Kapitel 2

Hamiltonsche Mechanik

2.1 Gradientenflüsse

Definition 2.1. *Sei X eine offene Teilmenge des $\mathbb{R}^n$ und $H : X \to \mathbb{R}$ eine differenzierbare Funktion. Dann definiert der Gradient von H ein Vektorfeld auf X. Ein entsprechender Fluss heißt Gradientenfluss.*

Wenn die Ableitung von H lokal Lipschitzstetig ist, dann ist wegen dem Satz von Picard-Lindelöff 1.24 der Gradientenfluss eindeutig.

Lemma 2.2. *Sei $H : X \to \mathbb{R}$ auf einer offenen Menge von $\mathbb{R}^n$ einmal stetig differenzierbar. Dann ist H auf jeder Integralkurve des Gradientenflusses monoton wachsend. Wenn der Gradient von H bei $x \in X$ nicht verschwindet ist H auf der Integralkurve des Gradientenflusses durch X sogar streng monoton wachsend.*

Beweis: Der Gradient von H ist definiert als

$$\nabla H(x) = \left(\frac{\partial H(x)}{\partial x_1}, \frac{\partial H(x)}{\partial x_2}, \dots, \frac{\partial H(x)}{\partial x_n}\right).$$

Wenn also $q(t)$ eine Integralkurve des Gradientenflusses durch $x \in H$ ist, dann gilt

$$\begin{aligned} \frac{d}{dt}H(q(t)) &= \frac{\partial H(g(t))}{\partial x_1}\dot{q}_1(t) + \dots \frac{\partial H(q(t))}{\partial x_n}\dot{q}_n(t) \\ &= \left(\frac{\partial H}{\partial x_1}(q(t))\right)^2 + \dots + \left(\frac{\partial H}{\partial x_n}(q(t))\right)^2 \geq 0 \end{aligned}$$

Gleichheit kann nur gelten, wenn der Gradient von H bei $q(t)$ verschwindet. **q.e.d.**

Um das Maximum einer Funktion H zu finden, kann man einfach den Gradientenfluss mit einem beliebigen Anfangswert möglichst lange fließen lassen. Dabei wird dann der Wert von H immer größer. Wenn man Glück hat landet man in einem Maximum.

Beispiel 2.3. *Sei $H(x) = x_1^2 + \ldots + x_n^2$ das Quadrat der euklidischen Länge. Um einen Vektor minimaler Länge zu finden, benutzen wir den negativen Gradientenfluss.*

$$-\nabla H(x) = -2x.$$

Dieses Vektorfeld besitzt die Integralkurven

$$g(t) = e^{-2t}x \text{ mit } q(0) = x.$$

Also konvergieren alle Bahnen dieses Gradientenflusses gegen das einzige Minimum $x = 0$ von dem Quadrat der euklidischen Länge H.

Korollar 2.4. *Sei H auf einer offenen Teilmenge X von $\mathbb{R}^n$ zweimal stetig differenzierbar. Dann besitzt der entsprechende lokale Fluss außer den Fixpunkten keine weiteren periodischen Orbits. Die Fixpunkte bestehen aus allen kritischen Punkten von H.*

Beweis: Offenbar sind alle kritischen Punkte von H auch Fixpunkte des entsprechenden Gradientenflusses. Wenn x kein kritischer Punkt von H ist, dann nimmt H auf der Integralkurve durch x nur noch Werte größer als $H(x)$ an und kann nicht mehr zu x zurückkehren. **q.e.d.**

Lemma 2.5. *Sei $F : X \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge X des $\mathbb{R}^n$. Wenn X konvex ist oder einfach zusammenhängend, dann ist F genau dann ein Gradientenvektorfeld, wenn für alle $1 \leq i < j \leq n$ gilt*

$$\frac{\partial F_i(x)}{\partial x_j} = \frac{\partial F_j(x)}{\partial x_i}.$$

Beweis: Sei $x_0 \in X$ ein beliebiger Punkt. Wir definieren für jeden Punkt x die Funktion

$$H : X \to \mathbb{R}, \quad x \mapsto H(x) = \int_0^1 (x - x_0) \cdot F(x_0 + t(x - x_0))dt.$$

Dann folgt

$$\begin{aligned}
\frac{\partial H}{\partial x_i} &= \int_0^1 F_i(x_0 + t(x - x_0))dt + \int_0^1 (x - x_0)t\frac{\partial F}{\partial x_i}(x_0 + t(x - x_0))dt \\
&= \int_0^1 F_i(x_0 + t(x - x_0))dt + \int_0^1 \sum_{j=1}^n (x - x_0)_j t\frac{\partial F_i}{\partial x_j}(x_0 + t(x - x_0))dt \\
&= \int_0^1 \frac{d}{dt}\Big(tF_i(x_0 + t(x - x_0))\Big)dt = F_i(x_0 + x - x_0) = F_i(x).
\end{aligned}$$

Also ist F das Gradientenvektorfeld von H. Umgekehrt erfüllt das Gradientenvektorfeld von H

$$\frac{\partial^2 H(x)}{\partial x_i \partial x_j} = \frac{\partial^2 H(x)}{\partial x_j \partial x_i} \qquad \textbf{q.e.d.}$$

Übungsaufgabe 2.6. *Sei H eine stetig differenzierbare Funktion auf dem $\mathbb{R}^n$, die nach unten beschränkt ist. Für ein $x_0 \in \mathbb{R}^n$ sei $t \mapsto q(t)$ die Integralkurve von dem negativen Gradientenvektorfeld von H durch den Punkt x_0. Zeige folgende Aussagen*

(i) *Es gibt eine Konstante $C > 0$, so dass $\|q(t_2) - q(t_1)\| \leq C\sqrt{t_2 - t_1}$ für alle $0 \leq t_1 \leq t_2$ im maximalen Definitionsbereich der Integralkurve q gilt.*

Hinweis: benutze den Schrankensatz, die Cauchy-Schwarzsche Ungleichung, und dass das Integral $H(q(t_2)) - H(q(t_1)) = \int_{t_1}^{t_2} \frac{dH(q(s))}{ds} ds$ beschränkt ist.

(ii) *Folgere aus (i), dass die Integralkurven durch alle $x \in \mathbb{R}^n$ für alle $t \in [0, \infty)$ definiert sind.*

(iii) *Zeige dass jeder Grenzwert $x \in \mathbb{R}^n$ von den Folgen $(q(t_n))_{n\in\mathbb{N}}$ mit streng monoton wachsenden unbeschrňkten Folgen $(t_n)_{n\in\mathbb{N}}$ in $[0, \infty)$ ein kritischer Punkt von H ist mit $H(x) = \inf\{H(q(t)) \mid t \in [0, \infty)\}$.*

Hinweis: Zeige mit Hilfe von Arzela-Ascoli dass eine Teilfolge von der Folge von Funktionen $(q_n)_{n\in\mathbb{N}}$ mit $q_n(t) = q(t + t_n)$ auf allen kompakten Teilmengen von $[0, \infty)$ gegen eine Integralkurve des negativen Gradientenvektorfeldes konvergiert.

(iv) *Sei $x \in \mathbb{R}^n$ ein Grenzwert einer Folge $(q(t_n))_{n\in\mathbb{N}}$ wie in (iii). Für ein $\delta > 0$ sei $W = \{t \in [0, \infty) \mid q(t) \in B(x, \delta)\}$ und $\int_W \|\dot{q}(t)\| dt < \infty$. Dann konvergiert q im Grenzwert $t \to \infty$ gegen x.*

Hinweis: Betrachte die Folge $(\tilde{t}_n)_{n\in\mathbb{N}}$ mit $\tilde{t}_n = \sup\{t \in [t_n, \infty) \mid [t_n, t] \subset W\}$. Zeige, dass es ein $N \in \mathbb{N}$ mit $\tilde{t}_N = \infty$ gibt, weil sonst die Folge $(q(\tilde{t}_n))_{n\in\mathbb{N}}$ gegen x konvergiert im Widerspruch zu $\|x - q(\tilde{t}_n)\| = \delta$.

(v) *Sei x wie in (iv) und $\Phi : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion mit $\Phi' > 0$, so dass für ein $\delta > 0$ folgende sogenannte Gradientenungleichung gilt: (siehe Sen-Zhong Huang: Gradient Inequalities, AMS 2006)*

$$\frac{1}{\Phi'(H(y))} \leq \|\nabla H(y)\| \quad \text{für alle} \quad y \in B(x, \delta).$$

Dann konvergiert q im Grenzwert $t \to \infty$ gegen x.

Hinweis: Zeige $\|\dot{q}(t)\| \leq -\frac{d}{dt}\Phi(H(q(t)))$ für alle t mit $q(t) \in B(x, \delta)$.

2.2 Hamiltonsche Systeme

Um das Gradientenvektorfeld einer Funktion zu definieren haben wir implizit die euklidische Metrik benutzt. Im Allgemeinen ist die Ableitung einer differenzierbaren Funktion auf einer offenen Teilmenge X eines Banachraumes V eine lineare Abbildung von V nach $\mathbb{R}$, also ein Element des Dualraumes von V. Um den Dualraum mit V zu identifizieren benötigen wir ein Skalarprodukt. Wir können auch andere (antisymmetrische) Bilinearformen benutzen um aus Funktionen Vektorfelder zu konstruieren.

Auf dem $\mathbb{R}^{2n}$ betrachten wir die Bilinearform

$$\omega : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}, \quad (v, w) \mapsto \omega(v, w) = v^t \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} w = \sum_{i+1^n} (v_i w_{n+i} - v_{n+i} w_i).$$

Hierbei betrachten wir v und w als Spaltenvektoren in $\mathbb{R}^{2n}$. Dann induziert eine differenzierbare Funktion H das Vektorfeld $X(H) = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right) \nabla H$ mit $\omega(v, X(H)) = v \cdot \nabla H$.

Definition 2.7. *Es sei $M \subset \mathbb{R}^{2n}$ offen und $H \in C^2(M, \mathbb{R})$. Die Differentialgleichung*

$$\dot{q} = \frac{\partial H(q,p)}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H(q,p)}{\partial q_i} \qquad \text{für } i = 1, \ldots, n$$

oder, in den Koordinaten $x = (q_1, \ldots, q_n, p_1, \ldots, p_n) = (q, p)$ $\dot{x} = X(H)(x)$ mit dem Hamiltonschen Vektorfeld $X(H) := J\nabla H$ und $J = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$ heißt Hamiltonsche Differentialgleichung.

Bemerkung 2.8. *H wird Hamiltonfunktion genannt. Wir setzen im Folgenden einfachheitshalber oft voraus, dass H ein Dynamisches System definiert.*

Definition 2.9. *Eine reelle Funktion auf dem Phasenraum M eines dynamischen Systems heißt Integral der Bewegung, wenn sie auf allen Bahnen konstant ist.*

Satz 2.10. *Die Hamiltonfunktion ist ein Integral der Bewegung.*

Beweis: $\frac{d}{dt} H(\Phi(t,x)) = H'(\frac{d}{dt}\Phi(t,x)) = H'(X(H))(\Phi(t,x)) = (\nabla H \cdot J\nabla H)(\Phi(t,x))$. Es gilt aber

$$(v, Jv) = (J^t v, v) = -(Jv, v) = -(v, Jv) = 0$$

für $v \in \mathbb{R}^{2n}$, also auch für $\nabla H(\Phi(t,x))$. **q.e.d.**

Dieser Satz erlaubt uns, das dynamische System auf die (oft Energieschalen genannten) Niveaumengen $H^{-1}(\{E\})$ zu restringieren. Nach dem Satz über die implizite Funktion sind dies für reguläre Werte von E von H Untermannigfaltigkeiten von M. Dies ermöglicht es für den Fall $n = 1$, also $M = \mathbb{R}^2$, für eine gegebene Funktion H die Orbits, allerdings ohne Zeitparametrisierung, aufzufinden: Zu $x \in M$ betrachten wir die Niveaumenge $H^{-1}(\{H(x)\})$. Ist $\nabla H(x) = 0$, so besteht der Orbit nur aus x. Sonst ist in einer Umgebung von x die Niveaumenge $H^{-1}(\{H(x)\})$ eine Kurve im Phasenraum $\mathbb{R}^2$, was man durch Verwendung des impliziten Funktionensatzes sehen kann. Um den Orbit zu bekommen, dehnen wir diese Kurve nach beiden Seiten so weit wie möglich aus. Die Orientierung erhalten wir durch die Richtung, die durch Drehung des Gradienten im Uhrzeigersinn um $\pi/2$ entsteht (J entspricht einer solchen Drehung).

Beispiel 2.11. *Die Orbits des eindimensionalen harmonischen Oszillators sind konzentrische Kreise um den Ursprung:* $H(q,p) := p^2 + q^2$.

Begriffe: Mit Phasenraum $M \subset \mathbb{R}^{2n} = \mathbb{R}^n_p \times \mathbb{R}^n_q$ nennen wir

- n Zahl der Freiheitsgrade
- $q \in \mathbb{R}^n_q$ Ort und $p \in \mathbb{R}^n_p$ Impuls
- $\dot{q} = \frac{\partial H}{\partial p}$ Geschwindigkeit
- H Hamiltonfunktion oder Gesamtenergie
- Besitzt der Phasenraum die Form $M = \mathbb{R}^n_p \times N$, so heißt $N \subset \mathbb{R}^n_q$ Konfigurationsraum.

2.3 Symplektische Gruppe

Damit die Hamiltonsche Differentialgleichung linear wird, muss $H : M \to \mathbb{R}$ bis auf eine Konstante ein homogenes Polynom zweiten Grades sein, also von der Form

$$H(x) = H(0) + \frac{1}{2}(x, Ax)$$

mit einer $2n \times 2n$ Matrix A. Die Funktionen zu A und A^t stimmen überein. Deshalb können wir o.B.d.A. $A = A^t$ annehmen. Dann ergibt sich $\nabla H(x) = Ax$ und $\dot{x} = JA \cdot x$. Die Differentialgleichungen bleiben invariant, wenn wir zu H eine Konstante dazuaddieren. Das entspricht physikalisch der Tatsache, dass nicht absolute Energiewerte, sondern nur Energiedifferenzen messbar sind. Wir setzen der Einfachheit halber $H(0) = 0$. Setzen wir $U = JA \in \mathcal{M}(2n, \mathbb{R})$, so stellen wir fest, dass

$$U^t J + JU = (JA)^t J + J^2 A = AJ^t J + J^2 A = -AJ^2 + J^2 A = 0$$

gilt. Das führt uns zu folgender Definition:

Definition 2.12. *Eine $2n \times 2n$ Matrix U und der entsprechenden Endomorphismus des $\mathbb{R}^{2n}$ heißen infinitesimal symplektisch, wenn folgendes gilt:*

$$U^t J + JU = 0.$$

Da die Bedingung an U linear ist, bilden die infinitesimal symplektischen Endomorphismen einen Unterraum $sp(n) \subset \mathcal{L}(\mathbb{R}^{2n})$.

Satz 2.13. *Für $U \in sp(n)$ gilt* $\mathrm{Spur}(U) = 0$.

Beweis: $\mathrm{Spur}(U) = \mathrm{Spur}(U^t) = -\mathrm{Spur}(JUJ^{-1}) = -\mathrm{Spur}(U)$. **q.e.d.**

Korollar 2.14. *Lineare Hamiltonsche Systeme sind volumenerhaltend.*

Beweis: $\Phi(t,x) = \exp(Ut)x$, und $\det(\exp(Ut)) = \exp(\mathrm{Spur}(U)t) = 1$. **q.e.d.**

Der Fluss eines linearen hamiltonschen Systems hat nicht nur die Eigenschaft volumenerhaltend zu sein. Es gilt auch

$$(\exp(Ut))^t J \exp(Ut) = \sum_{n=0}^{\infty} (U^t t)^n J \exp(Ut) = J \sum_{n=0}^{\infty} (-Ut)^n \exp(Ut) = J \quad \text{für} \quad t \in \mathbb{R}.$$

Diese Gleichung besagt, dass ein linearer Hamiltonscher Fluss die schiefsymmetrische Bilinearform

$$\omega : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \to \mathbb{R}, \quad (v,w) \mapsto \omega(v,w) = (v, Jw)$$

invariant lässt, d.h. $\omega(\Phi(t,v), \Phi(t,w)) = \omega(v,w)$, ähnlich wie eine orthogonale Transformation das kanonische Skalarprodukt des $\mathbb{R}^n$ invariant lässt. Den mechanischen Bewegungen entspricht damit eine besondere Art der Geometrie, die wir im nächsten Abschnitt eingehender untersuchen.

2.3.1 Symplektische Geometrie

Die den mechanischen Bewegungen zugrundeliegende symplektische Geometrie besitzt gewisse Ähnlichkeiten mit der Riemannschen Geometrie. Diese werden wir im Folgenden herausarbeiten.

Definition 2.15. *Sei V ein $\mathbb{R}$-Vektorraum der Dimension $n < \infty$ und $\omega : V \times V \to \mathbb{R}$ eine Bilinearform.*

(i) *Die Transponierte ω^t von ω ist durch $\omega^t(e_1.e_2) := \omega(e_2, e_1)$ gegeben.*

(ii) *ω heißt symmetrisch, wenn $\omega^t = \omega$, schiefsymmetrisch, wenn $\omega^t = -\omega$.*

(iii) *Durch ω wird die lineare Abbildung $V \to V'$, $\quad v \mapsto \omega(v,\cdot)$ mit $\omega(v,\cdot) : V \to \mathbb{R}$, $\quad w \mapsto \omega(v,w)$ in den Dualraum V' von V induziert.*

(iv) *ω heißt nicht degeneriert, wenn $\omega(v,\cdot) = 0$ nur für $v = 0$ gilt.*

(v) *Bezüglich einer Basis $(b_1, \ldots, b_n)$ von V ist die darstellende $n \times n$ Matrix J von ω durch $(J)_{ik} := \omega(b_i, b_k)$, $\quad i,k = 1,\ldots,n$ gegeben.*

(vi) *Der Rang von ω ist der (basisunabhängige) Rang der darstellenden Matrix.*

(vii) *Ist ω eine Bilinearform auf W und $A \in \mathcal{L}(V,W)$, dann heißt die Bilinearform*

$$A^*\omega : V \times V \to \mathbb{R}, \quad A^*\omega(v,w) := \omega(Av, Aw)$$

der pull-back von ω bezüglich A.

Satz 2.16. *Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum.*

(i) *Ist ω symmetrisch mit Rang r, dann ist die darstellende Matrix von ω bezüglich einer geeigneten Basis eine Diagonalmatrix, deren diagonale Einträge entweder ± 1 sind oder 0.*

(ii) *Ist ω schiefsymmetrisch mit Rang r, so ist $r = 2m, m \in \mathbb{N}_0$ und die darstellende Matrix von ω besitzt bezüglich einer geeigneten Basis die Form*

$$J = \begin{pmatrix} 0 & \mathbb{1} & 0 \\ -\mathbb{1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

mit $m \times m$-Einheitsmatrizen $\mathbb{1}$.

Beweis: Diese Aussagen werden oft in der Vorlesung Lineare Algebra bewiesen.

(i) Es gilt die Polarisationsidentität

$$\omega(v,w) = \frac{1}{4}(\omega(v+w, v+w) - \omega(v-w, v-w)).$$

Ist also $\omega \neq 0$, dann existiert ein Vektor $\hat{v}_1$ mit $c_1 = \omega(\hat{v}_1, \hat{v}_1) \neq 0$. Setze $v_1 := \hat{v}_1 / \sqrt{| c_1 |}$. Wir betrachten den von v_1 aufgespannten eindimensionalen Unterraum $V_1 \subset V$ und $V_2 := \{v \in V \mid \omega(v, v_1) = 0\}$. Es gilt $V_1 \cap V_2 = \{0\}$ und $V_1 + V_2 = V$, denn für $v \in V$ gilt

$$v - \omega(v_1, v_1)\omega(v, v_1)v_1 \in V_2.$$

Wir betrachten die Einschränkung von ω auf V_2 und fahren induktiv fort.

(ii) Für $\omega \neq 0$ existieren $\hat{v}_1, \hat{v}_{m+1} \in V$ mit $c_1 := \omega(\hat{v}_1, \hat{v}_{m+1}) \neq 0$. Setze $v_1 := \hat{v}_1/c_1$ und $v_{m+1} := \hat{v}_{m+1}$ Es ist

$$\omega(v_1, v_1) = \omega(v_{m+1}, v_{m+1}) = 0 \quad \text{und} \quad \omega(v_1, v_{m+1}) = -\omega(v_{m+1}, v_1) = 1.$$

Sei $V_1 := Span(v_1, v_{m+1}) \subset V$ und

$$V_2 := \{v \in V | \omega(v, w) = 0 \quad \text{für alle} \quad w \in V_1\}.$$

Es gilt $V_2 \cap V_1 = \{0\}$ und $V_2 + V_1 = V$, denn für $v \in V$ gilt

$$v + \omega(v, v_1) v_{m+1} - \omega(v, v_{m+1}) v_1 \in V_2.$$

Wir behandeln induktiv die Einschränkung von ω auf V_2 etc.

Definition 2.17.

