{\partial x} f + \frac{\partial E}{\partial y} g \end{aligned}$$

Bestimme nun die Koeffizienten a, b und c , sei dazu $\Delta = \det A$. Es gilt für die Terme in

$$\begin{aligned} (x^2) : \quad & 2aa_{11} + a_{21}b = -1 \\ (y^2) : \quad & ba_{12} + 2ca_{22} = -1 \\ (xy) : \quad & ba_{11} + 2aa_{12} + 2a_{21}c + ba_{22} = 0 \end{aligned}$$

Mit $B = -(a_{11} + a_{22})\Delta$ ist

$$\begin{aligned} a &= \frac{a_{12}^2 + a_{21}^2 + \Delta}{2B} \\ b &= \frac{-a_{11}a_{21} + a_{12}a_{22}}{B} \\ c &= \frac{a_{11}^2 + a_{12}^2 + \Delta}{2B} \end{aligned}$$

Damit ist $E \geq 0$ außerhalb von $(0, 0)$ und $E(0, 0) = 0$. Wenn nun noch $\frac{d}{dt} E < 0$ so ist E Ljapunoff-Funktion. Stelle also $\frac{d}{dt} E$ auf und transformiere in Polarkoordinaten:

$$\begin{aligned} \frac{d}{dt} E &= \frac{\partial E}{\partial x} F + \frac{\partial E}{\partial y} G \\ &= -(x^2 + y^2) + (2ax + by)f(x, y) + (bx + 2cy)g(x, y) \\ &= -r^2 + r^2 \left[(2a \cos \phi + b \sin \phi) \frac{f(r \cos \phi, r \sin \phi)}{r} + (b \cos \phi + 2c \sin \phi) \frac{g(r \cos \phi, r \sin \phi)}{r} \right] \end{aligned}$$

Nun findet man R sodass die Terme in der eckigen Klammer kleiner als $\frac{1}{2}$ sind für $r \leq R$. Es gilt also für $r \leq R$:

$$\frac{\partial E}{\partial x}F + \frac{\partial E}{\partial y}G \leq -\frac{1}{2}r^2 < 0$$

$\Rightarrow E$ ist strenge Ljapunoff-Funktion

$\Rightarrow (0, 0)$ ist asymptotisch stabil

□