- *Eine symplektische Form auf einem $\mathbb{R}$-Vektorraum V ist eine nicht degenerierte schiefsymmetrische Bilinearform $\omega : V \times V \to \mathbb{R}$.*
- *(V, ω) heißt dann symplektischer Vektorraum.*
- *Sind (V, ω) und (F, ω) symplektisch, so heißt eine lineare Abbildung $A : V \to W$ symplektisch, wenn $A^*\omega = \omega$.*

Bemerkung 2.18. (i) *Die symplektischen Abbildungen $A \in \mathcal{L}(V)$ sind diejenigen, die die symplektische Form ω erhalten, d.h. $A^*\omega = \omega$. Wegen dem vorangehenden Satz finden wir eine Basis von V, in der die darstellende Matrix von ω gleich $J = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$ ist. Es sei A die darstellende Matrix von A. Dann gilt*

$$A^t J A = J.$$

(ii) *Zwar sind symplektische Abbildungen volumenerhaltend, aber i.A. volumenerhaltende Abbildungen nicht symplektisch. Betrachten wir beispielsweise einen vierdimensionalen Vektorraum V mit Bais $v_1, \ldots, v_4$ und symplektischer Bilinearform ω mit Matrix $J = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$. Dann ist $A : V \to V$, $(v_1, v_2, v_3, v_4) \mapsto (-v_1, -v_2, v_3, v_4)$ volumenerhaltend, aber $\omega(Av_1), Av_3)) = -\omega(v_1, v_3)$. Läßt aber ein Endomorphismus A des $\mathbb{R}^2$ Die orientierte Fläche invariant, so ist A auch symplektisch. Denn für 2×2 Matrizen A gilt $\det(A)A^{-1} = J^t A^t J$, also $A^t J A = J$ genau dann, wenn der durch A gegebene Endomorphismus eine flächenerhaltende Abbildung ist.*

Satz 2.19. *Sei (V, ω) ein symplektischer Vektorraum, dann bildet die Menge der symplektischen Endomorphismen $A : V \to V$ unter der Komposition eine Gruppe, die symplektische Gruppe $Sp(V, \omega)$ genannt wird.*

Bemerkung 2.20. *Für einen Euklidischen Raum (V, ω) mit positiv definiter symmetrischer Bilinearform ω erhalten wir in ähnlicher Weise die orthogonale Gruppe.*

Beweis: Wir wählen eine Basis von V bezüglich der ω die darstellende Matrix $J = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$ hat. Eine lineare Abbildung ist genau dann symplektisch, wenn die entsprechende Matrix A die Gleichung $A^t JA = J$ erfüllt. Dann ist $J^{-1}A^tJ = -JA^tJ$ die inverse von A, also A invertierbar. Aus der Definition folgt, dass A^{-1} symplektisch ist. Wenn A, B zwei darstellende Matrizen von symplektischen Abbildungen sind, dann folgt

$$(AB)^t JAB = B^t A^t JAB = B^t JB = J.$$

Also ist auch die Komposition zweier symplektischer Abbildungen symplektisch. Die identische Abbildung ist offenbar symplektisch. **q.e.d.**

Wegen Satz 2.16 ist $Sp(V, \omega)$ isomorph zur Gruppe

$$Sp(n) = \{A \in \mathcal{L}(\mathbb{R}^{2n}) \mid (Av)^t JAw = v^t(A^t JA)w = v^t Jw \forall v, w \in \mathbb{R}^{2n} \iff A^t JA = J\}.$$

wobei ω bezüglich der kanonischen Basis des $\mathbb{R}^{2n}$ die darstellende Matrix J besitzt (ganz analog braucht man nach Satz 2.16 nur die orthogonale Gruppe $O(n)$ bezüglich des euklidischen Skalarproduktes zu betrachten.

Beispiel 2.21. *Der Fluss $\Phi(t, x) = \exp(tU)x$ mit $U = JA$ eines linearen Hamiltonschen Systems mit Hamiltonfunktion $H(x) = \frac{1}{2}(x, Ax)$ mit $A^t = A$ ist Element der symplektischen Gruppe $Sp(n)$.*

Der Betrag der Determinante eines symplektischen Endomorphismus ist gleich 1, denn es gilt ja

$$det(A^t JA) = (detA)^2 det(J) = det(J).$$

Übungsaufgabe 2.22. *Zeige in folgenden Schritten* $\det(A) = 1$ *für alle* $A \in Sp(n)$.

(i) *Sei ω die symplektische Form zu der Matrix $J = \left(\begin{smallmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{smallmatrix}\right)$ auf dem $\mathbb{R}^{2n}$. Dann definiert*

$$\begin{aligned}\omega^n : (\mathbb{R}^{2n})^{2n} \to \mathbb{R}, \quad (v_1, \ldots, v_{2n}) &\mapsto \omega^n(v_1, \ldots, v_{2n}) \\ &= \sum_{\sigma \in \Sigma_{2n}} (-1)^{|\sigma|} \omega\big(v_{\sigma(1)}, v_{\sigma(2)}\big) \cdot \ldots \cdot \omega\big(v_{\sigma(2n-1)}, v_{\sigma(2n)}\big)\end{aligned}$$

eine total antisymmetrische $2n$-lineare Abbildung. Zeige, dass jede symplektische Abbildung A diese Abbildung invariant läßt, d.h. dass für alle $v_1, \ldots, v_{2n} \in \mathbb{R}^{2n}$ auch $\omega^n(Av_1, \ldots, Av_{2n}) = \omega^n(v_1, \ldots, v_{2n})$ gilt.

(ii) *Zeige, dass jede total antisymmetrische* $2n$*-lineare Abbildung auf* $\mathbb{R}^{2n}$ *proportional zu der folgenden Abbildung ist: (benutze die Eigenschaften von* det*)*

$$(v_1, \ldots, v_{2n}) \mapsto \det(v_1, \ldots, v_{2n}).$$

Hier ist $(v_1, \ldots, v_{2n})$ *die Matrix, deren Spalten die Vektoren* $v_1, \ldots, v_{2n}$ *sind.*

(iii) *Folgere aus (ii), dass sich jede total antisymmetrische* $2n$*-lineare Abbildung* F *auf dem* $\mathbb{R}^{2n}$ *unter einer linearen Abbildung* A *folgendermaßen transformiert:*

$$F(Av_1, \ldots, Av_{2n}) = \det(A)F(v_1, \ldots, v_{2n}) \quad \text{für alle} \quad v_1, \ldots, v_{2n} \in \mathbb{R}^{2n}.$$

(iv) *Zeige dass* $\det(A) = 1$ *für alle* $A \in Sp(n)$ *gilt.*

Es folgt unmittelbar, dass das Produkt der $\mathbb{C}$-Nullstellen des charakteristischen Polynoms eines symplektischen Endomorphismus 1 ist.

Satz 2.23. *Sei* $A \in Sp(n)$ *und* $\lambda \in \mathbb{C}$ *Eigenwert von* A*. Dann sind auch* $\bar{\lambda}, \lambda^{-1}$ *und* $\bar{\lambda}^{-1}$ *Eigenwerte.*

Beweis: Dass mit λ auch $\overline{\lambda}$ Eigenwert ist, folgt daraus, dass die Koeffizienten des charakteristischen Polynoms $\det(\lambda\mathbb{1} - A)$ reell sind. Wir beweisen, dass $1/\lambda$ Eigenwert ist, indem wir das charakteristische Polynom von A betrachten. Dann gilt für $\lambda \in \mathbb{C}\backslash\{0\}$

$$\begin{aligned}
\det(\lambda\mathbb{1} - A) &= \det(J(\lambda\mathbb{1} - A)J^{-1}) & & \text{da} \quad \det J = 1 \neq 0 \\
&= \det(\lambda\mathbb{1} - JAJ^{-1}) & = \det(\lambda\mathbb{1} - (A^t)^{-1}) & \quad \text{da} \quad JAJ^{-1} = (A^t)^{-1} \\
&= \det(\lambda\mathbb{1} - A^{-1}) & = \det(A^{-1}(\lambda A - \mathbb{1})) & \\
&= \det(A^{-1}) \cdot \det(\lambda A - \mathbb{1}) & = \det(\lambda A - \mathbb{1}) & \quad \text{da} \quad \det(A) = 1 \\
&= (-\lambda)^{2n} \det(\lambda^{-1}\mathbb{1} - A) & = \lambda^{2n} \det(\lambda^{-1}\mathbb{1} - A). &
\end{aligned}$$

Da A invertierbar ist, sind alle komplexen Eigenwerte λ von A ungleich 0 und die Aussage folgt. **q.e.d.**

Auch die algebraischen Multiplizitäten der genannten Eigenwerte sind gleich.

Satz 2.24. *Sei* $A \in Sp(n)$ *und* $\lambda \in \mathbb{C}$ *Eigenwert von* A *mit Vielfachheit* k *(d.h. eine Nullstelle des charakteristischen Polynoms der Ordnung* k*). Dann sind auch* $\bar{\lambda}, \lambda^{-1}$ *und* $\bar{\lambda}^{-1}$ *Eigenwerte mit Vielfachheit* k*. Die Multiplizitäten etwaiger Eigenwerte* $+1$ *und* -1 *sind gerade.*

Beweis: Für $P(\lambda) = \det(\lambda\mathbb{1} - A)$ gilt $P(\lambda) = \lambda^{2n}P(\lambda^{-1})$. Einen Eigenwert $\lambda_0 \in \mathbb{C} \setminus \{0\}$ der Vielfachheit k können wir abspalten:

$$P(\lambda) = (\lambda - \lambda_0)^k Q(\lambda),$$

so dass für $\lambda \in \mathbb{C} \setminus \{0\}$ folgendes gilt:

$$\lambda^{2n} P(\lambda^{-1}) = (\lambda - \lambda_0)^k Q(\lambda) = (-\lambda\lambda_0)^k (\lambda^{-1} - \lambda_0^{-1})^k Q(\lambda).$$

Da $Q(\lambda)$ ein Polynom vom Grad $2n - k$ in λ ist, ist $(\lambda_0^k \lambda^{k-2n}) \cdot Q(\lambda)$ ein Polynom in λ^{-1}. Also ist λ_0^{-1} Nullstelle der Multiplizität $l \geq k$ von $P(\lambda^{-1})$. Vertauschen der Rollen von λ_0 und λ_0^{-1} zeigt $l = k$.

Es gilt genau dann $\lambda_0 = 1/\lambda_0$, wenn $\lambda_0 \in \{\pm 1\}$. Da insgesamt $2n$ Eigenwerte existieren und die Anzahl der Eigenwerte $\neq \pm 1$ gerade ist, muss die Multiplizität der 1 zusammen mit der der -1 gerade sein. Wegen $\det A = 1$ folgt dann auch einzeln gerade Multiplizität. **q.e.d.**

Beispiel 2.25. *Die Eigenwerte von $A \in Sp(1)$ sind entweder alle gleich* 1*, oder gleich* -1*, oder beide reell oder beide vom Betrag Eins und jeweils inverse zueinander.*

Nicht alle symplektischen Matrizen sind diagonalisierbar.

Beispiel 2.26. *Die Matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ gehört zu $Sp(1)$ ist aber nicht diagonalisierbar.*

Es gilt $Sp(n) \subset \mathcal{L}(\mathbb{R}^{2n})$ und damit wird die symplektische Gruppe zu einem metrischen Raum. Alle Einträge von $A^t JA - J$ definieren auf $A \in \mathcal{L}(\mathbb{R}^{2n})$ ein quadratisches Polynom. Weil alle diese Matrizen aber schiefsymmetrisch sind, sind höchstens $\frac{2n(2n-1)}{2}$ Einträge davon unabhängig, z.B. die $n(2n-1)$ oberen Dreieckseinträge. Wir behaupten jetzt, dass $n(2n-1)$ von diesen Polynome an allen Stellen $A \in Sp(n)$ linear unabhängige Ableitungen besitzen. Sei $A_0 \in Sp(n)$ und $A = A_0(\mathbb{1} + tU)$ mit $U \in \mathcal{L}(\mathbb{R}^{2n})$. Dann erhalten wir $A^t JA - J = t(U^t J + JU) + t^2(U^t JU)$. Weil A_0 invertierbar ist, ist der Kern der Ableitungen aller dieser Polynome bei $A = A_0$ isomorph zu $U \in sp(n)$. Offenbar ist eine Matrix $U \in \mathcal{L}(\mathbb{R}^{2n})$ genau dann in $sp(n)$, wenn JU symmetrisch ist. Deshalb hat $sp(n)$ die Dimension $(2n)^2 - \frac{2n(2n-1)}{2} = n(2n+1)$. Deshalb ist die Teilmenge $Sp(n) \subset \mathcal{L}(\mathbb{R}^{2n})$ tatsächlich die Nullstellenmenge von $n(2n-1)$ quadratischen Polynomen, deren Ableitungen auf $Sp(n)$ alle linear unabhängig sind. Wegen dem Satz der Impliziten Funktion, ist dann $Sp(n)$ eine $n(2n+1)$-dimensionale Untermannigfaltigkeit von dem Vektorraum $\mathcal{L}(\mathbb{R}^{2n})$.

Wir können Wege in $Sp(n)$, d.h. stetige Abbildungen $c : [0, 1] \to Sp(n)$ einführen. Wie verhalten sich dann die Eigenwerte von $c(t)$ bei Veränderung des Parameters t? Offensichtlich können Eigenwerte auf der reellen Achse bzw. dem Einheitskreis diese nicht verlassen, solange sie isoliert bleiben, Aus dieser Eigenschaft folgt eine Form von Stabilität Hamiltonscher Systeme unter kleinen Störungen.

2.3.2 Symplektische Algebra

Definition 2.27. *Eine lineare Abbildung $U \in \mathcal{L}(V)$ heißt infinitesimal symplektisch bezüglich einer symplektischen Bilinearform ω auf V, wenn*

$$\omega(Uv, w) + \omega(v, Uw) = 0 \quad \text{für alle} \quad v, w \in V.$$

Die Menge dieser Abbildungen bezeichnen wir mit $sp(V, \omega)$. Wegen der Linearität von ω und U ist $sp(V, \omega)$ ein Unterraum von $\mathcal{L}(V)$.

Definition 2.28. *Eine Liealgebra ist ein Vektorraum V mit einer bilinearen antisymmetrischen Abbildung*

$$[\cdot, \cdot] : V \times V \to V, \quad (v, w) \mapsto [v, w],$$

die die Jacobiidentität erfüllt:

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0 \quad \text{für alle} \quad a, b, c \in V.$$

Lemma 2.29. *Mit dem Kommutator $[A, B] = AB - BA$ von $A, B \in \mathcal{L}(V)$ ist $sp(V, \omega)$ eine Unter-Liealgebra von $\mathcal{L}(V)$.*

Beweis: $(\mathcal{L}(V), [\cdot, \cdot])$ ist eine Liealgebra, da der Kommutator bilinear und antisymmetrisch ist, und die Jacobi-Identität

$$\begin{aligned} &[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = \\ &\qquad A(BC - CB) - (BC - CB)A + B(CA - AC) - (CA - AC)B + \\ &\qquad\qquad + C(AB - BA) - (AB - BA)C = 0 \end{aligned}$$

für $A, B, C \in \mathcal{L}(V)$ erfüllt. Da für $A, B \in sp(V, \omega)$ und für $v, w \in V$

$$\begin{aligned} \omega([A, B]v, w) + \omega(v, [A, B]w) &= \omega(ABv, w) - \omega(BAv, w) + \omega(v, ABw) - \omega(v, BAw) \\ &= \omega(Bv, Aw) + \omega(Av, Bw) - \omega(Av, Bw) + \omega(Bv, Aw) = 0, \end{aligned}$$

ist auch $[A, B] \in sp(V, \omega)$. **q.e.d.**

Analog zu der obigen Aussage über die Gruppe $Sp(V, \omega)$ erhält man:

Satz 2.30. *Sei (V, ω) ein symplektischer Vektorraum und $U \in sp(V, \omega)$.*

(i) *Ist $\lambda \in \mathbb{C}$ Eigenwert von U mit Multiplizität k, so sind auch $-\lambda, \overline{\lambda}, -\overline{\lambda}$ Eigenwerte der Multiplizität k.*

(ii) *Ist Null Eigenwert, so besitzt er gerade Multiplizität.*

Beispiel 2.31. *Die Eigenwerte von $U \in sp(1)$ sind entweder beide reell oder imaginär.*

Eng verwandt mit dem Begriff der Liealgebra ist der der Liegruppe.

Beispiel 2.32. *Das wichtigste Beispiel einer Liegruppe ist $G = GL(n, \mathbb{R})$, die Gruppe der invertierbaren Matrizen in $\mathcal{L}(\mathbb{R}^n)$. Hier ist G als die offene Teilmenge derjenigen Matrizen A im n^2-dimensionalen Vektorraum $\mathcal{L}(\mathbb{R}^n)$ definiert, für die* $\det(A) \neq 0$. *Dadurch wird G zu einer sogenannten Untermannigfaltigkeit, und Matrizenmultiplikation und Inversion sind in den Matrizeneinträgen glatt. Betrachten wir die Abbildung*

$$\mathcal{L}(\mathbb{R}^n) \to GL(n, \mathbb{R}), \quad U \mapsto \exp(U),$$

dann liegt das Bild der Exponentialfunktion tatsächlich in G, denn

$$\det(\exp(U)) = \exp(\mathrm{Spur}(U)) > 0.$$

Da $g \in G$ mit $\det(g) < 0$ *existieren, ist* exp *nicht surjektiv ist. Für $A \in G$ mit* $\|A - \mathbb{1}\| < 1$ *ist* exp *invertierbar, denn die Potenzreihe von* ln *konvergiert:*

$$\ln(A) = \ln(\mathbb{1} + (A - \mathbb{1})) = -\sum_{k=1}^{\infty} \frac{(\mathbb{1} - A)^k}{k}.$$

Die Ableitung von exp *ist bei $g = \mathbb{1}$ bijektiv. Deshalb ist wegen dem Satz der inversen Funktion die Einschränkung von* exp *auf eine Umgebung von $0 \in \mathcal{L}(V)$ ein Homöomorphismus auf eine Umgebung von $\mathbb{1} \in GL(V)$. Die Liealgebra $\mathcal{L}(\mathbb{R}^n)$ bildet mit dem Kommutator eine Liealgebra. Der Kommutator misst den Mangel an Kommutativität in der Gruppenmultiplikation, denn für $U_1, U_2 \in \mathcal{L}(\mathbb{R}^n)$ ist*

$$\exp(\epsilon U_1)\exp(\epsilon U_2)\exp(-\epsilon U_1)\exp(-\epsilon U_2) = \mathbb{1} + \epsilon^2[U_1, U_2] + \mathrm{O}(\epsilon^3).$$

Man nennt daher $gl(n, \mathbb{R}) = \mathcal{L}(\mathbb{R}^n)$ die Liealgebra von $GL(n, \mathbb{R})$.

Analog können wir die Exponentialabbildung

$$sp(V, \omega) \to Sp(V, \omega), \quad U \mapsto \exp(U)$$

von der symplektischen Liealgebra in die symplektischen Liegruppen betrachten. Für $U \in sp(V, \omega)$ ist $\exp(U) \in Sp(V, \omega)$, denn aus von $\omega(Uv, w) = \omega(v, -Uv)$ folgt $\omega(\exp(U)v, w) = \omega(v, \exp(-U)w)$ und damit auch

$$\omega(\exp(U)v, \exp(U)w) = \omega(v, \exp(-U)\exp(U)w) = \omega(v, w).$$

Wir sehen, dass die Eigenwerte λ mit $Re(\lambda) > 0$ von U zu Eigenwerten $\exp(\lambda)$ mit Betrag > 1 von $\exp(U)$ werden. Erinnern wir uns nun an die Tatsache, dass für die quadratische Hamiltonfunktion $H(x) = \frac{1}{2}(x, Ax)$ die Differentialgleichung $\dot{x} = X(H)(x) = JAx$ die Lösung $\Phi_t(x) = \exp(U)x$ mit $U = JAt \in sp(n)$ besitzt, dann wird klar, dass der Fixpunkt 0 höchstens dann Liapunovstabil sein kann, wenn alle Eigenwerte von JA rein imaginär sind.

2.4 Variationsrechnung

Für das Folgende sind einige Kenntnisse aus der Variationsrechnung erforderlich. Die Variationsrechnung befaßt sich mit der Suche nach Funktionen, die kritische Punkte eines Funktionals sind, d.h. einer reellen Funktion auf einem undendlichdimensionalen Funktionenraum. In unseren Fall nennen wir diesen Raum Kurvenraum. Solche Funktionen nennt man Funktionale. Das Standardbeispiel eines Funktionals ist die Kurvenlänge in der euklidischen Ebene. Eine parametrisierte Kurve $\gamma = \{t, x : x = x(t), t_0 \leq t \leq t_1\}$ ist dabei eine Abbildung $t \mapsto x(t)$ von einem Intervall nach $\mathbb{R}^n$:

$$\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2} dt \quad \text{für} \quad \gamma : \quad [t_0, t_1] \to \mathbb{R}^n, \quad t \mapsto x(t).$$

Im allgemeinen versteht man unter einem Funktional eine Abbildung eines Raumes von Kurven nach $\mathbb{R}$ oder $\mathbb{C}$. Im Fall eines differenzierbaren Funktionals Φ ist das Differential $\Phi'(\gamma)$ von Φ an der Stelle γ eindeutig bestimmt als ein Element des Dualraumes des entsprechenden Funktionenraumes:

$$\lim_{\|h\| \to 0} \frac{|\Phi(\gamma + h) - \Phi(\gamma) - \Phi'(\gamma)(h)|}{\|h\|} \to 0.$$

Das Differential des Funktionals nennt man auch Variation des Funktionals und h die Variation der Kurve.

Beispiel 2.33. *Es sei $\gamma = \{t, x : x = x(t), t_0 \leq t \leq t_1\}$ eine Kurve in der t, x-Ebene, $\dot{x} = \frac{dx}{dt}$ und $L = L(a, b, c)$ eine differenzierbare Funktion dreier Variablen. Wir definieren ein Funktional Φ durch*

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t), t) dt.$$

Im Spezialfall $L = \sqrt{1 + b^2}$ ergibt sich die Bogenlänge γ.

Satz 2.34. *Wenn $(a, b, c) \to L(a, b, c)$ zweimal stetig differenzierbar ist, dann ist das Funktional*

$$\Phi : \quad \gamma \to \Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t) dt$$

differenzierbar, und seine Ableitung wird durch folgende Formel gegeben:

$$\Phi'(\gamma)(h) = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h dt + \left. \frac{\partial L}{\partial \dot{x}} h \right|_{t_0}^{t_1}.$$

Beweis: Es gilt

$$\Phi(\gamma+h)-\Phi(\gamma)=\int\limits_{t_0}^{t_1}\Big(L(x+h,\dot{x}+\dot{h},t)-L(x,\dot{x},t)\Big)\,dt$$
$$=\int\limits_{t_0}^{t_1}\left(\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}\right)dt+\mathrm{O}(h^2)=\Phi'(\gamma)(h)+\mathrm{O}(h^2)$$

mit $\Phi'(\gamma)(h)=\int_{t_0}^{t_1}\left(\frac{\partial L}{\partial x}h+\frac{\partial L}{\partial \dot{x}}\dot{h}\right)dt$. Eine partielle Integration ergibt

$$\int\limits_{t_0}^{t_1}\frac{\partial L}{\partial \dot{x}}\dot{h}dt=-\int\limits_{t_0}^{t_1}\left(\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right)hdt+\left.\frac{\partial L}{\partial \dot{x}}h\right|_{t_0}^{t_2}.$$ **q.e.d.**

Definition 2.35. *Unter einem kritischen Punkt eines differenzierbaren Funktionals $\Phi(\gamma)$ versteht man die Kurve γ, für die $\Phi'(\gamma)(h)=0$ für beliebiges h ist; genauso, wie ein kritischer Punkt einer Funktion eine Nullstelle des Differentials ist.*

Satz 2.36. *Damit die Kurve $\gamma=\{t,x:x=x(t),t_0\le t\le t_1\}$ ein kritischer Punkt des Funktionals $\Phi(\gamma)=\int_{t_0}^{t_1}L(x.\dot{x},t)dt$ im Raum der Kurven wird, die durch die Punkte $x(t_0)=x_0$ und $x(t_1)=x_1$ verlaufen, ist es notwendig und hinreichend, dass*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}-\frac{\partial L}{\partial x}=0$$

längs der Kurve $x(t)$ gilt.

Lemma 2.37. *Wenn eine stetige Funktion $f(t)$ auf $t\in[t_0,t_1]$ für jede glatte Funktion $h(t)$ mit $h(t_0)=h(t_1)=0$ die Beziehung $\int_{t_0}^{t_1}f(t)h(t)dt=0$ erfüllt, dann ist $f=0$.*

Beweis des Lemmas. Es sei $f(t^*)>0, t_0<t^*<t_1$. Infolge der Stetigkeit ist $f(t)>c$ in einer gewissen Umgebung des Punktes t^*:

$$t_0<t^*-\epsilon<t<t^*+\epsilon<t_1.$$

Ferner sei $h(t)=0$ außerhalb der Umgebung, $h(t)\ge 0$ innerhalb der Umgebung und $h(t)=1$ auf der halben Umgebung (d.h., $t^*-\frac{\epsilon}{2}<t<t^*+\frac{\epsilon}{2}$). Dann ist offensichtlich

$$\int\limits_{t_0}^{t_1}f(t)h(t)dt=\int\limits_{t^*-\epsilon}^{t^*+\epsilon}f(t)h(t)dt\ge\int\limits_{t^*-\frac{\epsilon}{2}}^{t^*+\frac{\epsilon}{2}}f(t)h(t)dt\ge\epsilon c>0.$$

Der so erhaltene Widerspruch zeigt, dass $f(t^*) = 0$ für alle $t_0 < t^* < t_1$ ist. **q.e.d.**
Beweis des Satzes. Aus dem vorangehenden Satz ergibt sich

$$\Phi'(\gamma)(h) = \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \right) h dt + \frac{\partial L}{\partial \dot{x}} h \Big|_{t_0}^{t_1}.$$

Der zweite Summand ist gleich 0 wegen $h(t_0) = h(t_1) = 0$. Wenn γ ein kritischer Punkt ist, gilt $\Phi(\gamma)(h) = 0$ für alle h mit $h(t_0) = h(t_1) = 0$. Deshalb ist für solche stetigen h

$$\int_{t_0}^{t_1} f(t)h(t)dt = 0 \quad \text{mit} \quad f(t) = \frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}}.$$

Nach dem Lemma ist $f = 0$. Umgekehrt folgt aus $f = 0$ offenbar $\Phi'(\gamma) = 0$. **q.e.d.**

Beispiel 2.38. *Man überprüfe, dass ein kritischer Punkt der Bogenlänge eine Gerade ist. Es gilt nämlich* $L = \sqrt{1 + \dot{x}^2}$, *und damit* $\frac{\partial L}{\partial x} = 0$ *und* $\frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1+\dot{x}^2}}$. *Dann folgt aus*

$$\frac{d}{dt}\left(\frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} \right) = 0 \quad \textit{auch} \quad \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} = c \quad \textit{und} \quad \dot{x} = c_1 \quad \textit{bzw.} \quad x = c_1 t + c_2.$$

Definition 2.39. *Die Gleichung*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.$$

heißt Euler-Lagrangesche Gleichung des Funktionals

$$\Phi = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt.$$

Sei x ein Vektor im n-dimensionalen Raum, $\gamma = \{t, x : x = x(t), t_0 \le t \le t_1\}$ eine Kurve im $(n+1)$-dimensionalen Raum $\mathbb{R} \times \mathbb{R}^n$ und $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ eine Funktion von $2n + 1$ Argumenten. Analog zum Vorhergehenden läßt sich der folgende Satz beweisen:

Satz 2.40. *Damit die Kurve* γ *kritischer Punkt des Funktionals* $\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt$ *im Raum der Kurven* γ *im* $\mathbb{R} \times \mathbb{R}^n$ *von der Form* $t \mapsto (t, x(t))$ *ist, welche die beiden gegebenen Punkte* $(t_0, x_0), (t_1, x_1)$ *verbinden, ist es notwendig und hinreichend, dass die Euler-Lagrangesche Gleichung*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0$$

längs der Kurve γ *erfüllt wird.* **q.e.d.**

Das ist ein System von n Gleichungen zweiter Ordnung, und die Lösung hängt von $2n$ willkürlichen Konstanten ab. Sie gestatten die $2n$ Bedingungen $x(t_0) = x_0$ und $x(t_1) = x_1$ zu erfüllen.

Wichtige Bemerkung 2.41. *Die Eigenschaft einer Kurve γ, kritischer Punkt eines Funktionals zu sein, hängt nicht von der Wahl des Koordinatensystems ab.*

Zum Beispiel wir ein und dasselbe Funktional - die Bogenlänge einer Kurve - in kartesischen Koordinaten und in Polarkoordinaten durch verschiedene Formeln

$$\Phi_{\text{kar}} = \int_{t_0}^{t_1} \sqrt{\dot{x}^2 + \dot{y}^2 dt} \quad \text{und} \quad \Phi_{\text{pol}} = \int_{t_0}^{t_1} \sqrt{\dot{r}^2 + r^2\dot{\varphi}^2 dt}$$

gegeben. Aus der Beziehung $(x, y) = (r\cos\varphi, r\sin\varphi)$ folgt nämlich

$$\begin{aligned}(\dot{x}, \dot{y}) &= (\dot{r}\cos\varphi - r\dot{\varphi}\sin\varphi, \dot{r}\sin\varphi + r\dot{\varphi}\cos\varphi)\\ \dot{x}^2 + \dot{y}^2 &= (\dot{r}^2 + r^2\dot{\varphi}^2)(\cos^2\varphi + \sin^2\varphi) = (\dot{r}^2 + r^2\dot{\varphi}^2).\end{aligned}$$

Die kritischen Punkte sind die gleichen - es sind Geraden in der Ebene. Die Gleichungen der Geraden in kartesischen Koordinaten und in Polarkoordinaten ergeben sich durch verschiedene Funktionen $(x(t), y(t))$ und $(r(t), \varphi(t))$. Beide erfüllen die Euler-Lagrangesche Gleichung

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0,$$

nur ist im ersten Fall

$$x_{\text{kar}} = (x, y) \quad \text{und} \quad L_{\text{kar}} = \sqrt{\dot{x}^2 + \dot{y}^2},$$

und im zweiten

$$x_{\text{pol}} = (r, \varphi), \quad L_{\text{pol}} = \sqrt{\dot{r}^2 + r^2\dot{\varphi}^2}.$$

Somit können wir leicht die Differentialgleichung der Gesamtheit aller Geraden in beliebigen Koordinaten aufschreiben.

2.5 Lagrangesche Gleichungen

Hier wird das Variationsprinzip hergeleitet, dessen kritischen Punkte die Lösung der Newtonschen Gleichungen für die Bewegung eines Systems von Teilchen mit den Koordinaten q_i im Kraftfeld eines Potentials U ist. Wir vergleichen die Newtonsche Bewegungsgleichung

$$\frac{d}{dt}m\dot{q}_i = F = -\frac{\partial U}{\partial q_i}$$

mit der Euler-Lagrangeschen Gleichung

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.$$

Satz 2.42. *Die Lösungen der Bewegungsgleichung des mechanischen Systems von Teilchen im Kraftfeld eines Potentials U sind kritische Punkte des Funktionals $\Phi(\gamma) = \int_{t_0}^{t_1} Ldt$, wobei $L = T - U$ die Differenz von kinetischer und potentieller Energie ist.*

Beweis: Mit $U = U(q)$ und $T = \sum m_i \frac{\dot{q}_i^2}{2}$ haben wir

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} = m_i \dot{q}_i \quad \text{und} \quad \frac{\partial L}{\partial q_i} = -\frac{\partial U}{\partial q_i} = F_i.$$ **q.e.d.**

Folgerung 2.43. *Es eien $q_1, \ldots, q_{3n}$ beliebige Koordinaten im Konfigurationsraum eines Systems von n Massenpunkten. Dann wird die Änderung von q in der Zeit durch die Euler-Lagrangeschen Gleichungen beschrieben:*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \textit{ mit } L = T - U.$$

Beweis. Nach dem vorhergehenden Satz ist die Bewegung ein kritischer Punkt des Funktionals $\int Ldt$. Also sind in einem beliebigen Koordinatensystem die in diesem System aufgeschriebenen Euler-Lagrangeschen Gleichungen erfüllt. **q.e.d.**

Definition 2.44. *In der Mechanik sind die folgenden Bezeichnungen gebräuchlich:*

- $L(q, \dot{q}, t) = T - U$ *ist die Lagrange-Funktion*
- q_i *sind die generalisierten Koordinaten*
- $\dot{q}_i$ *die generalisierten Geschwindigkeiten*
- $\frac{\partial L}{\partial \dot{q}_i} = p_i$ *die kanonisch konjugierten (generalisierten) Impulse*
- $\frac{\partial L}{\partial q_i}$ *die verallgemeinerten Kräfte*
- $\int L(q, \dot{q}, t)dt$ *ist die Wirkung*
- $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial \dot{q}_i} = 0$ *sind die Lagrangeschen Gleichungen.*

Der obige Satz wird das Hamiltonsche Prinzip der kleinsten Wirkung genannt, weil die Bewegung $q(t)$ in einigen Fällen nicht nur extremal ist, sondern auch den kleinsten Wert des Funktionals der Wirkung $\int L dt$ liefert.

Beispiel 2.45. *Für einen freien Massenpunkt im $\mathbb{R}^3$ ist*

$$L = T = \frac{m\dot{q}^2}{2} = \frac{m}{2}(\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2).$$

Hier sind die generalisierten Geschwindigkeiten die Komponenten des Geschwindigkeitsvektors und die kanonisch konjugierten Impulse $p_i = m\dot{q}_i$ die Komponenten des Impulsvektors, und die Lagrangeschen Gleichungen stimmen mit den Newtonschen $\frac{dp}{dt} = 0$ überein. Die kritischen Punkte sind Geraden. Aus dem Hamiltonschen Prinzip folgt, dass die Geraden nicht nur die Kürzesten sind (d.h. die kritischen Punkte der Bogenlänge $\int \sqrt{\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2 dt}$), sondern auch die kritischen Punkte der Wirkung

$$\int_{t_0}^{t_1} (\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2) dt.$$

Übungsaufgabe 2.46. *Man beweise, dass dieser kritische Punkt ein Minimum ist.*

Beispiel 2.47. *Wir betrachten die Bewegung im Zentralfeld in der Ebene in Polarkoordinaten $q_1 = r, q_2 = \varphi$. Aus der Beziehung $(x, y) = (r\cos\varphi, r\sin\varphi)$ erhalten wir*

$$(\dot{x}, \dot{y}) = \dot{r}\cos\varphi - r\dot{\varphi}\sin\varphi, \dot{r}\sin\varphi + r\dot{\varphi}\cos\varphi)$$

und damit für die kinetische Energie

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2)(\cos^2\varphi + \sin^2\varphi) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2)$$

und die Lagrange-Funktion

$$L(q, \dot{q}) = T(q, \dot{q}) - U(q) \quad \textit{mit} \quad U = U(q_1).$$

Die kanonisch konjugierten Impulse sind $p_i = \frac{\partial L}{\partial \dot{q}_i}$, d.h.

$$p_1 = m\dot{r} \quad \textit{und} \quad p_2 = mr^2\dot{\varphi}.$$

Wir können p_2 als z-Komponente M_z des Drehimpulses interpretieren:

$$\begin{aligned} M &= m(x, y, 0) \wedge (\dot{x}, \dot{y}, 0) = (0, 0, x\dot{y} - y\dot{x}) \\ &= (0, 0, mr\cos\varphi(\dot{r}\sin\varphi + r\dot{\varphi}\cos\varphi) - mr\sin\varphi(\dot{r}\cos\varphi - r\dot{\varphi}\sin\varphi)) = (0, 0, mr^2\dot{\varphi}). \end{aligned}$$

Die Lagrangesche Gleichung $\dot{p}_i = \frac{\partial L}{\partial q_i}$ *nimmt folgende Form an:*

$$\dot{p}_1 = m\ddot{r} = mr\dot{\varphi}^2 - \frac{\partial U}{\partial r} \qquad \dot{p}_2 = 2mr\dot{r} + mr^2\ddot{\varphi} = \frac{\partial L}{\partial q_2} = -\frac{\partial U}{\partial \varphi} = 0.$$

Deshalb lautet die zweite Lagrangesche Gleichung $\dot{p}_2 = 0$ *also* p_2 *konstant. Das ist der Drehimpulserhaltungssatz. Im allgemeinen Fall, d.h. wenn das Feld nicht zentral, also* $U = U(r,\varphi)$ *ist, folgt* $\dot{p}_2 = -\frac{\partial U}{\partial \varphi}$. *Wegen*

$$\begin{aligned} dU &= \frac{\partial U}{\partial r}dr + \frac{\partial U}{\partial \varphi}d\varphi = -F_x dx - F_y dy \\ &= -(F_x \cos\varphi + F_y \sin\varphi)dr + r(F_x \sin\varphi - F_y \cos\varphi)d\varphi \end{aligned}$$

gilt dann auch

$$(x,y,0)\wedge(F_x,F_y,0) = (0,0,xF_y - yF_x) = (0,0,-r(F_x\sin\varphi - F_y\cos\varphi) = \left(0,0,-\frac{\partial U}{\partial \varphi}\right).$$

Diese Gleichung läßt sich also in

$$\frac{d}{dt}M = \frac{d}{dt}mx\wedge\dot{x} = m\dot{x}\wedge\dot{x} + mx\wedge\ddot{x} = x\wedge F = N$$

umformen. Die zeitliche Änderung des Drehimpulses bezüglich der z-Achse ist gleich dem Drehmoment der Kraft F *bezüglich der z-Achse.*

Das betrachtete Beispiel legt eine Verallgemeinerung des Drehimpulserhaltungssatzes nahe.

Definition 2.48. *Die Koordinate* q_i, *heißt zyklisch, wenn die Lagrange-Funktion von ihr nicht abhängt:* $\frac{\partial L}{\partial q_i} = 0$.

Satz 2.49. *Der zu einer zyklischen Koordinate* q_i *gehörende kanonisch konjugierte Impuls* $p_i = \frac{\partial L}{\partial \dot{q}_i}$ *bleibt erhalten.*

Beweis: Auf Grund der Lagrangeschen Gleichung gilt $\displaystyle\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} = 0.$ **q.e.d.**

2.6 Legendre-Transformation

Als nächstes wollten wir die Lagrangesche Mechanik in die Hamiltonsche Mechanik transformieren. Durch diese Transformation werden alle Funktionen nicht mehr als

Funktionen von q und $\dot{q}$ sondern als Funktionen von q und den konjugierten Impulsen $p = \frac{\partial L}{\partial \dot{q}}$ aufgefasst. Diese zweiten Variablen sind aber nur implizit gegeben. Die Transformation, die das leistet heißt Legendre-Transformation.

Die Legendre Transformation ist ein mathematisches Verfahren, welches auf einem Übergang von Funktionen eines linearen Raumes zu Funktionen in einem dualen Raum beruht. Die Legendre-Transformation ist verwandt mit der projektiven Dualität und den Tangentialkoordinaten der algebraischen Geometrie oder mit der Konstruktion des dualen Banachraumes in der Analysis. Sie kommt in der Physik (z.B. bei der Definition thermodynamischer Größen) oft vor.

Definition 2.50. *Es sei $y = f(x)$ eine konvexe stetige Funktion auf einem reellen Intervall. Für alle $p \in \mathbb{R}$ sei $F(p, x) = px - f(x)$. Für alle $p \in \mathbb{R}$, für die die Funktion $x \mapsto F(p, x)$ in Abhängigkeit von x auf dem Definitionsbereich von f ein Maximum annimmt, sei $x(p)$ die Koordinate eines solchen Maximums und g definiert durch*

$$g(p) = F(p, x(p)) = px(p) - f(x(p)).$$

Wenn f streng konvex ist, dann ist $x(p)$, wenn es existiert auch eindeutig.

Übungsaufgabe 2.51. *Man zeige, daß des Definitionsbereich von g ein Punkt, eine beschränktes Intervall oder ein einseitig unbeschränktes Intervall oder ganz $\mathbb{R}$ sein kann, wenn die Funktion f auf der ganzen x-Achse definiert ist. Man zeige ferner: Ist die Funktion f in einem abgeschlossenen Intervall definiert, so ist sie auf der ganzen p-Achse definiert.*

Beispiel 2.52. (i) *Es sei $f(x) = x^2$. Dann ist*

$$F(p, x) = px - x^2 \quad \text{und} \quad x(p) = \frac{p}{2} \quad \text{und} \quad g(p) = \frac{1}{4}p^2.$$

(ii) *Es sei $f(x) = \frac{mx^2}{2}$. Dann ist*

$$F(p, x) = px - \frac{mx^2}{2} = -\frac{m}{2}\left(x - \frac{p}{m}\right)^2 + \frac{p^2}{m}.$$

Also gilt $g(p) = \frac{p^2}{2m}$.

(iii) *Es sei $f(x) = \frac{x^\alpha}{\alpha}$ mit $1 < \alpha$. Dann hat für alle $p \in \mathbb{R}$ die Funktion $x \mapsto F(p, x) = px - f(x)$ eine monoton fallende Ableitung $x \mapsto p - x^{\alpha-1}$ und bei $x = p^{\frac{1}{\alpha-1}}$ ein eindeutiges Maximum. Also folgt*

$$g : \mathbb{R} \to \mathbb{R}, \quad p \mapsto g(p) = p^{1+\frac{1}{\alpha-1}} - \frac{p^{\frac{\alpha}{\alpha-1}}}{\alpha} = \frac{\alpha - 1}{\alpha} p^{\frac{\alpha}{\alpha-1}} = \frac{p^\beta}{\beta} \quad \text{mit} \quad \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

(iv) *Es sei $f(x)$ ein konvexes Polygon. Dann ist auch g ein konvexes Polygon, wobei den Ecken von $f(x)$ die Kanten von $g(p)$ entsprechen und den Kanten von $f(x)$ die Ecken von $g(p)$. Beispielsweise geht ein Winkel in eine Strecke über.*

Wir nehmen an, daß die Funktion f differenzierbar ist. Dann muss f' monoton wachsend sein. Es ist leicht zu prüfen, daß die Legendre-Transformation eine konvexe Funktion in eine konvexe überführt. Somit können wir sie zweimal anwenden.

Lemma 2.53. *Sie f eine differenzierbare konvexe Funktion auf einem offenen Intervall I. Dann verläuft der Graph von f oberhalb aller Tangenten an den Graphen von f.*

Beweis: Weil f konvex ist, ist f' monoton wachsend. Dann folgt für alle $x_0, x \in I$:

$$f(x) - f(x_0) - (x - x_0)f'(x_0) = \int_{x_0}^{x} (f'(t) - f'(x_0))dt \begin{cases} \geq 0 & \text{für} \quad x \geq x_0 \\ \geq 0 & \text{für} \quad x \leq x_0. \end{cases}$$

Offenbar ist $x \mapsto f(x_0) + (x - x_0)f'(x_0)$ die Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$. Deshalb veräuft die Tangente an alle Punkte unterhalb des Graphen.**q.e.d.**

Satz 2.54. *Die Legendre-Transformation ist für differenzierbare konvexe Funktionen involutiv, d.h. ihr Quadrat ist gleich der identischen Transformation: Wenn f bei der Legendre-Transformation in g übergeht, liefert die Legendre-Transformation von g wieder f.*

Beweis: Um die Legendre-Transformierte der Funktion g von der Variablen p zu bestimmen, müssen wir definitionsgemäß eine neue unabhängige Variable betrachten (die wir mit x bezeichnen), die Funktion

$$G(x, p) = xp - g(p)$$

aufstellen und den Punkt $p(x)$ finden, in welchem G ein Maximum hat:

$$G(x, p) \leq G(x, p(x)) \quad \text{für alle} \quad p.$$

Dann wird durch die Legendre-Transformation von $g(p)$ eine Funktion von x gegeben, die gleich $G(x, p(x))$ ist. Wir beweisen, daß $G(x, p(x)) = f(x)$ ist. Dazu bemerken wir, daß für festes p im Definitionsbereich von g die Abbildung $x \mapsto G(x, p) = xp - g(p)$ eine einfache geometrische Bedeutung hat: Es ist die Tangente an den Graphen von f in dem Punkt $(x(p), f(x(p)))$, welche die Steigung p hat. Tatsächlich ist bei festem p die Funktion $G(x, p)$ eine lineare Funktion von x, wobei $\frac{\partial G}{\partial x} = p$ ist, und bei $x = x(p)$ haben wir $G(x, p) = xp - g(p) = xp - xp + f(x(p)) = f(x(p))$ gemäß der Definition von $g(p)$. Weil f differenzierbar ist und $x(p)$ ein kritischer Punkt von $x \mapsto px - f(x)$ muss auch $f'(x(p)) = p$ gelten.

Wir halten nun den Punkt $x = x_0$ fest und verändern p. Dann sind die Werte $G(x,p)$ die Funktionswerte der Schnittpunkte der Tangenten an den Graphen von f mit der Geraden $x = x_0$. Aus dem vorangehenden Lemma folgt, dass all diese Tangenten unterhalb des Graphen von f verlaufen und dass deshalb das Maximum $G(x,p)$ bei festem $x(p_0)$ gleich $f(x)$ ist und für $p = p(x_0) = f'(x_0)$ angenommen wird. **q.e.d.**

Folgerung 2.55. *Es sei eine Schar von Geraden $y = px - g(p)$ gegeben. Dann hat die Einhüllende dieser Schar die Gleichung $y = f(x)$, wobei f die Legendre-Transformation der Funktion g ist.*

Definition 2.56. *Zwei Funktionen f und g, die Legendre-Transformierte voneinander sind, nennt man nach Young zueinander dual.*

Nach Definition der Legendre-Transformation gilt für alle p auch $F(x,p) = px - f(x) \leq g(p)$ für beliebige x. Dann gilt diese Ungleichung für beliebige x und p. Hieraus folgt die Youngsche Ungleichung

$$px \leq f(x) + g(p).$$

2.6.1 Der Fall mehrerer Variabler

Es sei jetzt $x \mapsto f(x)$ eine konvexe Funktion auf einer konvexen Teilmenge $\Omega \subset \mathbb{R}^n$; d.h. der für je zwei Punkte $x, y \in \Omega$ liegt der Graph von f unter der Verbindungsstrecke zwischen den beiden Punkten:

$$f(x + t(y - x) \leq f(x) + t(f(y) - f(x) \quad \text{für alle} \quad t \in [0,1].$$

Wenn f zweimal differenzierbar ist, sind die zweiten Richtungsableitungen in alle Richtungen nicht negativ. Deshalb ist die Hessische, also die Matrix der zweiten partiellen Ableitungen von f, eine nicht negative Matrix, d.h. für alle $dx \in \mathbb{R}^n$ gilt

$$\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j \geq 0.$$

Dann wird die Legendre Transformation $p \mapsto g(p)$ der Funktion f definiert als

$$g(p) = F(p, x(p)) = max\{F(p,x) \mid x \in \Omega\} \quad \text{mit} \quad F(p,x) = p \cdot x - f(x).$$

Der Definitionsbereich besteht aus allen p, für die das Maximum in der obigen Definition existiert. Wenn f differenzierbar ist, dann gilt $p = \frac{\partial f}{\partial x}$. Alle vorangegangenen Betrachtungen, einschließlich der Youngschen Ungleichung, lassen sich ohne Änderungen auf diesen Fall übertragen.

Übungsaufgabe 2.57. **(i)** *Weil für jede konvexe Funktion f auf einer konvexen Teilmenge $\Omega \subset \mathbb{R}^n$ die Einschränkung auf die Schnittmenge von Ω mit einer beliebigen Geraden im $\mathbb{R}^n$ eine konvexe Funktion ist, verläuft für jede differenzierbare konvexe Funktion auf einer offenen konvexen Teilmenge $\Omega \subset \mathbb{R}^n$ die Tangentialebene an jeden Punkt unterhalb des Graphen von f.*

(ii) *Zeige, dass für eine differenzierbare konvexe Funktion f auf einer offenen konvexen Teilmenge $\Omega \subset \mathbb{R}^n$ für $x_0 \in \Omega$ die Menge $\{x \in \Omega \mid f'(x) = f'(x_0)\}$ konvex ist.*

(iii) *Es sei $f : V \to \mathbb{R}$ eine konvexe differenzierbare Funktion auf einem endlichdimensionalen Vektorraum V, so dass die Abbildung $f' : V \to V' = \mathcal{L}(V, \mathbb{R})$ in den Dualraum von V surjektiv ist. Man zeige, dass für alle p das folgende Maximum existiert und eine Funktion $g : V' \to \mathbb{R}$ definiert:*

$$g(p) = F(p, x(p)) = max\{F(p, x) \mid x \in \Omega\} \quad mit \quad F(p, x) = \langle p, x\rangle - f(x).$$

(iv) *Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine quadratische Form $x \mapsto f(x) = \sum f_{ij}x_jx_j$ für alle $x \in \mathbb{R}^n$. Man zeige, dass ihre Legendre-Transformierte ebenfalls eine quadratische Form $g : \mathbb{R}^n \to \mathbb{R}$ mit $p \mapsto g(p) = \sum g_{ij}p_ip_j$ ist, wobei die Werte beider Formen in den entsprechenden Punkten übereinstimmen. D.h. es gilt*

$$f(x(p)) = g(p) \quad \text{für alle} \quad p \in \mathbb{R}^n \quad \text{und} \quad g(p(x)) = f(x) \quad \text{für alle} \quad x \in \mathbb{R}^n.$$

2.7 Hamiltonsche Gleichungen

Durch die Legendre Transformation geht das System der Lagrangschen Differentialgleichungen zweiter Ordnung in ein besonderes symmetrisches System von $2n$ Gleichungen erster Ordnung über, in das System der Hamiltonschen Gleichungen (oder der kanonischen Gleichungen).

Wir betrachten das System der Lagrangeschen Gleichungen

$$\dot{p} = \frac{\partial L}{\partial q} \quad \text{mit} \quad p = \frac{\partial L}{\partial \dot{q}}$$

zu einer gegebenen Lagrange-Funktion $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, welche wir bezüglich des zweiten Arguments $\dot{q}$ als konvex annehmen.

Satz 2.58. *Die Lagrangschen Gleichungen sind zu folgendem System von $2n$ Gleichungen erster Ordnung äquivalent, die Hamiltonsche Gleichungen genannt werde:*

$$\dot{p} = -\frac{\partial H}{\partial q} \quad und \quad \dot{q} = \frac{\partial H}{\partial p}.$$

Hierbei ist $H(q,p,t) = \max\{p\dot{q} - L(q,\dot{q},t) \mid \dot{q} \in \mathbb{R}^n\}$ für festes q und t die Legendre-Transformierte der Lagrange Funktion, als Funktion von $\dot{q}$.

Beweis: Auf Grund der Definition ist für festes q und t die Legendre-Transformierte von $L(q,\dot{q},t)$ bezüglich $\dot{q}$ eine Funktion $H(p) = p\dot{q} - L(\dot{q})$, in der $\dot{q}$ durch p mittels der Formel $p = \frac{\partial L}{\partial \dot{q}}$ ausgedrückt ist und die noch von den Parametern q,t abhängt. Diese Funktion H nennt man Hamilton-Funktion. Ihr totales Differential

$$dH = \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial t}dt,$$

ist gleich dem totalen Differential von $p\dot{q} - L$ für $p = \frac{\partial L}{\partial \dot{q}}$:

$$dH = \dot{q}dp - \frac{\partial L}{\partial q}dq - \frac{\partial L}{\partial t}dt.$$

Beide Ausdrücke für dH müssen übereinstimmen. Deshalb gilt

$$\dot{q} = \frac{\partial H}{\partial p} \quad \text{und} \quad \frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} \quad \text{und} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

Unter Verwendung der Lagrangschen Gleichungen $\dot{p} = \frac{\partial L}{\partial q}$ erhalten wir die Hamiltonschen Gleichungen. Wenn also $q(t)$ den Lagrangschen Gleichungen genügt, so erfüllt $(p(t), q(t))$ die Hamiltonschen Gleichungen. Umgekehrt folgt aus $p = \frac{\partial L}{\partial \dot{q}}$ und den Hamiltonschen Gleichungen auch die Lagrangeschen Gleichungen. Somit sind das Lagrangsche und das Hamiltonsche System äquivalent. **q.e.d.**

Bemerkung 2.59. *Der so bewiesene Satz bezieht sich auf alle Variationsprobleme und nicht nur auf die Lagrangeschen Gleichungen der Mechanik.*

Beispiel 2.60. *Im Beispiel der Newtonschen Gleichungen hat die Lagrange-Funktion die Form $L = T - U$, wobei die kinetische Energie T eine quadratische Form in $\dot{q}$ ist*

$$T = \frac{1}{2}\sum a_{ij}\dot{q}_i\dot{q}_j \quad \textit{mit} \quad a_{ij}(q,t),$$

und die potentielle Energie nur von q abhängt $U = U(q)$.

Satz 2.61. *Unter diesen Voraussetzungen ist die Hamilton-Funktion H die Gesamtenergie $H = T + U$.*

Der Beweis basiert auf dem folgenden Lemma über die Legendre-Transformierte einer quadratischen Form:

Lemma 2.62. *Die Werte einer quadratischen Form $f(x)$ und ihrer Legendre- Transformierten $g(p)$ stimmen in den entsprechenden Punkten überein: $f(x) = g(p)$.*

Beispiel 2.63. *Für die Form $f(x) = \frac{mx^2}{2}$ ist $p = mx$ und $g(p) = \frac{p^2}{2m} = \frac{mx^2}{2} = f(x)$.*

Beweis des Lemmas: Nach dem Eulerschen Satz über homogene Funktionen zweiten Grades ist

$$f'(x)(x) = 2f(x).$$

Folglich ist $g(p(x)) = p \cdot x - f(x) = f'(x)(x) - f(x) = 2f(x) - f(x) = f(x)$. **q.e.d.**
Beweis des Satzes: Für festes q und t gilt wegen dem Lemma

$$H = p\cdot\dot{q} - L(\dot{q}) = 2T(\dot{q}) - (T(\dot{q}) - U) = T(\dot{q}) + U. \qquad \textbf{q.e.d.}$$

Beispiel 2.64. *Für die eindimensionale Bewegung gilt*

$$m\ddot{q} = -\frac{\partial U}{\partial q}.$$

In diesem Fall ist $T = \frac{m}{2}\dot{q}^2$, $U = U(q)$, $p = \dot{q}$, $H = \frac{p^2}{2m} + U(q)$, und die Hamiltonschen Gleichungen erhalten die Form

$$m\dot{q} = p \qquad\qquad \dot{p} = -\frac{\partial H}{\partial q}$$

Dieses Beispiel ermöglicht es, sich schnell zu erinnern, in welcher der beiden Hamiltonschen Gleichungen sich das Minuszeichen befindet.

Aus dem Satz über die Äquivalenz der Bewegungsgleichungen mit dem Hamiltonschen System ergibt sich eine Reihe wichtiger Folgerungen. Zum Beispiel nimmt der Energieerhaltungssatz eine einfache Gestalt an:

Korollar 2.65. *Offensichtlich gilt $\frac{dH}{dt} = \frac{\partial H}{\partial t}$. Insbesondere ist für ein System dessen Hamiltonfunktion nicht explizit von der Zeit abhängt ($\frac{\partial H}{\partial t} = 0$) die Hamiltonfunktion eine Erhaltungsgröße: $H(p(t), q(t)) = const.$*

Beweis: Betrachten wir die Änderung von H entlang der Trajektorie $H(p(t), q(t), t)$, dann ist infolge der Hamiltonschen Gleichungen

$$\frac{dH}{dt} = \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \left(\frac{\partial H}{\partial q}\right)\left(\frac{\partial H}{\partial p}\right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \qquad \textbf{q.e.d.}$$

Beim Betrachten von Zentralfeldern stellen wir fest, dass sich das Problem durch die Einführung von Polarkoordinaten auf ein eindimensionales zurückführen läßt. Es zeigt sich nun, dass jede Symmetrie eines Problems es zuläßt, ein Koordinatensystem q so auszuwählen, dass die Hamilton-Funktion von gewissen Koordinaten nicht abhängt. Damit ist es also möglich, gewisse erste Integrale zu finden, die ein solches Problem in ein Problem mit einer geringeren Anzahl von Koordinaten überführen.

Definition 2.66. *Wenn die Hamilton-Funktion $H(p_1, p_2, \ldots, p_n; q_1, \ldots, q_n; t)$ nicht von der Koordinate q_i abhängt, also $\frac{\partial H}{\partial q_i} = 0$ gilt, nennt man sie zyklisch*

Dieser Name stammt von dem Spezialfall der Winkelkoordinate φ im Zentralfeld. Offenbar ist die Koordinate q_1 dann und nur dann zyklisch, wenn sie in der Lagrange-Funktion nicht auftritt $\frac{\partial L}{\partial q_i} = 0$. Aus der Hamiltonschen Form der Bewegungsgleichung ergibt sich

Korollar 2.67. *Es sei q_1 eine zyklische Koordinate. Dann ist p_1 ein Integral der Bewegung. Die Änderung der übrigen Koordinaten mit der Zeit ist dabei dieselbe wie in einem System mit $n-1$ unabhängigen Koordinaten $q_2, \ldots, q_n$ mit der Hamilton-Funktion*

$$H(c, p_2, \ldots, p_n; q_2, \ldots, q_n; t),$$

die noch vom Parameter $c = p_1$ abhängt.

Beweis: Setzen wir $p' = (p_2, \ldots, p_n)$ und $q' = (q_2, \ldots, q_n)$, dann bekommt das Hamiltonsche System die Form

$$\frac{d}{dt}q' = \frac{\partial H}{\partial p'} \qquad \frac{d}{dt}q_1 = \frac{\partial H}{\partial p_i} \qquad \frac{d}{dt}p' = -\frac{\partial H}{\partial q'} \qquad \frac{d}{dt}p_1 = 0.$$

Die letzte Gleichung zeigt, dass $p_1 = const$ ist. Deshalb geht die Größe p_1 im Gleichungssystem für p' und q' nur als Parameter in die Hamiltonfunktion ein. Nach dem das System von $2n-2$ Gleichungen gelöst ist, nimmt die Gleichung für q_1 die Gestalt

$$\frac{d}{dt}q_1 = \frac{\partial H(p_1, p'(t), q'(t), t)}{\partial p_1}$$

an und läßt sich leicht integrieren. **q.e.d.**

Korollar 2.68. *Jedes System mit zwei Freiheitsgraden $(n = 2)$, das eine zyklische Koordinate hat, ist integrierbar.*

Beweis: In diesem Fall ist nämlich das System für p', q' eindimensional und kann unverzüglich mit Hilfe des Integrals $H(p', q') = c$ integriert werden. **q.e.d.**

2.8 Ein Satz von Liouville

Wir werden sehen, dass der Fluss des eines Hamiltonschen Vektorfeldes infolge der Hamiltonschen Gleichungen das Volumen im Phasenraum unverändert läßt. Hieraus folgt z.B., dass das Hamiltonsche System nicht asymptotisch stabil sein kann. Wir betrachten der Einfachheit halber den Fall, dass die Hamilton- Funktion die Zeit nicht explizit enthält $H = H(q, p)$.

Definition 2.69. *Der $2n$-dimensionale Raum mit den Koordinaten $p_1, \dots, p_n$ und $q_1, \dots, q_n$ wird Phasenraum genannt.*

Beispiel 2.70. *Im Fall $n = 1$ ist dies die Phasenebene mit dem Systems $\ddot{x} = -\frac{\partial U}{\partial x}$.*

Ebenso wie in diesem einfachsten Beispiel ergeben die rechten Seiten der Hamiltonschen Gleichungen ein Vektorfeld: In jedem Punkt (q, p) des Phasenraumes gibt es einen $2n$-dimensionalen Vektor $(-\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p})$. Wir setzen voraus, dass sich jede Lösung der Hamiltonschen Gleichungen auf die gesamte Zeitachse fortsetzen läßt. Dafür ist es wegen Lemma 1.40 hinreichend, dass die Niveaumengen der Funktion H kompakt sind.

Definition 2.71. *Der Phasenfluß ist die einparametrige Gruppe der Transformationen des Phasenraumes*

$$\Phi(t, \cdot) : (p(0), q(0)) \mapsto (p(t), q(t)),$$

wobei $p(t), q(t)$ die Lösungen des Systems der Hamiltonschen Gleichungen sind.

Übungsaufgabe 2.72. *Man beweise, dass $\{\Phi(t, \cdot) \mid t \in \mathbb{R}\}$ eine Gruppe ist.*

Satz 2.73 (Liouville)**.** *Der Phasenfluß von einer zweimal differenzierbaren Hamiltonfunktion H läßt das Volumen unverändert: Für jeden beliebigen Bereich D ist das Volumen von $\Phi(t, D)$ unabhängig von t gleich dem Volumen von D.*

Wir werden eine etwas allgemeinere Behauptung beweisen die ebenfalls von Liouville stammt. Gegeben sei ein System gewöhnlicher Differentialgleichungen auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$:

$$\dot{x} = F(x) \quad \text{mit} \quad F : \Omega \to \mathbb{R}^n$$

dessen Lösungen auf der gesamten Zeitachse fortgesetzt werden können. Es sei $\Phi(t, \cdot)$ der entsprechende Fluss:

$$\frac{\partial \Phi(t, x)}{\partial t} = F(\Phi(t, x)).$$

Ferner sei $D(0)$ eine Teilmenge von Ω und $v(0)$ sein Volumen. Wir definieren

$$v(t) = \text{Volumen}(D(t)) \quad \text{für} \quad D(t) = \Phi(t, D(0)).$$

Satz 2.74. *Wenn die Divergenz $\nabla \cdot F$ des Vektorfeldes F verschwindet, dann läßt $\Phi(t, \cdot)$ das Volumen unverändert: $v(t) = v(0)$.*

Wir beweisen zunächst folgedes Lemma:

Lemma 2.75. *Es gilt die Beziehung*

$$\left.\frac{dv}{dt}\right|_{t=0} = \int\limits_{D(0)} \nabla \cdot F dx^n.$$

Beweis: Für beliebiges t und nach Definition der Jacobischen Determinante gilt

$$v(t) = \int\limits_{D(0)} \left|\det \frac{\partial \Phi(t,x)}{\partial x}\right| dx^n.$$

Für $t = 0$ ist $\Phi(0,x) = x$ und $\frac{\partial \Phi(0,x)}{\partial t} = F(x)$. Daraus folgt

$$\left.\frac{\partial \Phi(t,x)}{\partial t}\right|_{t=0} = \mathbb{1} \qquad \left.\frac{d}{dt}\frac{\partial \Phi(t,x)}{\partial x}\right|_{t=0} = F'(x).$$

Wegen $\nabla \cdot F(x) = \operatorname{Spur} F'(x)$ folgt die Aussage aus dem Lemma 1.63. **q.e.d.**
Beweis des Satzes von Liouville: Wegen der Eigenschaft (ii) in Definition 1.33 folgt aus dem Lemma auch

$$\left.\frac{dv(t)}{dt}\right|_{t=t_0} = \int\limits_{D(t_0)} \nabla \cdot F dx^n.$$

Weil die Divergenz identisch verschwindet, isl also $v(t)$ konstant. **q.e.d.**

Speziell für das Hamiltonsche System gilt für zweimal differenzierbare H

$$\nabla \cdot F = \frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \frac{\partial}{\partial q}\left(\frac{\partial H}{\partial p}\right) = 0.$$

Damit ist der Satz von Liouville ebenfalls bewiesen. **q.e.d.**

Der Liouvillesche Satz erlaubt es, bei der Untersuchung mechanischer Systeme Methoden der sogenannten Ergodentheorie anzuwenden. In der Ergodentheorie benutzt man die Maßtheorie um dynamische Systeme zu untersuchen. Natürlicherweise spielen dabei solche dynamischen eine herausragende Rolle, bei denen ein invariantes Maß existiert. Wegen dem Satz von Liouville erhalten alle Hamiltonschen Vektorfelder das Volumenmaß. Zum Abschluss dieses Abschnittes zeigen wir noch einen grundlegenden Satz über solche Systeme, die ein Maß invariant lassen. Dabei betrachten wir den einfachten und allgemeinsten Fall einer Abbildung, die ein Maß invariant läßt.

Poincaréscher Wiederkehrsatz 2.76. *Es sei A eine volumenerhaltende, stetige Abbildung, die ein offenes Gebiet $\Omega \subset \mathbb{R}^n$ endlichen Volumens auf sich selbst abbildet: $A[\Omega] \subset \Omega$. Dann besitzt jede offene Teilmenge $O \subset \Omega$ eine messbare Teilmenge S mit Volumen Null, so dass alle Elemente $x \in O \setminus S$ unendlich oft nach O zurückkehren, d.h. die Folge $(A^n(x))_{n\in\mathbb{N}}$ enthält unendlich viele Elemente in O.*

Beweis: Wir bezeichnen mit der Abbildung A^{-1} die Abbildung von den messbaren Teilmengen von Ω auf sich selber, die jeder messbaren Menge das Urbild unter A zuordnet. Entsprechend bezeichen wir für alle $n \in \mathbb{N}_0$ mit A^{-n} die n-fache Iteration von A^{-1}, also die Abbildung, die jede messbare Teilmenge $X \subset \Omega$ auf das Urbild von X unter der n-fachen Iteration von A abbildet. Insbesondere sei $A^0[X] = X$ für alle messbaren Teilmengen $X \subset \Omega$. Die offene Menge O hat natütlich ein positives Volumen

$$\mathrm{vol}(O) = \int\limits_O d^n x > 0.$$

Dann sei U die Menge aller Punkte $x \in O$, für die die Folgen $(A^n(x))_{n\in\mathbb{N}}$ unendlich viele Elemente in O besitzen:

$$U = O \cap \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} A^{-n}[O].$$

Weil A volumenerhaltend ist, und Ω auf sich selber abbildet, gilt $\mathrm{vol}(\Omega) = \mathrm{vol}(A[\Omega])$. Deshalb hat das Komplement $\Omega \setminus A[\Omega]$ von dem Bild von Ω unter A in Ω kein Volumen

$$\mathrm{vol}(\Omega \setminus A[\Omega]) = \mathrm{vol}(\Omega) - \mathrm{vol}(A[\Omega]) = 0.$$

Für jede messbare Teilmenge $X \subset \Omega$ wird $A^{-1}[X]$ durch A auf das relative Komplement von $\Omega \setminus A[\Omega]$ in X abgebildet. Deshalb gilt für jede messbare Teilmenge $X \subset \Omega$ auch $\mathrm{vol}(A^{-1}[X]) = \mathrm{vol}(X)$. Daraus folgt für alle $N \in \mathbb{N}_0$

$$\mathrm{vol}(O) = \mathrm{vol}(A^{-1}[O]) = \mathrm{vol}(A^{-N}[O]) \quad \text{und} \quad \mathrm{vol}\left(\bigcup_{n=0}^{\infty} A^{-n}[O]\right) = \mathrm{vol}\left(\bigcup_{n=N}^{\infty} A^{-n}[O]\right).$$

Die Folge $(\bigcup_{n=N}^{\infty} A^{-n}[O])_{N\in\mathbb{N}_0}$ enthält immer kleinere Mengen, die alle das gleiche Volumen haben. Dann gilt für alle $N \in \mathbb{N}_0$:

$$0 = \mathrm{vol}\left(\bigcup_{n=N}^{\infty} A^{-n}[O] \setminus \bigcup_{n=N+1}^{\infty} A^{-n}[O]\right) \geq \mathrm{vol}\left(O \cap \bigcup_{n=N}^{\infty} A^{-n}[O] \setminus \bigcup_{n=N+1}^{\infty} A^{-n}[O]\right) \geq 0.$$

Das Komplement von U in O ist gleich

$$S = O \setminus U = \bigcup_{N=0}^{\infty}\left(O \cap \bigcup_{n=N}^{\infty} A^{-n}[O] \setminus \bigcup_{n=N+1}^{\infty} A^{-n}[O]\right).$$

Weil alle diese Mengen kein Volumen haben, hat auch S kein Volumen. **q.e.d.**

Wir haben in dem Beweis sogar für alle messbaren Mengen $O \subset \Omega$ gezeigt, dass die Menge aller Punkte $x \in O$, für die die Folge $(A^n(x))_{n\in\mathbb{N}}$ unendlich viele Element in O besitzt, das gleiche Volumen wir O hat.

Dieser Satz wird zum Beispiel auf den Phasenfluß eines mechanischen Systems angewandt, in dem U im Unendliche anwächst, so dass folgende Mengen kompakt sind:

$$\Omega = \{(q,p) \mid T(q,p) + U(q) \leq E\}.$$

Weil der Phasenfluss diese Mengen invariant läßt, kann man dann den Poincaréschen Wiederkehrsatz anwenden. Wie sich das Teilchen bewegt, ist zwar unbekannt. Jedoch sagt das Poincarésche Theorem eine unendlichfache Rückkehr in jede Umgebung der Ausgangslage voraus. Dies ist eine der wenigen allgemeinen Schlußfolgerungen über den Charakter der Bewegung. Die Einzelheiten der Bewegung sind schon in dem folgenden Fall nicht mehr bekannt:

$$\ddot{x} = \frac{\partial U}{\partial x} \quad \text{mit} \quad x \in \mathbb{R}^2.$$

Eine etwas paradoxe Schlußfolgerung aus dem Liouvilleschen Satz und dem Poincaréschen Wiederkehrsatz ist folgende Aussage: Wenn man die Trennwand öffnet, die eine Kammer mit Gas von der mit dem Vakuum trennt, so konzentrieren sich die Gasmoleküle nach einer gewissen Zeit wieder in der ersten Kammer. Die Lösung des Rätsels besteht darin,dass eine "gewisse Zeit" größer ist als die Zeit der Existenz des Sonnensystems. Siehe beispielsweise P. Halmos, Lectures on Ergodic Theory, AMS 2006.

Beispiel 2.77. *Es sei Ω ein Kreis und A die Drehung um den Winkel $2\pi\alpha$. Die Abbildung $t \mapsto \exp(2\pi it)$ ist ein Gruppenhomomorphismus von $\mathbb{R}$ nach $\mathbb{S}^1$. Der Kern besteht aus $\mathbb{Z}$. Deshalb ist $\mathbb{S}^1$ isomorph zu $\mathbb{R}/\mathbb{Z}$. Das Lebesguemaß induziert auf $\mathbb{R}/\mathbb{Z}$ ein translationsinvariantes Maß. Deshalb ist dieses Maß unter der Multiplikation mit $\exp(2\pi i\alpha)$ invariant. Wenn $\alpha = \frac{m}{n}$ rational ist, dann ist A^n die identische Transformation. Wenn α nicht rational ist, so folgt aus dem Poincarésche Wiederkehrsatz, dass für alle $\delta > 0$ ein $x \in (-\delta, \delta)$ existiert, so dass $x + \frac{n\alpha}{2\pi} \in (-\delta, \delta) \mod \mathbb{Z}$ für unendliche viele $n \in \mathbb{N}$ gilt. Weil für irrationale α aber gilt $n\alpha \neq 0 \mod \mathbb{Z}$, gibt es für alle $\delta > 0$ also unendlich viele $n \in \mathbb{N}$ mit $n\alpha \in (-2\delta, 0) \cup (0, 2\delta) \mod \mathbb{Z}$. Dann liegen in jedem Intervall der Länge 4δ von $\mathbb{R}/\mathbb{Z}$ undendlich viele Elemente von der Folge $(n\alpha)_{n\in\mathbb{N}}$. Deshalb folgt*

Korollar 2.78. *Ist $\frac{\alpha}{2\pi}$ keine rationale Zahl, so ist für $z \in \mathbb{C}$ mit $|z| = 1$ die Menge der komplexen Zahlen $\{\exp(ik\alpha)z \mid k \in \mathbb{N}\}$ eine in $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ dichte Teilmenge.* **q.e.d.**

Beispiel 2.79. *Es sei $\Omega = \mathbb{R}^2/\mathbb{Z}^2$ ein zweidimensionaler Torus und $\varphi = (\varphi_1, \varphi_2)$ seien Winkelkoordinaten auf ihm. Das konstante Vektorfeld $\alpha \in \mathbb{R}^2$ ist divergenzfrei und induziert auf Ω einen volumenerhaltenden Fluss:*

$$\Phi(t,\cdot) : \mathbb{R}^2/\mathbb{Z}^2 \to \mathbb{R}^2/\mathbb{Z}^2, \quad \varphi \mapsto \Phi(t,\varphi) = \varphi + t\alpha \mod \mathbb{Z}^2.$$

Um die Bahnen dieses dynamischen Systems zu untersuchen, betrachten wir für alle $\varphi_1 \in \mathbb{R}/\mathbb{Z}$ *die Bahn durch den Punkt* $(\varphi_1, 0) \in \mathbb{R}^2/\mathbb{Z}^2$. *Wenn* $\alpha_2 \neq 0$, *dann geht die Bahn für* $t = \frac{n}{\alpha_2}$ *mit* $n \in \mathbb{Z}$ *wieder durch einen Punkt von der Form* $(\varphi_1 + \frac{n\alpha_1}{\alpha_2}, n) = (\varphi_1 + \frac{n\alpha_1}{\alpha_2}, 0) \mod \mathbb{Z}^2$. *Also können wir durch die Translation* $t \mapsto t + \frac{n}{\alpha_2}$ *ein zeitdiskretes System* $\varphi_1 \mapsto \varphi_1 + \frac{n\alpha_1}{\alpha_2}$ *auf* $\varphi_1 \in \mathbb{R}/\mathbb{Z}$ *definieren. Dann folgt aus Korollar 2.78*

Satz 2.80. *Ist das Verhältnis* $\frac{\alpha_1}{\alpha_2}$ *irrational, so ist für alle* $\phi \in \mathbb{R}^2/\mathbb{Z}^2$ *die Umwicklung* $\{\Phi(t, \varphi) \mid t \in \mathbb{R}\}$ *des Torus überall dicht auf dem Torus.*

Übungsaufgabe 2.81. *Man zeige, dass im Fall eines irrationalen* ω *die Lissajoussche Figur* $\{(\cos t, \cos \omega t) \mid t \in \mathbb{R}\}$ *im Quadrat* $(x, y) \in [-1, 1] \times [-1, 1]$ *überall dicht ist.*

Übungsaufgabe 2.82. *Es sei* Ω *ein* n*-dimensionaler Torus* $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. *Es sei* $\alpha \in \mathbb{R}^n$ *und* $\Phi(t, \cdot)$ *die volumenerhaltende Transformation*

$$\Phi(t, \cdot) : \mathbb{R}^n/\mathbb{Z}^n \to \mathbb{R}^n/\mathbb{Z}^n, \quad \varphi \mapsto \varphi + at.$$

Man untersuche, unter welchen Bedingungen an α *und* φ *die folgenden Mengen dicht in* $\mathbb{R}^n/\mathbb{Z}^n$ *liegen:*

(i) $\{\Phi(t, \varphi) \mid t \in \mathbb{R}\}$.

(ii) $\{\Phi(t, \varphi) \mid t \in \mathbb{Z}\}$.

Die Transformationen aus den vorangehenden Beispielen sind eng mit der Mechanik verknüpft. Da aber das Poincarésche Theorem abstrakt ist, ist es nicht an mechanische Anwendungen gebunden.

Übungsaufgabe 2.83. *Wir betrachten jeweils die erste Ziffer der Zahlen* 2^n:

$$(n = 0, 1, 2, \ldots) : \quad 1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, \ldots$$

(i) *Finde ein maßerhaltendes zeitdiskretes dynamisches System, so dass die obige Folge durch einen Orbit dieses Systems beschrieben werden kann.*

(ii) *Zeige, dass jede Ziffer in* $\{1, \ldots, 9\}$ *in der obigen Folge undendlich oft vorkommt.*

(iii) *Zeige, dass es nur ein Wahrscheinlichkeitsmaß auf dem zeitdiskreten dynamischen Systeme in (i) gibt, das unter* $\Phi(1, \cdot)$ *invariant ist.*

Hinweise: benutze, dass das Lebesguemaß auf $\mathbb{R}$ *eindeutig dadurch charakterisiert ist, dass es translationsinvariant ist und die Menge* $[0, 1]$ *das Volumen Eins hat.*

(iv) *Berechne die Wahrscheinlichkeiten, mit der 7 und 8 in der obigen Folge auftauchen.*

2.9 Kommutator von Vektorfeldern

Jedem differenzierbaren Vektorfeld $A : \Omega \to \mathbb{R}^n$ auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ können wir zwei Objekte zuordnen:

(i) Den lokalen Fluß Φ_A auf der offenen Teilmenge $W_A \subset \mathbb{R} \times \Omega$, der die maximalen Integralkurven folgender gewöhnlichen Differentialgleichung beschreibt:

$$\frac{\partial \Phi_A(t,x)}{\partial t} = A(\Phi_A(t,x)).$$

(ii) Den Differentialoperator erster Ordnung L_A, der Richtungsableitungen in Richtung des Vektorfeldes A:

$$(L_A f)(x) = A(x) \cdot \nabla f(x) = A_1(x)\frac{\partial f(x)}{\partial x_1} + \ldots + A_n(x)\frac{\partial f(x)}{\partial x_n}.$$

Übungsaufgabe 2.84. *Man zeige, daß der Operator L_A linear ist:*

$$L_A(\lambda f + \mu g) = \lambda L_A f + \mu L_A g \quad \text{für alle} \quad \lambda, \mu \in \mathbb{R} \quad \text{und differenzierbare} \quad f, g$$

und beweise die Leibnizsche Formel

$$L_A(fg) = (L_A f)g + f(L_A g) \quad \text{für differenzierbare} \quad f, g.$$

Übungsaufgabe 2.85. (i) *Man gebe Bedingungen an, unter denen das Vektorfeld A eindeutig einen lokalen Fluss Φ_A definiert.*

(ii) *Umgekehrt gebe man Bedingungen an, unter denen ein lokaler Fluss Φ eindeutig ein Vektorfeld A bestimmt.*

(iii) *Man zeige induktiv in n, dass es für einen Differentialoperator L auf den glatten Funktionen einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$, der linear ist und die Leibnizregel erfüllt, ein glattes Vektorfeld A gibt, so dass $Lf = L_A f$ für alle Polynome f gilt.*

(iv) *Man zeige, dass (iii) im Fall $n = 1$ sogar für alle glatten Funktionen f gilt. Benutze dabei, dass für jede glatte Funktion f die Funktion $\frac{f - f(x_0)}{x - x_0}$ auch glatt ist.*

(v) *Zeige induktiv in n, dass (iii) für $\Omega \subset \mathbb{R}^n$ und alle glatten Funktionen f gilt.*

Auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ seien zwei Vektorfelder A und B gegeben. Wir sagen, dass die entsprechenden lokalen Flüsse kommutieren, wenn

$$\Phi_A(s, \Phi_B(t,x)) = \Phi_B(t, \Phi_A(s,x))$$

für alle s, t, x gilt, so dass einerseits $(t, x) \in W_B$ und $(s, \Phi_B(t, x)) \in W_A$ gilt und andererseits $(s, x) \in W_A$ und $(t, \Phi_A(s, x)) \in W_B$. Diese Menge ist immer eine Umgebung von $(s, t, x) \in \{0\} \times \{0\} \times \Omega$.

Übungsaufgabe 2.86. *Man zeige, dass die lokalen Flüsse Φ_A und Φ_B von zwei differenzierbaren Vektorfeldern A und B auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ genau dann kommutieren, wenn $\Phi_A(s, \Phi_B(t, x)) = \Phi_B(t, \Phi_A(s, x))$ für alle (s, t, x) in einer Umgebung von $\{0\} \times \{0\} \times \Omega \subset \mathbb{R} \times \mathbb{R} \times \Omega$ gilt.*

Man gebe ein Beispiel von nicht kommutierenden Flüssen an.

Um den Grad der Nichtkommutativität der beiden Flüße Φ_A und Φ_B zu messen, betrachten wir die beiden Punkte $\Phi_A(s, \Phi_B(t, x))$ und $\Phi_B(t, \Phi_A(s, x))$. Zur Abschätzung des Abstands zwischen diesen Punkten vergleichen wir die Werte für eine glatte Funktion f auf Ω. Die Differenz

$$\triangle(s, t, x) = f(\Phi_A(s, \Phi_B(t, x)) - f(\Phi_B(t, \Phi_A(s, x))$$

ist offenbar eine differenzierbare Funktion, welche für $s = 0$ und für $t = 0$ gleich 0 ist. Deshalb verschwinden bei $(s, t) = (0, 0)$ alle partiellen Ableitungen nur nach s oder nur nach t. Die erste nicht verschwindende Ableitung bei $(s, t) = (0, 0)$ ist also die gemischte zweite partielle Ableitung nach s und t. Diesen Term berechnen wir jetzt.

Lemma 2.87. *Die gemischte partielle Ableitung von $\triangle$ nach s, t an der Stelle $(s, t) = (0, 0)$ ist der Kommutator der Differentiation in den Richtungen von A und B:*

$$\left.\frac{\partial^2}{\partial t \partial s}\right|_{s=t=0} f(\Phi_A(s, \Phi_B(t, x)) - f(\Phi_B(t, \Phi_A(s, x)) = (L_B L_A f - L_A L_B f)(x).$$

Beweis: Gemäß der Definition L_A gilt

$$\left.\frac{\partial}{\partial s}\right|_{s=0} f(\Phi_A(s, \Phi_B(t, x)) = (L_A f)(\Phi_B(t, x))).$$

Wenn wir die Funktion $L_A f$ mit g bezeichnen, ergibt sich nach Definition von L_B

$$\left.\frac{\partial}{\partial t}\right|_{t=0} g(\Phi(t, x)) = (L_B g)(x),$$

und damit

$$\left.\frac{\partial^2}{\partial t \partial s}\right|_{s=t=0} (L_B L_A f)(x).$$

Wegen dem Schwarzschen Lemma kommutieren die partiellen Ableitungen, und wir können die gemischte Ableitung des anderen Terms in umgekehrter Reihenfolge ausrechnen. Dann folgt die Behauptung. **q.e.d.**

Wir betrachten jetzt den Kommutator von $L_B L_A - L_A L_B$. Auf den ersten Blick ist das ein Differentialoperator zweiter Ordnung.

Lemma 2.88. *Der Operator $L_B L_A - L_A L_B$ ist ein linearer Differentialoperator erster Ordnung, der linear ist und die Leibnizregel erfüllt.*

Beweis: Es seien $A = (A_1, \ldots, A_n)$ und $B = (B_1, \ldots, B_n)$ die Komponenten der Vektorfelder A bzw. B auf der offenen Menge $\Omega \subset \mathbb{R}^n$. Dann gilt

$$L_B L_A f = \sum_{i=1}^n B_i \frac{\partial}{\partial x_i} \sum_{j=1}^n A_j \frac{\partial}{\partial x_j} f \qquad = \sum_{i,j=1}^n B_i \frac{\partial A_j}{\partial x_i} \frac{\partial}{\partial x_j} f + \sum_{i,j=1}^n B_i A_j \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Wenn wir hiervon $L_A L_B f$ subtrahieren, verschwinden die Terme mit den zweiten Ableitungen von f, und übrig bleibt

$$(L_B L_A - L_A L_B) f = \sum_{i,j=1}^n \left(B_i \frac{\partial A_j}{\partial x_i} - A_i \frac{\partial B_j}{\partial x_i} \right) \frac{\partial f}{\partial x_j}.$$

Also gilt $[L_B, L_A] = L_C$ mit $C(x) = A'(x)B(x) - B'(x)A(x)$. Dann folgt auch, dass der Kommutator linear ist und die Leibnizregel erfüllt. **q.e.d.**

Man rechnet auch leicht allgemein nach, dass der Kommutator zweier Differentialoperatoren L und L', die beide die Leibnizregel erfüllen, wieder die Leibnizregel erfüllt:

$$\begin{aligned}(LL' - L'L)(fg) &= L((L'f)g + f(L'g)) - L'((Lf)g + f(Lg)) \\ &= (LL'f)g + (L'f)(Lg) + (Lf)(L'g) + f(LL'g) \\ &- (L'Lf)g - (Lf)(L'g) - (L'f)(Lg) - f(L'Lg) \\ &= (LL' - L'Lf)g + f(LL' - L'Lg).\end{aligned}$$

Definition 2.89. *Da jeder lineare Differentialoperator erster Ordnung durch ein Vektorfeld gegeben ist, entspricht der Operator $L_B L_A - L_A L_B$ ebenfalls einem Vektorfeld*

$$C : \Omega \to \mathbb{R}^n, \quad x \mapsto C(x) = A'(x)B(x) - B'(x)A(x) \quad \textit{und} \quad L_C = L_B L_A - L_A L_B.$$

Diese Vektorfeld heißt der Kommutator der beiden Vektorfelder B und A und wird mit $C = [B, A]$ bezeichnet.

Übungsaufgabe 2.90. *Es sei A das lineare Vektorfeld, das jedem Punkt $x \in \mathbb{R}^3$ auf das Vektorprodukt von x mit dem Vektor $a \in \mathbb{R}^3$ abbildet und B das lineare Vektorfeld, dass jeden Punkt $x \in \mathbb{R}^3$ auf das Vektorprodukt von x mit dem Vektor $b \in \mathbb{R}^3$ abbildet. Die entsprechenden Flüsse ergeben Rotationen um die Achsen a und b. Berechne den Kommutator $[A, B]$ dieser beiden Vektorfelder.*

Satz 2.91. *Der Kommutator verwandelt den linearen Raum von Vektorfeldern auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ in eine Lie Algebra.*

Beweis: Linearität und Schiefsymmetrie des Kommutators sind evident. Wir beweisen nun die Jacobische Identität. Gemäß der Definition der Poissonklammer gilt

$$\begin{aligned} L_{[A,B],C]} &= L_C L_{[A,B]} - L_{[A,B]} L_C \\ &= L_C L_B L_A - L_C L_A L_B + L_A L_B L_C - L_B L_A L_C. \end{aligned}$$

Insgesamt sind in der Summe $L_{[[A,B],C]} + L_{[[B,C],A]} + L_{[[C,A],B]}$ zwölf Terme vorhanden; jeder erscheint zweimal mit engegengesetztem Vorzeichen. **q.e.d.**

Satz 2.92. *Seien A und B differenzierbare Vektorfelder auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$. Die beiden lokalen Flüsse Φ_A und Φ_B kommutieren dann und nur dann, wenn der Kommutator $[A, B]$ der beiden Vektorfelder verschwindet.*

Beweis: Wenn die beiden lokalen Flüsse Φ_A und Φ_B kommutieren, dann folgt aus Lemma 2.87, dass auch die beiden Vektorfelder kommutieren, also $[A, B] = 0$.

Umgekehrt nehmen wir jetzt an, dass der Kommutator $[A, B]$ der beiden Vektorfelder A und B verschwindet. Wir wollen zeigen, dass die beiden lokalen Flüsse dann auch kommutieren. Offenbar ist für $(s, x) \in W_A$ und hinreichend kleine t die Abbildung $(t, x) \mapsto \Phi_A(-s, \Phi_B(t, \Phi_A(s, x)))$ ein lokaler differenzierbarer Fluss. Weil jeder differenzierbare Fluss zu einem Vektorfeld gehört, gehört auch dieser Fluss zu einem Vektorfeld, dass wir mit B_s bezeichnen wollen. Dieses Vektorfeld berechnet sich durch:

$$B_s(x) = \frac{\partial}{\partial t}\bigg|_{t=0} \Phi_A(-s, \Phi_B(t, \Phi_A(s, x))) = \Phi'_A(-s, \Phi_A(s, x))B(\Phi_A(s, x)).$$

Hier bezeichnet der Strich die Ableitung nach der Raumkoordinate in Ω. Wir wollen jetzt für $x \in \Omega$ die Abbildung $s \mapsto B_s(x)$ nach s differenzieren. Dafür bemerken wir, dass B_s eine Verkettung von drei Abbildungen ist. Die letzte $\Phi'_A(-s, \Phi_A(s, x))$ ist eine lineare Abbildung von $\mathbb{R}^n$ auf sich selber, die von s abhängt. Um sie nach s zu differenzieren, ist es einfacher die Umkehrabbildung nach s zu differenzieren. Aus $\Phi_A(-s, \Phi_A(s, x)) = x$ folgt, dass die Ableitungen dieser beiden Abbildungen zueinander invers sind, also $\Phi'_A(-s, \Phi_A(s, x))$ die Umkehrabbildung von $\Phi'_A(s, x)$ ist. Weil die partiellen Ableitungen nach s und x vertauschen, gilt

$$\frac{\partial \Phi'_A(s, x)}{\partial s} = \frac{\partial^2 \Phi_A(s, x)}{\partial s \partial x} = \frac{\partial^2 \Phi_A(s, x)}{\partial x \partial s} = \frac{\partial A(x)\Phi_A(s, x)}{\partial x} = A'(\Phi_A(s, x))\Phi'_A(s, x).$$

Dann folgt für die inverse lineare Abbildung

$$\begin{aligned}\frac{d}{ds}\Phi'_A(-s,\Phi(s,x)) &= -\Phi'_A(-s,\Phi(s,x))A'(\Phi_A(s,x))\Phi'_A(s,x)\Phi'_A(-s,\Phi(s,x))\\ &= -\Phi'_A(-s,\Phi(s,x))A'(\Phi_A(s,x)).\end{aligned}$$

Deshalb berechnet sich die Ableitung von $s \mapsto B_s(x)$ durch

$$\begin{aligned}\frac{d}{ds}B_s(x) &= \Phi'_A(-s,\Phi_A(s,x))B'(\Phi_A(s,x))A(\Phi_A(s,x))\\ &\quad - \Phi'_A(-s,\Phi_A(s,x))A'(\Phi_A(s,x))B(\Phi_A(s,x))\\ &= \Phi'_A(-s,\Phi_A(s,x))[A,B](\Phi_A(s,x)).\end{aligned}$$

Wenn also der Kommutator verschwindet, dann ist $s \mapsto B_s(x)$ für alle $x \in \Omega$ konstant. Für $s = 0$ gilt aber $B_0 = B$. Deshalb ist dann dieses Vektorfeld B_s für alle s gleich B. Dann stimmen aber auch die beiden Flüsse überein. Also folgt, dass für hinreichend kleine s die folgenden beiden Flüsse übereinstimmen:

$$\Phi_A(-s,\cdot) \circ \Phi_B(t,\cdot) \circ \Phi_A(s,\cdot) = \Phi_B(t,\cdot).$$

Dann folgt sofort $\Phi_B(t,\cdot) \circ \Phi_A(s,\cdot) = \Phi_A(s,\cdot) \circ \Phi_B(t,\cdot)$. **q.e.d.**

Dieser Satz zeigt, dass die Vektorfelder als die Lie-Algebra der Gruppe der lokalen Diffeomorphismen einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ angesehen werden kann. Dabei definiert der Kommutator von Vektorfeldern die Verknüpfung der Lie-Algebra, also die Lieklammen. Umgekehrt läßt sich die Lieklammer einer Lie-Algebra mit Hilfe des Kommutators von Vektorfeldern definieren.

Es sei G die Gruppe der linearen invertierbaren Abbildungen von $\mathbb{R}^n$ auf sich selber. Ähnliche Aussagen gelten auch für Untergruppen dieser Gruppe. Für $g \in G$ ist die Rechtstranslation R_g die Abbildung

$$R_g : G \to G, \quad R_g h = hg.$$

Die Ableitung von R_g im Punkt g bildet die $n \times n$ Matrizen auf sich selber ab. Somit entspricht jeder $n \times n$-Matrix A auch ein Vektorfeld auf G: Es besteht aus den Rechtstranslationen $(R_g)'A$ und wird rechtsinvariantes Vektorfeld genannt. Offensichtlich ist ein rechtsinvariantes Vektorfeld auf einer Gruppe eindeutig durch seinen Wert im Einselement bestimmt.

Übungsaufgabe 2.93. **(i)** *Man zeige, dass die rechtsinvarianten Vektorfelder auf G durch die Abbildungen $g \mapsto Ag$ gegeben sind, wobei A eine $n \times n$-Matrix ist.*

(ii) *Man zeige, dass der Kommutator zweier rechtsinvarianter Vektorfelder wieder ein rechtsinvariantes Vektorfeld ist, und genau dem Kommutator der entsprechenden $n \times n$-Matrizen entspricht.*

2.10 Poissonklammer

Wir werden in diesem Abschnitt sehen, dass die Hamiltonschen Vektorfelder auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ eine Unteralgebra der Lieschen Algebra aller Vektorfelder bilden. Die Hamiltonfunktionen bilden sogar selber eine Liesche Algebra: Die Operation in dieser Algebra wird Poissonklammer von Funktionen genannt. Die ersten Integrale des Hamiltonschen Phasenflusses bilden eine Unteralgebra der Lieschen Algebra der Hamilton-Funktionen.

Es sei Ω eine offene Teilmenge von dem Phasenraum $(q,p) \in \mathbb{R}^{2n}$. Jede auf Ω zweimal stetig differenzierbare Funktion $H : \Omega \to \mathbb{R}$ entspricht ein lokaler Fluss $\Phi_H(t,\cdot) = \Phi_{X(H)}$ auf Ω, nämlich der Phasenfluß des Hamiltonschen Vektorfeldes $X(H)$.

Definition 2.94. *Die Poissonklammer $\{g,f\}$ zweier differenzierbarer Funktionen f und g auf einer offenen Menge Ω des Phasenraumes $(q,p) \in \mathbb{R}^{2n}$ is definiert als die Ableitung von f in Richtung des Hamiltonschen Vektorfeldes $X(g)$:*

$$\{g,f\}(x) = \left.\frac{d}{dt}\right|_{t=0} f(\Phi_{X(g)}(t,x)).$$

Somit ist die Poissonklammer zweier Funktionen auf Ω ebenfalls eine Funktion auf Ω.

Bemerkung 2.95. *In der klassischen Dynamik definiert man üblicherweise die Poissonklammer mit einem umgekehrten Vorzeichen. Dann ist aber die Abbildung $f \mapsto X(f)$ von den Funktionen in die Hamiltonschen Vektorfelder mit der in der Differentialgeometrie üblichen Definition kein Lie Algebra Homomorphismus. Deshalb definiert Arnold beispielsweise die Lieklammer von Vektorfeldern mit dem umgekehrten Vorzeichen, als es in der Differentialgeometrie üblich ist. Weil sich die Konventionen der klassischen Mechanik mit denen der Differentialgeometrie wiedersprechen, müssen wir eine Wahl treffen. Wir haben uns für die Konventionen der Differentialgeometrie entschieden, und definieren deshalb die Poissonklammer mit dem umgekehrten Vorzeichen, als es in der klassischen Mechanik üblich ist.*

Korollar 2.96. *Die Funktion f ist dann und nur dann ein Integral des Phasenflusses der Hamilton-Funktion H, wenn ihre Poissonklammer mit H identisch verschwindet:*

$$\{H,f\} = 0.$$

q.e.d.

Wir können die Poissonklammer auch in einer etwas anderen Art definieren, wenn wir den Isomorphismus J zwischen dem Gradienten von H und den Vektorfeldern auf einer offenen Teilmenge Ω des Phasenraumes $(q,p) \in \mathbb{R}^{2n}$ verwenden. Das Hamiltonsche Vektorfeld $X(H)$ ist durch $X(H) = J\nabla H$ definiert. Hieraus ergibt sich

Korollar 2.97. *Die Poissonklammer der Funktion g und f ist gleich dem Wert der Richtungsableitung in Richtung von dem Hamiltonschen Vektorfeld $X(g)$:*

$$\{g, f\} = X(g) \cdot \nabla f. \qquad \textbf{q.e.d.}$$

Setzen wir $X(g) = J\nabla g$ in die Formel ein, so ergibt sich

Korollar 2.98. *Die Poissonklammer der Funktionen g und f ist gleich dem antisymmetrischen Produkt der Gradienten der beiden Funktionen g und f:*

$$\{g, f\} = (J\nabla g) \cdot \nabla f = -\nabla g \cdot J\nabla f = \nabla f \cdot J\nabla g. \qquad \textbf{q.e.d.}$$

Jetzt gilt offensichtlich

Korollar 2.99. *Die Poissonklammer der Funktionen f und g ist eine schiefsymmetrische Bilinearform der Funktionen f und g:*

$$\{f, g\} = -\{g, f\}, \quad \{\lambda f + \mu g, h\} = \lambda\{f, h\} + \mu\{g, h\} \quad \text{mit} \quad \lambda, \mu \in \mathbb{R}. \qquad \textbf{q.e.d.}$$

Die vorangegangenen Betrachtungen sind zwar evident. Sie führen zu nichttrivialen Schlußfolgerungen, darunter zu folgender Verallgemeinerung des Noetherschen Satzes.

Satz 2.100. *Wenn eine Hamilton-Funktion H auf einer offenen Teilmenge Ω des Phasenraumes $(q, p) \in \mathbb{R}^{2n}$ gegeben ist, und auf den Integralkurven eines Hamiltonschen Vektorfeldes $X(g)$ einer Funktion g auf Ω konstant ist, so ist g ein Integral des Systems mit der Hamilton-Funktion H.*

Beweis: Unter dieser Bedingung ist H ein Integral des Flusses $\Phi_{X(g)}$, daraus folgt $\{g, H\} = 0$ und $\{H, g\} = 0$, und F ist ein Integral des Flusses. **q.e.d.**

Die Poissonklammer zweier differenzierbarer Funktionen berechnet sich also durch

$$\begin{aligned}\{f, g\}(q, p) &= \nabla g(q, p) \cdot J\nabla f(q, p) = \nabla_p g(q, p) \cdot \nabla_q f(q, p) - \nabla_q g(q, p) \nabla_p f(q, p) \\ &= \sum_{i=1}^{n} \left(\frac{\partial g(q, p)}{\partial p_i} \frac{\partial f(q, p)}{\partial q_i} - \frac{\partial g(q, p)}{\partial q_i} \frac{\partial f(q, p)}{\partial p_i} \right).\end{aligned}$$

Insbesondere sind die Poissonklammern der Basisfunktionen $q_1, \ldots, q_n$ und $p_1, \ldots, p_n$ gegeben durch

$$\{q_i, q_j\} = \{p_i, p_j\} = 0 \text{ für } i, j = 1, \ldots, n \text{ und } -\{q_i, p_j\} = \{p_j, q_i\} = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j. \end{cases}$$

Korollar 2.101. *Die Poissonklammer erüllt für differenzierbare Funktionen f, g, h auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ folgende Leibnizregel:*

$$\{h, fg\} = \{h, f\}g + f\{h, g\}.$$

Beweis: Das folgt daraus, dass die Poissonklammer $\{h, fg\}$ gleich der Richtungsableitung von fg in Richtung des Vektorfeldes $X(h)$ ist, und die Richtungsableitung jedes Vektorfeldes die Leibnizregel erfüllt. **q.e.d.**

Satz 2.102. *Die Poissonklammer erfüllt für zweimal differenzierbare Funktionen f, g, h auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ folgende Jacobiidentität:*

$$\{h, \{g, f\}\} + \{f, \{h, g\}\} + \{g, \{f, h\}\} = 0.$$

Korollar 2.103. *(Satz von Poisson). Die Poissonklammer zweier Integrale f und g der Bewegung f, g eines Systems mit der Hamilton-Funktion H ist ebenfalls ein Integral der Bewegung.*

Beweis: Aus der Jacobischen Identität folgt

$$\{H, \{g, f\}\} = -\{f, \{H, g\}\} + \{g, \{H, f\}\} = 0.$$ **q.e.d.**

Somit können wir aus der Kenntnis zweier erster Integrale ein drittes, viertes usw. bekommen. Allerdings sind nicht alle so ermittelten Integrale grundsätzlich neu, da es nicht mehr als $2n$ unabhängige Funktionen auf $\mathbb{R}^{2n}$ gibt. Manchmal bekommen wir Funktionen von alten Integralen oder Konstanten, die gleich 0 sein können. Mitunter ergeben sich jedoch neue Integrale.

Übungsaufgabe 2.104. *Auf dem Phasenraum $\mathbb{R}^{2n}$ ist der Drehimpuls M definiert als das Vektorprodukt von q mit p. Man berechne die Poissonklammern der Komponenten von p und M.*

Daraus folgt der

Satz 2.105. *Wenn zwei Komponenten M_1 und M_2, des Drehimpulses eines mechanischen Systems Erhaltungsgrößen sind, ist es die dritte ebenfalls.* **q.e.d.**

Beweis der Jacobischen Identität: Wir betrachten die Summe

$$\{h, \{g, f\}\} = \{g, \{f, h\}\} + \{f, \{h, g\}\},$$

die eine Linearkombination von zweiten partiellen Ableitungen der zugehörigen Funktionen f, g und h ist, deren Koeffizienten Produkte der erste Ableitungen der beiden anderen Funktionen sind. Die zweiten Ableitungen von f tauchen nur in den Termen

$$\{h, \{g, f\}\} + \{g, \{f, h\}\} = \{h, \{g, f\}\} - \{g, \{h, f\}\}$$

auf und sind gleich

$$(L_{X(h)}L_{X(g)} - L_{X(g)}L_{X(h)})f = [X(h), X(g)] \cdot \nabla f.$$

Wir haben gesehen, dass diese Richtungsableitung von f in Richtung des Kommutators $[X(h), X(g)]$ nur erste Ableitungen von f enthält. Also verschwinden die Koeffizienten vor den zweiten Ableitungen von f und dann wegen der Symmetrie auch vor den zweiten Ableitungen von g und h. Dann muss dieser Ausdruck verschwinden. **q.e.d.**

Korollar 2.106. *Es seien $X(g)$ und $X(h)$ Hamiltonsche Vektorfelder zu den Hamilton Funktionen g und h. Dann ist die Richtungsableitung von f in Richtung des Kommutators $[X(h), X(g)]$ gegeben durch*

$$\{h, \{g, f\}\} - \{g, \{h, f\}\} = \{\{h, g\}, f\}$$

also die Richtungsableitung von f in Richtung von $X(\{h, g\})$. Deshalb ist der Kommutator $[X(h), X(g)]$ zweier Hamiltonscher Vektorfelder wieder ein Hamiltonsches Vektorfeld $X(\{h, g\})$ von der Poissonklammer der beiden Funktionen h und g. **q.e.d.**

Definition 2.107. *Ein linearer Unterraum einer Lieschen Algebra wird Unteralgebra genannt, wenn der Kommutator zweier beliebiger Elemente des Unterraumes in ihm liegt. Die Unteralgebra einer Lieschen Algebra ist ebenfalls eine Liesche Algebra.*

Korollar 2.108. *Die Hamiltonschen Vektorfelder auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ bilden eine Unteralgebra der Lieschen Algebra aller Vektorfelder.* **q.e.d.**

Der Poissonsche Satz über die ersten Integrale läßt sich damit umformulieren zu

Korollar 2.109. *Die Integrale des Hamiltonschen Phasenflusses bilden eine Unteralgebra der Lieschen Algebra aller Funktionen.* **q.e.d.**

Die Liesche Algebra der Hamilton Funktionen kann auf natürliche Weise auf die Liesche Algebra der Hamiltonschen Vektorfelder abgebildet werden. Dazu ordnen wir jeder Funktion H das Hamiltonsche Vektorfeld H der Hamilton-Funktion H zu.

Korollar 2.110. *Die Abbildung der Lieschen Algebra von Funktionen auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ auf die Liesche Algebra der Hamiltonschen Vektorfelder ist ein Homomorphismus von Lieschen Algebren. Sein Kern besteht aus den lokal konstanten Funktionen. Wenn Ω zusammenhängend ist, dann ist der Kern eindimensional und besteht aus den konstanten Funktionen.*

Beweis: Die Abbildung ist linear, und Jacobiidentität zeigt, daß die Abbildung die Poissonklammer von Funktionen in den Kommutator von Vektorfeldern überführt. Der Kern besteht aus Funktionen f, für die $J\nabla f$ identisch verschwindet. Da J invertierbar ist, sind das genau die lokal konstanten Funktionen. **q.e.d.**

Korollar 2.111. *Damit die Phasenflüsse der Hamilton-Funktionen f und g kommutieren, ist es notwendig und hinreichend, daß die Poissonklammer der Funktionen f und g (lokal) konstant ist.* **q.e.d.**

Definition 2.112. *Eine differenzierbare Abbildung Φ von einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ auf eine offene Teilmenge $\Omega' \subset \mathbb{R}^{2n}$ heißt kanonisch, wenn*

$$\{f \circ \Phi, g \circ \Phi\} = \{f, g\} \circ \Phi$$

für alle differenzierbaren Funktionen f und g auf Ω' gilt.

Satz 2.113. *Eine differenzierbare Abbildung Φ von einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ auf eine offene Teilmenge $\Omega' \subset \mathbb{R}^{2n}$ ist genau dann kanonisch, wenn ihre Ableitung auf ganz Ω symplektisch ist, also die Jacobimatrix zu $Sp(n)$ gehört:*

$$(\Phi'(x))^t J \Phi'(x) = J \quad \textit{für alle} \quad x \in \Omega.$$

Beweis: Seien f und g differenzierbare Funktionen auf Ω'. Dann gilt für alle $x \in \Omega$:

$$\begin{aligned}
\{f \circ \Phi, g \circ \Phi\}(x) &= -\nabla(f \circ \Phi)(x) \cdot J\nabla(g \circ \Phi)(x) \\
&= -\left((\Phi'(x))^t (\nabla f)(\Phi(x))\right) \cdot J(\Phi'(x))^t (\nabla g)(\Phi(x)) \\
&= -(\nabla f)(\Phi(x) \left(\Phi'(x) J (\Phi'(x))^t\right) (\nabla g)(\Phi(x)).
\end{aligned}$$

Die Abbildung Φ ist also kanonisch, wenn $\Phi'(x)J(\Phi'(x))^t = J$ für alle $x \in \Omega$ gilt. Wegen $J^{-1} = -J = J^t$ sind für eine symplektische Matrix A auch die Matrizen A^{-1}, JAJ^{-1} und $A^t = JA^{-1}J^{-1}$ symplektisch. Deshalb ist obige Bedingung äquivalent dazu, dass die Ableitung von Φ symplektisch ist. Wenn die Abbildung Φ kanonisch ist, dann muss obige Bedingung für alle differenzierbaren Funktionen f und g gelten. Daraus folgt auch, dass die Ableitung $\Phi'(x)$ für alle $x \in \Omega$ symplektisch ist. **q.e.d.**

Satz 2.114. *Ein stetig differenzierbares Vektorfeld auf einer einfach zusammenhängenden offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ ist genau dann ein Hamiltonsches Vektorfeld, wenn der entsprechende lokale Fluss aus kanonischen Abbildung besteht.*

Beweis: Sei Φ_F der lokale Fluss eines Vektorfeldes F auf $\Omega \subset \mathbb{R}^{2n}$. Für $t = 0$ ist der lokale Fluss die identische Abbildung, und damit offensichtlich kanonisch. Die Ableitung der Jacobimatrix $\frac{\partial \Phi_F(t,x)}{\partial x}$ nach t berechnet sich durch

$$\frac{\partial^2 \Phi_F(t,x)}{\partial t \partial x} = \frac{\partial^2 \Phi_F(t,x)}{\partial x \partial t} = \frac{\partial F(\Phi_F(t,x)}{\partial x} = F'(\Phi_F(t,x))\Phi'_F(t,x).$$

Ein Vektorfeld ist genau dann ein Hamiltonsches Vektorfeld, wenn es von der Form $J\nabla f(x)$ für eine differenzierbare Funktion f ist. Also ist für ein Hamiltonsches Vektorfeld $X(f)$ die Ableitung der Jacobimatrix des lokalen Flusses $\Phi_{X(f)}$ nach t von der

Form J mal der Hessischen von f, also das Produkt von J mit einer symmetrischen Matrix, also einer Matrix in $sp(n)$. Daraus folgt

$$\frac{\partial}{\partial t}\left(\frac{\partial \Phi_F(t,x)}{\partial x}\right)^t J\frac{\partial \Phi_F(t,x)}{\partial x} = (\Phi'(t,x))^t \left((F'(\Phi(t,x))^t J + JF'(\Phi(t,x))\right)\Phi'(t,x) = 0.$$

Also besteht der lokale Fluss aus kanonischen Abbildungen. Wenn umgekehrt der lokale Fluss aus kanonischen Abbildungen besteht, dann ist $F'(x)$ für alle $x \in \Omega$ ein Element von $sp(n)$. Daraus folgt, dass $-JF'(x)$ für alle $x \in \Omega$ symmetrisch ist. Aus Lemma 2.5 folgt dann, dass $-JF'$ die Hesse'sche einer Funktion auf Ω ist, und F ein Hamiltonsches Vektorfeld ist. **q.e.d.**

Wir bekommen noch eine weitere Verallgemeinerung des Noetherschen Satzes: Ist ein Fluß bekannt, der mit dem betrachteten kommutiert, so läßt sich ein Integral der Bewegung errechnen.

Definition 2.115. *Es sei $\Omega \subset \mathbb{R}^{2n}$ eine offene Teilmenge. Dann heißen alle Vektorfelder, die lokal Hamiltonsche Vektorfelder sind, lokale Hamiltonsche Vektorfelder. Ein Phasenfluß, der durch ein lokal Hamiltonsches Vektorfeld gegeben ist, wird lokal Hamiltonscher Fluß genannt.*

Korollar 2.116. *Ein Vektorfeld auf einer offene Teilmenge $\Omega \subset \mathbb{R}^{2n}$ ist genau dann ein lokales Hamiltonsches Vektorfeld, wenn sein Fluss aus kanonischen Transformationen besteht.*

Beweis: Jeder Punkt $x \in \Omega$ besitzt in Ω eine konvexe und damit einfach zusammenhängende Umgebung. Auf solchen Mengen haben wir die Aussage gezeigt. **q.e.d.**

Nicht alle lokalen Hamiltonschen Vektorfelder sind auch Hamiltonsche Vektorfelder:

Beispiel 2.117. *Es sei $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ der zweidimensionale Torus, dessen Punkt durch das Koordinatenpaar $(q,p) \mod \mathbb{Z}$ gegeben sind. Nun betrachten wir die Familie von Verschiebungen $\Phi(t,(q,p)) = (p+t,q)$. Das ist ein differenzierbare Fluss, der dann von einem Vektorfeld induziert ist. Für alle t sind diese Abbildungen offenbar kanonische Abbildungen. Also ist Φ Fluss eines lokalen Hamiltonschen Vektorfeldes. Aus $\dot{p} = -\frac{\partial H}{\partial q}, \dot{q} = \frac{\partial H}{\partial p}$ würde $\frac{\partial H}{\partial p} = 0, \frac{\partial H}{\partial q} = -1$, d.h. $H = -q + C$ folgen. Aber q ist nur eine lokale Koordinate auf $\mathbb{T}^2$, und die Abbildung $H : \mathbb{T}^2 \to \mathbb{R}$ mit $\frac{\partial H}{\partial p} = 0$ und $\frac{\partial H}{\partial q} = 1$ existiert nicht. Somit ist $\Phi(t,\cdot)$ kein Hamiltonscher Phasenfluß.*

Übungsaufgabe 2.118. *Man zeige, daß auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ die lokal Hamiltonschen Vektorfelder eine Unteralgebra der Lieschen Algebra aller Vektorfelder bilden. Dabei ist der Kommutator zweier lokal Hamiltonscher Vektorfelder F und*

G sogar ein Hamiltonsches Vektorfeld, dessen Hamilton Funktion eindeutig durch die Felder Vektorfelder F und G mittels der Formel

$$H(x) = F(x) \cdot JG(x)$$

gegeben ist. Somit bilden die Hamiltonschen Vektorfelder ein Ideal in der Lieschen Algebra der lokal Hamiltonschen Vektorfelder.

2.11 Integrable Systeme

Um ein System von $2n$ gewöhnlichen Differentialgleichungen zu integrieren, müssen $2n$ erste Integrale bekannt sein. Wenn ein System von Hamiltonschen Differentialgleichungen gegeben ist, stellt sich heraus, daß es in vielen Fällen genügt, nur n erste Integrale zu kennen, denn jedes von ihnen erlaubt, die Ordnung des System nicht nur um 1, sondern gleich um 2 herabzusetzen. Wir erinnern uns, daß die Funktion f und nur dann ein Integral der Bewegung eines Systems mit der Hamilton-Funktion H ist, wenn die Poissonklammer $\{f, H\}$ identisch verschwindet.

Definition 2.119. *Zwei Funktionen f und g auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ sind miteinander in Involution, wenn ihre Poissonklammer identisch verschwindet.*

Wenn mehrere Funktionen $f = (f_1, \ldots, f_m)$ auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ alle paarweise miteinander in Involution sind, dann bleiben die entsprechenden Phasenflüsse durch einen Punkt $x_0 \in \Omega$ alle innerhalb der entsprechenden Niveaumengen

$$\{x \in \Omega \mid f_1(x) = f_1(x_0), \ldots, f_m(x) = f_m(x_0)\} = f^{-1}[\{f(x_0)\} \quad \text{mit} \quad x_0 \in \Omega.$$

Wenn für eine der Funktionen $f_1, \ldots, f_m$ diese Niveaumengen mit den entsprechenden Niveaumengen der restlichen Funktionen übereinstimmen, dann ist der Funktionswert dieser Funktion an allen Punkten durch die Funktionswerte der restlichen Funktionen eindeutig bestimmt. Wenn alle diese Funktionen stetig differenzierbar sind, dann gibt es Punkte $x_0 \in \Omega$, an denen der Rang der Ableitung der Abbildung

$$f : \Omega \to \mathbb{R}^m \quad x \mapsto f(x) = (f_1(x), \ldots, f_m(x))$$

maximal ist. An diesen Punkten x_0 gibt es dann eine Auswahl dieser Funktionen, deren Gradienten alle an diesem Punkt x_0 linear unabhängig sind, und für die die Gradienten aller dieser Funktionen in x_0 von den Gradienten der ausgewählten Funktionen linear abhängen. Wegen der Stetigkeit der Gradienten, und weil die Dimension des Unterraumes von $\mathbb{R}^{2n}$, der von den Gradienten $\nabla f_1(x), \ldots, \nabla f_m(x)$ auf gespannt wird in dem Punkt $x = x_0$ maximal ist, muss dann das gleiche in einer Umgebung von x_0 gelten.

Aus dem Satz der impliziten Funktion folgt, dass die Niveaumengen der ausgewählten Funktionen durch die Punkte in einer Umgebung von x_0 einen Tangentialraum besitzen, der das orthogonale Komplement ihrer linear unabhängigen Gradienten ist. Weil alle anderen Gradienten aber lokal von diesen Gradienten linear abhängigen, verschwinden die Gradienten aller Funktionen auf den Niveaumengen der ausgwählten Funktionen. Deshalb stimmen in einer Umgebung von x_0 die Niveaumangen der ausgewählten Funktionen mit den Niveaumangen aller Funktionen überein. In diesem Fall sind in einer Umgebung von x_0 die Funktionswerte aller Funktionen eindeutig durch die Funktionswerte der ausgewählten Funktionen bestimmt. Wir sagen dann, dass alle Funktionen algebraisch nur von den ausgewählten Funktionen abhängen. Solche algebraisch abhängige Funktionen beinhalten dann keine zusätzliche Information. Deshalb wollen wir sie weglassen. Wir können dann zumindest auf der Umgebung von solchen Punkten x_0, an denen der Rang der Ableitung von der funktion $f = (f_1, \ldots, f_m)$ maximal ist, annehmen, dass all Gradienten $\nabla f_1, \ldots, \nabla f_m$ linear unabhängig sind. Offenbar kann es maximal n solcher Funktionen geben, die zusätzlich alle paarweise miteinder in Involution sind.

Wenn die Bahn eines Hamiltonschen Vektorfeldes durch einen Punkt x_0 nach längeren Zeit in die Umgebung des Punktes x_0 zurückkehrt, dann muss die Hamiltonfunktion an allen diesen Punkten die gleichen Werte annehmen. Nun kann man sich vorstellen, das die Bahn eines Hamiltonschen Vektorfeldes sogar in einer höherdimensionalen Menge einer Umgebung von x_0 dicht liegt, und deshalb die Hamiltonfunktion wegen ihrer Stetigkeit sogar auf dieser höherdimensionalen Menge konstant sein muss. Allerdings kann das nur global passieren, dadurch dass die Bahn unendlich oft in Umgebungen von x_0 zurückkehrt. Deshalb beeinflussen auch die globalen Eigenschaften der Phasenflüsse von Hamiltonschen Vektorfeldern, das lokale Verhalten der entsprechenden Hamiltonfunktionen. Um diesen Zusammenhang in den Blick zu bekommen, sollten wir annehmen, dass die Phasenflüsse der betrachteten Hamiltonschen Vektorfelder zumindest für alle Punkte aus der Niveaumenge global existieren. Solche Hamiltonschen Systeme, die maximal viele algebraisch unabhängige Integraler der Bewegung haben, die alle paareweise miteinander in Involution sind, heißen integrabel:

Definition 2.120. *Ein Hamiltonsches System einer Hamiltonfunktion H auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$ heißt vollständig integrabel, wenn es neben der Hamiltonfunktion H noch $n-1$ andere Funktionen gibt, die*

(i) *alle paarweise miteinander in Involution sind,*

(ii) *deren Gradienten linear unabhängig sind,*

(iii) *und deren Hamiltonschen Vektorfelder vollständig sind.*

Definition 2.121. *Eine Untergruppe $\Gamma \subset \mathbb{R}^n$ heißt diskret, wenn es eine offene Umgebung von $0 \in \mathbb{R}^n$ gibt, deren Schnittmenge mit Γ nur die Null enthält.*

Liouville hat nun beweisen, dass man diese Systeme tatsächlich integrieren kann. Hier ist eine exakte Formulierung dieses Satzes.

Satz 2.122. *(Liouville) Sei $\Omega \subset \mathbb{R}^{2n}$ eine offenen Teilmenge, x_0 ein Punkt in Ω, und $f = (f_1, \ldots, f_n)$ zweimal stetig differenzierbare Funktionen auf Ω,*

(i) *die alle paarweise miteinander in Involution sind,*

(ii) *deren Gradienten auf der Niveaumenge $N = f^{-1}[\{f(x_0)\}]$ linear unabhängig sind,*

(iii) *und deren Hamiltonsche Vektorfelder $X(f_1), \ldots, X(f_n)$ durch alle Punkte $x \in N$ der Niveaumenge Integralkurven auf ganz $\mathbb{R}$ besitzen.*

Dann sind die Zusammenhangskomponenten von N homöomorph zu dem Quotienten $\mathbb{R}^n/\Gamma$, wobei Γ eine diskrete Untergruppe vom $\mathbb{R}^n$ ist. Die Phasenflüsse der Hamiltonschen Vektorfelder $X(f_1), \ldots, X(f_n)$ wirken dabei wie die Translationen des $\mathbb{R}^n$.

Beweis: Wegen dem Satz der impliziten Funktion gibt es für alle $x \in N$ eine offene Umgebung U und einen Homöomorphismus Ψ von U auf eine offene Teilmenge $V \subset \mathbb{R}^{2n}$, so dass $U \cap N$ durch Ψ auf $V \cap (\mathbb{R}^n \times \{0\}) = V \cap \{x \in \mathbb{R}^2 n \mid x_{n+1} = 0, \ldots, x_{2n} = 0\}$ ist. Dabei sind Ψ und die Umkehrabbildung Ψ^{-1} sogar zweimal stetig differenzierbar. Insbesondere besitzt jeder Punkt von N dann eine Umgebung, die homöomorph zu einer offenen Teilmenge des $\mathbb{R}^n$ ist. Weil die Gradienten $\nabla f_1, \ldots, \nabla f_n$ auf der Niveaumenge lineare unabhängig sind, und weil J invertierbar ist, sind auch die Hamiltonschen Vektorfelder $X(f_1), \ldots, X(f_n)$ auf N linar unabhängig. Also besitzt die n-dimensionale Niveaumenge N n linear unabhängige Vektorfelder $X(f_1), \ldots, X(f_n)$, die alle paareweise miteinander kommutieren. Weil $f_1, \ldots, f_n$ alle paarweise miteinander in Involution sind, verschwinden die Richtungsableitungen dieser Funktionen in Richtung einer der Vektorfelder $X(f_1), \ldots, X(f_n)$ auf ganz N. Insbesondere werden diese Vektorfelder durch Ψ' auf Vektoren in $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^n \times \mathbb{R}^n$ abgebildet.

Wir definieren jetzt folgende Abbildung von $\mathbb{R}^n \times N$ nach N:

$$\Phi : \mathbb{R}^n \times N \to N, \quad (t, x) \mapsto \Phi(t, x) = \Phi_{X(f_1)}(t_1, \Phi_{X(f_n)}(t_2, \ldots, \Phi_{X(f_n)}(t_n, x) \cdots).$$

Hierbei bezeichnet $\Phi_{X(f_1)}, \ldots, \Phi_{X(f_n)}$ die Flüsse der Vektorfelder $X(f_1), \ldots, X(f_n)$. Wegen der Bedingung (iii) sind sie auf ganz N für alle $(t_1, \ldots, t_n) \in \mathbb{R}^n$ definiert. Weil die Funktionen $f_1, \ldots, f_n$ alle paareweise miteinander in Involution, sind sie auf dem Bild von $t \mapsto \Phi(t, x)$ lokal konstant. Deshalb liegt das Bild von Φ in N. Weil die Funktionen $f_1, \ldots, f_n$ zweimal stetig differenzierbar sind, sind die Vektorfelder $X(f_1), \ldots, X(f)n)$

einmal stetig differenzierbar, und die Abbildung F zweimal stetig differenzierbar. Wegen der Bedingung (i) kommutieren alle diese Flüsse miteinander. Deshalb können wir die Reihenfolge der Abbildungen $\Phi_{X(f_1)}(t_1,\cdot),\ldots,\Phi_{X(f_n)}(t_n,\cdot)$ in der Definition von F beliebig vertauschen. Aus der Flusseigenschaft (ii) der lokalen Flüsse folgt

$$\Phi(t+t',x)) = \Phi(t,\Phi(t',x)) = \Phi(t',\Phi(t,x)) \quad \text{für alle} \quad t,t' \in \mathbb{R}^n.$$

Wir zeigen jetzt, dass für alle $x \in N$ das Bild von $t \mapsto \Phi(t,x)$ in N abgeschlossen und offen ist. Dazu benutzen wir die Abbildungen Ψ, die wir oben eingeführt haben. Die Abbildungen $\Psi^{-1} \circ \Phi$ besitzen lokal invertierbare Ableitungen. Wegen dem Satz der inversen Funktion sind sie dann lokal bijektiv, also insbesondere lokal auch surjektiv. Deshalb ist das Bild von $t \mapsto \Phi(t,x)$ eine Umgebung aller seiner Elemente, und damit offen. Wenn eine Folge im Bild von $t \mapsto \Phi(t,x)$ gegen ein Element $y \in N$ konvergiert, dann besitzt y auch eine Umgebung U mit einem entsprechenden Homöomorphismus Ψ. Dann gibt es eine Umgebung von y in N, die im Bild von $t \mapsto \Phi(t,y)$ enthalten ist. Dann gibt es in dieser Umgebung auch unendlich viele Element im Bild von $t \mapsto \Phi(t,x)$. Aus der Gleichung $\Phi(t',y) = \Phi(t,x)$ folgt dann $\Phi(t-t',x) = \Phi(-t',\Phi(t,x) = \Phi(-t',\Phi(t',y) = y$. Also liegt auch y im Bild von $t \mapsto \Phi(t,x)$. Deshalb ist das Bild von $t \mapsto \Phi(t,x)$ auch abgeschlossen. Das Bild unter einer stetigen Abbildung einer zusammenhängenden Menge ist wieder zusammenhängend. Deshalb ist für alle $x \in N$ das Bild von $t \mapsto \Phi(t,x)$ zusammenhängend und in der Zusammenhangskomponente von N enthalten, die x enthält. Weil das Bild aber offen und abgeschlossen ist, stimmt es dann mit der Zusammenhangskomponente überein, die x enthält.

Sei jetzt N_0 eine Zusammenhangskomponente von N, also das Bild der Abbildung $t \mapsto \Phi(t,x)$ für alle $x \in N_0$. Als letztes zeigen wir noch, dass die beiden Mengen

$$\{t \in \mathbb{R}^n \mid \Phi(t,x) = x \text{ für ein } x \in N_0\} \qquad \{t \in \mathbb{R}^n \mid \Phi(t,x) = x \text{ für alle } x \in N_0\}$$

übereinstimmen und diskrete Untergruppen von $\mathbb{R}^n$ sind. Weil für alle $x \in N_0$ die Abbildung $t \mapsto \Phi(t,x)$ eine surjektive Abbildung von $t \in \mathbb{R}^n$ auf N_0 ist, gibt es für zwei $x,y \in N_0$ ein $t \in \mathbb{R}^n$ mit $\Phi(t,x) = y$ bzw. $\Phi(-t,y) = x$. Dann folgt aus der Kommutativität von dem Fluss Φ, dass die beiden Mengen

$$\{t \in \mathbb{R}^n \mid \Phi(t,x) = x\} \qquad\qquad \{t \in \mathbb{R}^n \mid \Phi(t,y) = y\}$$

übereinstimmen. Beide Mengen sind wegen der Kommuativität von $t \mapsto \Phi(t,\cdot)$ Untergruppen von $\mathbb{R}^n$. Weil die Abbildungen $t \mapsto \Psi^{-1} \circ \Phi(t,x)$ lokale Homöomorphismen sind, gibt es eine Umgebung der $\{0\} \in \mathbb{R}^n$, die nur das Element 0 in dieser Menge enthält. Also ist diese Menge eine diskrete Untergruppe vom $\mathbb{R}^n$. **q.e.d.**

Wir wollen jetzt solche diskreten Untergruppen vom $\mathbb{R}^n$ untersuchen.

Lemma 2.123. *Für eine Untergruppe $\Gamma \subset \mathbb{R}^n$ ist folgendes äquivalent:*

(i) *Es gibt eine offenen Umgebung der Null im $\mathbb{R}^n$, die schnittfremd zu $\Gamma \setminus \{0\}$ ist.*

(ii) *Jede beschränkte Menge im $\mathbb{R}^n$ enthält höchstens endlich viele Elemente von Γ.*

(iii) *Entweder ist $\Gamma = \{0\}$ oder es gibt ein Element kürzester Länge in $\Gamma \setminus \{0\}$.*

(iv) *Die Schnittmenge von Γ mit jedem linearen Unterraum vom $\mathbb{R}^n$ ist eine diskrete Untergruppe dieses linearen Unterraumes.*

(v) *Für $\gamma \in \Gamma \setminus \{0\}$ ist $\Gamma \cap \mathbb{R}\gamma$ eine diskrete Untergruppe von $\mathbb{R}\gamma$. Wenn $\Gamma \not\subset \mathbb{R}\gamma$, dann gibt es ein Element in $\Gamma \setminus \mathbb{R}\gamma$ mit minimalem Abstand zu $\mathbb{R}\gamma$.*

(vi) *Es gibt endlich viele linear unabhängige Elemente $\gamma_1, \ldots \gamma_m$ in Γ, die Γ erzeugen.*

Beweis: Wenn (i) gilt, dann gibt es ein $\varepsilon > 0$, so dass $B(0, \varepsilon) \cap \Gamma = \{0\}$ gilt. Weil die Differenz zweier Elemente von Γ wieder in Γ liegt, ist dann die Länge der Differenz zweier Elemente in Γ entweder gleich Null oder gößer als ε. Dann enthält jeder Ball $B(x, \frac{\varepsilon}{2})$ höchstens ein Element in Γ. Wenn A eine beschränkte Menge ist, dann ist der Abschluss von A eine kompakte Menge und besitzt eine endliche Überdeckung von Bällen vom Radius $\frac{\varepsilon}{2}$. Als enthält A höchstens endlich viele Elemente.

Wenn $\Gamma \neq \{0\}$ dann gibt es eine beschränkte Menge, die ein Element von $\Gamma \setminus \{0\}$ enthält. Also folgt (iii) aus (ii).

Aus (iii) folgt, dass es ein $\varepsilon > 0$ gibt, so dass $B(0, \varepsilon) \cap \Gamma = \{0\}$ gilt. Daraus folgt (i). Also sind (i)-(iii) äquivalent.

Weil die Schnittmenge einer offenen Umgebung der Null in $\mathbb{R}^n$ mit jedem linearen Unterraum vom $\mathbb{R}^n$ eine offene Umgebung der Null in dem linearen Unterraum ist, folgt (iv) aus (i). Die Umkehrung folgt, weil $\mathbb{R}^n$ ein linearer Unterraum vom $\mathbb{R}^n$ ist.

Es gelte (i)-(iv), und γ sei ein Element von $\Gamma \setminus \{0\}$. Dann gibt es für alle $x \in \mathbb{R}\gamma$ ein $n \in \mathbb{Z}$, so dass $x - n\gamma \in B(0, \|\gamma\|)$ gilt. Für jedes Element von $\Gamma \cap (\mathbb{R}^n \setminus \mathbb{R}\gamma)$, dessen Abstand zu $\mathbb{R}\gamma$ kleiner als δ ist, gibt es in $B(0, \delta + \|\gamma\|)$ eine Element von Γ mit dem gleichen Abstand zu $\mathbb{R}\gamma$. Wegen (ii) gibt es dann ein Element in $\Gamma \setminus \mathbb{R}\gamma)$ mit minimalem Abstand zu $\mathbb{R}\gamma$. Wenn umgekehrt (v) gilt und $\gamma \in \Gamma \setminus \{0\}$, dann gibt es ein $\varepsilon > 0$ so dass alle Elemente von $(\Gamma \setminus \{0\}) \cap \mathbb{R}\gamma$ und $\Gamma \setminus \mathbb{R}\gamma$ länger als ε sind. Also gilt (i).

Wir beweisen jetzt unter den Bedingungen (i)-(v) induktiv in m die Aussage (vi). Das erste γ_1 definieren wir als das kürzeste Element von $\Gamma \setminus \{0\}$, soweit es existiert. Dann ist $\Gamma \cap \mathbb{R}\gamma_1 = \mathbb{Z}\gamma_1$, weil es sonst in $\mathbb{R}\gamma_1$ ein kürzeres Element geben würde. Sei $P_1 : \mathbb{R}^n \to \mathbb{R}^n$ mit $x \mapsto P_1(x) = x - \frac{\langle \gamma_1, x \rangle}{\langle \gamma_1, \gamma_1 \rangle} \gamma_1$ die orthogonale Projektion auf das orthogonale Komplement von $\mathbb{R}\gamma_1$. Dann ist der Kern von P_1 gleich $\mathbb{R}\gamma_1$. Wegen (v) ist das Bild von Γ unter P_1 eine diskrete Untergruppe vom $\mathbb{R}^n$. Offenbar gilt $\Gamma \simeq \mathbb{Z}\gamma_1 \oplus$

$P_1(\Gamma)$. Wir fahren jetzt induktiv fort und erhalten dadurch induktiv linear unabhängige Elemente $\gamma_1, \ldots, \gamma_m \in \Gamma$, die Γ erzeugen. Wenn umgekehrt (vi) gilt, dann gibt es lineare Abbildungen $l_1, \ldots, l_m$ auf $\mathbb{R}^n$, so dass $\gamma = l_1(\gamma)\gamma_1 + \ldots + l_m(\gamma)\gamma_m$ für alle $\gamma \in \Gamma$ gilt. Weil $l_1, \ldots, l_m$ auf beschränkten Mengen beschränkt sind, folgt (ii). **q.e.d.**

Die Anzahl der Erzeugenden von Γ in (vi) ist die Dimension des von Γ erzeugten Unterraumes und deshalb unabhängig von der Wahl der Erzeugenden. Sie wird Rang der diskreten Untergruppe genannt. Wegen (vi) sind für diskrete Untergruppen $\Gamma \subset \mathbb{R}^n$ vom Rang n die Quotienten $\mathbb{R}^n/\Gamma$ isomorph zu $\mathbb{R}^n/\mathbb{Z}^n$, also n-dimensionale Tori. Wir können $\mathbb{R}^n$ in eine direkte Summe des Unterraumes, der von Γ erzeugt wir und dessen orthogonalem Komplement zerlegen. Der Quotient $\mathbb{R}^n/\Gamma$ ist isomorph zu der direkten Summe von dem orthogonalen Komplement und dem Quotienten des von Γ erzeugten Unterraumes durch Γ. Wenn der Rang von Γ nicht n ist, ist also $\mathbb{R}^n/\Gamma$ nicht kompakt.

Korollar 2.124. *Unter den Bedingungen von dem Satz von Liouville ist eine Zusammenhangskomponente der Niveaumenge N genau dann kompakt, wenn sie isomorph zu einem Torus ist.* **q.e.d.**

Zum Abschluss diese Abschnittes wollen wir zeigen, dass alle Hamiltonschen Systeme lokal vollständig integrabel sind:

Satz 2.125. *Sei $\Omega \subset \mathbb{R}^{2n}$ eine offene Teilmenge und H eine zweimal stetig differenzierbare Funktion auf Ω. Dann besitzt jeder Punkt $x \in \Omega$, an dem ∇H nicht verschwindet, auf einer Umgebung $U \subset \Omega$ von x zweimal stetig differenzierbare Funktionen $f_1 = H, f_2 \ldots, f_n$, die auf U die Bedingungen (i)-(ii) aus der Definition 2.121 erfüllen.*

Bevor wir diesen Satz beweisen können brauchen wir ein wenig Vorbereitung. Zwei dynamische Systeme $\Phi : \mathbb{R} \times \Omega \to \Omega$ und $\tilde{\Phi} : \mathbb{R} \times \tilde{\Omega} \to \tilde{\Omega}$ sind offenbar äquivalent, wenn es einen Homöomorphismus $\Psi : \Omega \to \tilde{\Omega}$ gibt, so dass Folgendes gilt:

$$\tilde{\Phi}(t, \cdot) = \Psi \circ \Phi(t, \cdot) \circ \Psi^{-1} \quad \text{für alle} \quad t \in \mathbb{R}.$$

Aus der Gruppeneigenschaft folgt, dass das äquivalent dazu ist, dass das für kleine t gilt. Entsprechend induzieren zwei Vektorfelder F und $\tilde{F}$ auf den offenen Teilmengen Ω und $\tilde{\Omega}$ äquivalent lokale Flüsse, wenn für alle $x \in \tilde{\Omega}$ und hinreichend kleine t

$$\Phi_{\tilde{F}}(t, x) = \Psi(\Phi_F(t, \Psi^{-1}(x)))$$

gilt. Diese Gleichung ist offenbar äquivalent dazu, dass Ψ die Integralkurven von F auf die Integralkurven von $\tilde{F}$ abbildet. Wenn wir das nach t differenzieren folgt

$$\tilde{F}(x) = \Psi'(\Psi^{-1}(x))F(\Psi^{-1}(x)).$$

Diese Formel beschreibt wie sich Vektorfelder unter differenzierbaren Homöomorphismen transformieren. Unter den Koordinatenvektorfeldern einer offenen Teilmenge $\Omega \subset \mathbb{R}^n$ verstehen wir die konstanten Vektorfelder auf die Einheitsvektoren des $\mathbb{R}^n$.

Lemma 2.126. *Sei $\Omega \subset \mathbb{R}^n$ eine offene Teilmenge und $X_1, \ldots, X_m$ kommutierende stetig differenzierbare Vektorfelder auf Ω. Alle Punkte $x \in \Omega$, in denen $X_1, \ldots, X_m$ linear unabhängig sind, besitzen eine Umgebung $U \subset \Omega$ mit einem zweimal stetig differenzierbaren Homöomorphismus Φ von einer offenen Teilmenge $O \subset \mathbb{R}^n$ auf U, das die ersten m Koordinatenvektorfelder von O in die Vektorfelder $X_1, \ldots, X_m$ transformiert.*

Beweis: Sei $x_0 \in \Omega$ ein Punkt, an dem die Vektorfelder $X_1, \ldots, X_m$ linear unabhängig sind. Dann gibt es in $(t, x + x_0) \in \mathbb{R}^n \times \Omega$ eine Umgebung von $(0, 0)$, auf der die Flüsse

$$(t, x) \mapsto \Phi_{X_1}(t_1, \Phi_{X_2}(t_2, \ldots, \Phi_{X_m}(t_m, x + x_0) \cdots)$$

definiert sind. Wir wählen einen lineare Unterraum V der Dimension $n - m$ vom $\mathbb{R}^n$, dessen Schnittmenge mit der linearen Hülle von $X_1(x_0), \ldots, X_m(x_0)$ gleich $\{0\}$ ist. Die Einschränkung der obigen Abbildung auf die Schnittmenge der Umgebung von $(0, 0)$ mit $(t, v) \in \mathbb{R}^n \times V$ definiert eine zweimal stetig differenzierbare Abbildung nach Ω, deren Ableitung im Punkt $(t, v) = (0, 0)$ invertierbar ist. Wegen dem Satz der inversen Funktion gibt es dann einen zweimal stetig differenzierbaren Homöomorphismus Φ von einer Umgebung O von $(0, 0) \in \mathbb{R}^n \times V$ auf eine Umgebung U von x_0 mit zweimal stetig differenzierbarer Umkehrabbildung. Weil die lokalen Flüsse $\Phi_{X_1}(t_1, \cdot), \ldots, \Phi_{X_m}(t_m, \cdot)$ kommutieren, werden die Integralkurven der Vektorfelder $X_1, \ldots, X_m$ durch Φ^{-1} auf die Integralkurven der ersten m Koordinatenvektorfelder von $\mathbb{R}^m \times V$ abgebildet. Weil die Vektorfelder $X_1, \ldots, X_m$ durch Φ^{-1} auf stetig differenzierbare Vektorfelder abgebildet werden, die eindeutig durch ihre Integralkurven bestimmt sind, transformiert Φ die ersten m Koordinatenvektorfelder von $O \subset \mathbb{R}^m \times V \simeq \mathbb{R}^n$ in diese Vektorfelder.**q.e.d.**

Lemma 2.127. *Seien $f_1, \ldots, f_m$ zweimal stetig differenzierbare Funktionen auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$, die alle paarweise miteinander in Involution sind. Dann besitzen alle Punkte $x \in \Omega$, an denen $\nabla f_1(x), \ldots, \nabla f_m(x)$ lineare unabhängig sind, eine Umgebung U und einen zweimal stetig differenzierbaren Homöomorphismus Ψ auf eine offene Teilmenge $O \subset \mathbb{R}^{2n}$, so dass $X(f_1), \ldots, X(f_m)$ durch Ψ auf die ersten m Koordinatenvektorfelder abgebildet werden, und $f_1|_U = \Psi_{2n-m+1}, \ldots, f_m|_U = \Psi_{2n}$ gilt.*

Beweis: Die Abbildung $f = (f_1, \ldots, f_m) : \Omega \to \mathbb{R}^m$ hat in einem Punkt $x_0 \in \Omega$, in dem $\nabla f_1(x_0), \ldots, \nabla f_m(x_0)$ linear unabhängig sind, eine Ableitung vom Rang m. Offenbar ist die Aussage äquivalent zu der analogen Aussage für die Funktion $f - f(x_0)$. Deshalb können wir $f(x_0) = 0$ annehmen. Sei K der Kern von $f'(x_0)$ als linearer Unterraum von $\mathbb{R}^{2n}$. Sei jetzt W ein m dimensionaler Unterraum von $\mathbb{R}^{2n}$, den $f'(x_0)$

isomorph nach $\mathbb{R}^m$ abbildet. Dann gibt es eine offene Teilmenge $\tilde{\Omega}$ von $K \times W$, die durch $(l, w) \mapsto x_0 + k + w$ diffeomorph auf Ω abgebildet wird. Im Folgenden identifizieren wir $\tilde{\Omega}$ mit Ω. Die zweimal stetig differenzierbare Abbildung

$$\tilde{\Psi} : \tilde{\Omega} \to W \times \mathbb{R}^m, \quad (k, w) \mapsto (w, f(x_0 + k + w)$$

hat im Punkt $(0, 0) \in \tilde{\Omega}$ eine invertierbare Ableitung. Wegen dem Satz der inversen Funktion bildet Ψ eine Umgebung $\tilde{U} \subset \tilde{\Omega}$ bijektiv mit zweimal stetig differenzierbarer Umkehrabbildung Ψ^{-1} auf eine offene Umgebung $\tilde{O}$ von $(0, 0)$ ab.

Weil die Funktionen $f_1, \ldots, f_m$ paarweise miteinander in Involution sind, kommutieren die Hamiltonschen Vektorfelder $X(f_1), \ldots, X(f_m)$, und ihre Integralkurven verbleiben in den entsprechenden Niveaumengen von f. Wenn wir wie oben beschrieben Ω mit $\tilde{\Omega}$ identifizieren, dann werden durch $\tilde{\Psi}$ diese Vektorfelder in stetig differenzierbare kommutierende Vektorfelder $\tilde{X}_1, \ldots, \tilde{X}_m$ auf $\tilde{O}$ transformiert. Die Werte dieser Vektorfelder bei $(0, 0)$ liegen also in K und spannen einen m dimensionalen Unterraum X von K auf. Sei $\tilde{V}$ ein $2n - 2m$-dimensionaler Unterraum von K, dessen Schnittmenge mit X gleich $\{0\}$ ist. Dann können wir $\tilde{O}$ mit einer offenen Teilmenge von $X \times \tilde{V} \times \mathbb{R}^m$ identifizieren. Wir wenden das vorangehende Lemma auf die stetig differenzierbaren Vektorfelder $\tilde{X}_1, \ldots, \tilde{X}_m$ von $\tilde{O}$ an, und wählen in dem Beweis dort die Menge V als $\tilde{V} \times \mathbb{R}^m$. Dann erhalten wir eine Homöomorphismus Φ von einer offenen Umgebung O von $(0, 0) \in X \times V \times \mathbb{R}^m$ auf eine offene Umgebung von $(0, 0) \in \tilde{O}$. Sei jetzt U das Urbild dieser offenen Umgebung von $(0, 0)$ in $\tilde{O}$ unter $\tilde{\Psi}$, aufgefasst als Teilmenge von Ω, und $\Psi = \Phi^{-1} \circ \tilde{\Psi}$. Dann ist Ψ eine bijektive zweimal stetig differenzierbare Abbildung von U auf O, mit zweimal stetig differenzierbarer Umkehrabbildung. Die Hamiltonschen Vektorfelder $X(f_1), \ldots, X(f_m)$ werden durch Ψ in die ersten m Koordinatenvektorfelder transformiert. Weil die Funktionen $f_1, \ldots, f_m$ längs der Integralkurven dieser Vektorfelder konstant sind, läßt Φ^{-1} die letzten m Koordinaten invariant und es gilt $f|_U = \Psi_{2n_m+1}, \ldots, f_m|_U = \Psi_{2n}$. **q.e.d.**

Beweis von Satz 2.125: Wenn in der Situation des vorangehenden Lemmas m kleiner als n ist, dann sind die Funktionen $\Psi_{m+1} \ldots, \Psi_{2n-m}$ konstant auf den Integralkurven der Hamiltonschen Vektorfelder $X(f_1) \ldots, X(f_m)$. Deshalb sind diese Funktionen mit allen Funktionen $f_1, \ldots, f_m$ in Involution. Deshalb gibt es dann lokal ein zusätzliches Integral der Bewegung. Außerdem sind die Ableitungen von diesen Funktionen auch lineare unabhängig von den Ableitungen von $f_1, \ldots, f_m$. Wir wenden also das vorangehende Lemma mehrfach an und erhalten auf einer immer kleineren Umgebung eines Punktes $x_0 \in \Omega$ mit $\nabla H(x_0) \neq 0$ eine immer größere Anzahl an Integralen der Bewegung.**q.e.d.**

Im Fall von n glatten Funktionen $f_1, \ldots, f_n$ auf einer offenen Menge $\Omega \subset \mathbb{R}^{2n}$, die alle paarweise miteinander in Involution sind, besitzt wegen dem vorangehenden Lemma jeder Punkt $x_0 \in \Omega$, an dem die Gradienten $\nabla f_1, \ldots, \nabla f_n$ linear unabhängig sind, einen Diffeomorphismus einer offenen Umgebung U von x_0 auf eine offene Teilmenge

$O \subset \mathbb{R}^{2n}$, die die Hamiltonschen Vektorfelder in die ersten n Koordinatenvektorfelder transformiert und deren letzte n Koordinaten mit den Funktionen übereinstimmen. Diese Abbildungen müssen nicht kanonisch sein.

Übungsaufgabe 2.128. *Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine zweimal stetig differenzierbare Abbildung und $\Psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ die Abbildung mit $(q,p) \mapsto \Psi(q,p) = (q + f(p), p)$.*

(i) *Zeige, dass Ψ die Hamiltonschen Vektorfelder der letzten n Komponenten p von $(q,p) \in \mathbb{R}^{2n}$ auf die ersten n Koordinatenvektorfelder von $\mathbb{R}^{2n}$ abbildet.*

(ii) *Charakterisiere die Funktionen f, für die Ψ kanonisch ist.*

Übungsaufgabe 2.129. *Auf $\mathbb{R}^{2n}$ seien folgende Funktionen $f_1, \ldots, f_n$ gegeben*

$$f_i : \mathbb{R}^{2n} \to \mathbb{R}, \quad (q,p) \mapsto f_i(q,p) = q_i^2 + p_i^2.$$

(i) *Zeige, dass alle diese Funktionen paareweise miteinander in Involution sind.*

(ii) *Bestimme die Menge aller Punkte an denen die Gradienten dieser Funktionen linear unabhängig sind. Zeige, dass diese Menge eine Vereinigung von Niveaumengen ist, und charakterisiere die entsprechenden Funktionswerte.*

(iii) *Gebe eine möglichst große Teilmenge von Ω an, auf der die Hamiltonfunktion $H = f_1 + \ldots + f_n$ vollständig integrabel ist.*

(iv) *Zeige, dass alle Niveaumengen Tori sind.*

(v) *Zeige dass die Hamiltonschen Vektorfelder $X(f_1), \ldots, X(f_n)$ vollständig sind.*

2.12 Winkelwirkungsvariable

In diesem Abschnitt wollen wir für ein vollständig integrables System eine kanonische Abbildung konstruieren, die die Hamiltonschen Vektorfelder in die ersten n Koordinatenvektorfelder überführt und deren letzte n Einträge die Algebra der Integrale der Bewegung erzeugt. Dabei nehmen wir an, dass die Niveaumengen kompakt sind.

Lemma 2.130. *Seien $f = (f_1, \ldots, f_n)$ zweimal stetig differenzierbare Funktionen auf einer offenen Teilmenge $\Omega \subset \mathbb{R}^{2n}$, deren Gradienten auf einer kompakten Zusammenhangskomponente N_0 einer Niveaumenge $f^{-1}[\{f(x_0)\}]$ linear unabhängig sind. Dann gibt es einen zweimal stetig differenzierbaren Homöomorphismus Ψ mit zweimal stetig differenzierbarer Umkehrabbildung Ψ^{-1} von einer offene Teilmenge $U \subset \Omega$, die die kompakte Zusammenhangskomponente N_0 enthält, nach $\mathbb{R}^n/\mathbb{Z}^n \times O$. Hier ist O eine offene Teilmenge von $\mathbb{R}^n$. Die letzten n Komponenten von ψ stimmen mit der Einschränkungen $f|_U$ überein.*

Beweis: Wir schränken f auf eine offene Umgebung von N_0 ein, deren Schnittmenge mit der Niveaumenge $f^{-1}[\{f(x_0)\}]$ gleich N_0 ist. Jeder Punkt $x \in N_0$ besitzt eine Umgebung, wie wir sie im Lemma 2.127 konstruiert haben. Weil die Niveaumenge kompakt ist, gibt es eine endliche Teilüberdeckung. Dann gibt es eine offene Umgebung O von $f(x_0) \in \mathbb{R}^n$, so dass für alle $y \in O$ die Niveaumengen $f^{-1}[\{y\}]$ in dieser endlichen Teilüberdeckung enthalten sind. Alle diese Niveaumengen sind dann kompakt, und die Hamiltonschen Vektorfelder sind auf ihnen vollständig. Wegen dem Satz von Liouville, sind alle diese Niveaumengen isomorph zu $\mathbb{R}^n/\Gamma(y)$. Die diskreten Untergruppen $\Gamma(y)$ haben dann für alle $y \in O$ den Rang n. Offenbar hängen die Erzeugenden von $\Gamma(y)$ zweimal stetig differenzierbar von $y \in O$ ab. Deshalb gibt es einen Isomorphismus von den Niveaumengen auf $\mathbb{R}^n/\mathbb{Z}^n$, der zweimal stetig differenzierbar von $y \in O$ abhängt. Dann folgt die Aussage aus dem Lemma 2.127 **q.e.d.**

Man kann die letzten n Komponenten dieser Abbildung Ψ so verändern, dass es eine kanonische Transformation wird. Um das zu zeigen benutzt man den Formalismus des Erzeugenden Funktionals einer kanonischen Abbildung.

Lemma 2.131. *Sei $(q,p) \mapsto (Q,P)$ eine stetig differenzierbare bijektive Abbildung mit stetig differenzierbarer Umkehrabbildung von einer offenen einfach zusammenhängenden Teilmenge $\Omega \subset \mathbb{R}^{2n}$ auf eine offene einfach zusammenhängende Teilmenge $\Omega' \subset \mathbb{R}^{2n}$. Das Vektorfeld*

$$F = (p,0) + \sum_{i=1}^{n} Q_i \nabla P_i$$

ist genau dann ein Gradientenvektorfeld ist, wenn $(q,p) \mapsto (Q,P)$ kanonisch ist.

Beweis: Wegen Lemma 2.5 ist ein Vektorfeld F genau dann ein Gradientenvektorfeld, wenn die Ableitung F' eine symmetrische Matrix ist. Für das Vektorfeld $F = (p,0)$ gilt offenbar $F' - (F')^t = J$.

Für das Vektorfeld $F = \sum_{i=1}^n Q_i \nabla P_i$ fallen in $F' - (F')^t$ die zweiten Ableitungen von P wegen dem Schwarzen Lemma weg. Dann gilt $F' - (F')^t = -(\Phi')^t J \Phi'$ (Übungsaufgabe), wobei Φ' die Jacobimatrix von $(q,p) \mapsto (Q,P)$ ist.

Dann folgt, dass das Vektorfeld $F = (p,0) + \sum_{i=1}^n Q_i \nabla P_i$ genau dann ein Gradientenvektorfeld ist, wenn die Jacobimatrix Φ' der Abbildung $(q,p) \mapsto (Q,P)$ auf Ω in $Sp(n)$ liegt. Dann folgt die Aussage aus dem Satz 2.113. **q.e.d.**

Wir schreiben für das Gradientenvektorfeld $F = (p,0) + \sum_{i=1}^n Q_i \nabla P_i$ der Funktion S auch $dS = p \cdot dq + Q \cdot dP$, und nennen solche Gradientenvektorfelder auch Einsformen. Deshalb definiert $p \cdot dq + Q \cdot dP$ auf einer einfach zusammenhängenden Teilmenge eine Funktion S. Wir nehmen jetzt an, dass diese Funktion nur von den Variablen q und P abhängt. Dann folgt aus $dS(q,P) = p \cdot dq + Q \cdot dP$

$$p = \frac{\partial S}{\partial q} \quad \text{und} \quad Q = \frac{\partial S}{\partial P}.$$

Wir konstruieren in der Situation von Lemma 2.130 eine kanonische Transformation $(q,p) \mapsto (Q,P)$, deren erste n Komponenten Q mit den ersten n Komponenten von Ψ übereinstimmen, und deren letzte n Komponenten nur von f abhängen. Wir nehmen dafür an, dass in einem Punkt $x_0 \in N_0$ die Ableitung von f nach p invertierbar ist. Das können wir durch eine geeignete Wahl von Koordinaten immer erreichen. Das entsprechende erzeugende Funktional $S(q,P)$ soll nur von q und P abhängen. Weil P dann die Integrale der Bewegung darstellt, sind für festes P auch die Funktionen $(f_1,\ldots,f_n)$ und damit auch die Niveaumengen fixiert. Die Funktion S stimmt auf den Niveaumengen mit dem Wegintegral $\int pdq$ überein. Wir wollen die Funktion S zunächst auf den Niveaumengen konstruieren.

Übungsaufgabe 2.132. *Zeige in der Situation von Lemma 2.130, dass es für alle $y \in O$ eine zweimal stetig differenzierbare Funktion S auf $\mathbb{R}^n$ gibt, so dass*

$$\frac{\partial S(t)}{\partial t_i} = X(f_i)(\Phi(t,y)) \cdot (p,0)(\Phi(t,y)) \quad \textit{für alle } t \in \mathbb{R}^n \textit{ und } S(0) = 0 \textit{ gilt.}$$

Dabei ist Φ die Wirkung von $\mathbb{R}^n$ auf der Niveaumenge $f^{-1}[\{f(y)\}]$, die wir im Satz von Liouville konstruiert haben.

Satz 2.133. *In der Situation von Lemma 2.130 wählen wir Erzeugende $\gamma_1(y),\ldots,\gamma_n(y)$ der diskreten Untergruppe von der Niveaumenge durch y. Wir definieren die Funktionen*

$$P_1(y) = S(\gamma_1(y)) \quad ,\ldots, \quad P_n(y) = S(\gamma_n(y)).$$

Wenn die Ableitung der Abbildung P in $f(x_0)$ invertierbar ist, und die Ableitung $\frac{\partial f}{\partial p}$ im Punkt x_0 invertierbar ist, dann ist die Abbildung $(\mathbb{1}_{\mathbb{R}^n/\mathbb{Z}^n} \times P) \circ \Psi$ eine kanonische Abbildung von einer offenen Umgebung von N_0 auf das kartesische Produkt von $\mathbb{R}^n/\mathbb{Z}^n$ mit einer offenen Umgebung von $f(x_0 \in O$.

Beweis: Weil $\frac{\partial f}{\partial p}$ im Punkt x_0 invertierbar ist, ist eine Umgebung von x_0 eindeutig durch $p(q,f)$ und q parametriesierbar. Weil die Ableitung von P im Punkt $f(x_0)$ invertierbar ist, ist eine Umgebung von x_0 auch durch $p(q,P)$ und q parametrisierbar. Dort definieren wir die Funktion S durch $S(q,P) = \int_{q_0}^{q} p(q,P)dq$. Für festes P stimmt dieses S mit dem aus der obigen Übungsaufgabe überein, und ist mehrdeutig. Die Differenz zwischen den Zweigen sind ganzzahlige Linearkombinationen von den Einträgen von P. Deshalb sind die Ableitungen der Differenzen nach P lokal konstant. Die verschiedenen Zweige von Q unterscheiden sich dann nur um ganze Zahlen. Insbesondere sind die Differenzen von den Werten von Q an den Erzeugenden $\gamma_1(y),\ldots,\gamma_n(y)$ und 0 gleich den kanonischen Einheitsvektoren von $\mathbb{R}^n$. **q.e.d.**

Die Komponenten der Variablen Q werden Winkelvariable genannt und die Komponenten der Variable P Wirkungsvariable.