

Fachbereich Informatik und Mathematik
ISMI - Institut für Stochastik
& Mathematische Informatik

Analysis III

SS 2011

H. Dinges

14. Juli 2011

Inhaltsverzeichnis

1	Stetige Differenzierbarkeit	1
1.1	Partielle Ableitungen und Stammfunktionen	1
1.2	Krummlinige Koordinaten; Tangentialvektoren	10
1.3	Extremwert-Aufgaben	22
1.4	Der Begriff der Mannigfaltigkeit	31
1.5	Umformungen für spezielle Differentialgleichungen	44
1.6	Lösungskurven; dynamische Systeme	54
1.7	Rückblick und Ausblick	64
2	Einschlägige Themen der Linearen Mathematik	67
2.1	Lineare Differentialgleichungen	67
2.2	Die Idee des Randoperators, Orientierung	83
2.3	Tensoren und ihre Indizes	90
2.4	Die Algebra der schiefsymmetrischen Formen	104
3	Differentialformen	113
3.1	Vorbereitungen	113
3.1.1	Topologische Vorbemerkung: Zerlegungen der Eins	113
3.1.2	Gerichtete Volumina	115
3.1.3	Maßtheoretische Vorbereitung: Wechsel der Integrationsvariablen	116

1 Stetige Differenzierbarkeit

1.1 Partielle Ableitungen und Stammfunktionen

Der Begriff der Stammfunktion auf einem Intervall

In der Analysis I haben wir uns mit stetigen reellwertigen Funktionen $F(\cdot)$ auf einem offenen Intervall (a, b) befasst, zu welchen eine stetige Funktion $f(\cdot)$ existiert, sodass gilt

$$\lim_{h \rightarrow 0} \frac{1}{h} (F(\cdot + h) - F(\cdot)) = f(\cdot) \quad \text{gleichm\u00e4\u00dfig auf Kompakten.}$$

Der lokal gleichm\u00e4\u00dfige Limes der stetigen ‘Differenzenquotienten’ ist nat\u00fcrlich stetig. Wir haben eine Funktion der beschriebenen Art eine **stetig differenzierbare** Funktion genannt; und wir haben notiert $f(\cdot) = F'(\cdot) = \frac{d}{dx} F(\cdot)$.

Die traditionelle Herangehensweise an den Begriff der Differenzierbarkeit weiss nichts von lokal gleichm\u00e4\u00dfiger Konvergenz. Man beginnt anders

Definition 1.1. Eine stetige Funktion $F(\cdot)$, die in einer Umgebung $U = (a, b)$ von \tilde{x} definiert ist, heisst differenzierbar im Punkt \tilde{x} , wenn eine Zahl \tilde{a} existiert, sodass

$$F(\tilde{x} + h) - F(\tilde{x}) = \tilde{a} \cdot h + o(|h|) \quad \text{f\u00fcr } h \rightarrow 0.$$

Die Zahl \tilde{a} heisst die Ableitung (oder der Differentialquotient) der Funktion F im Punkt \tilde{x} und man notiert $\tilde{a} = F'(\tilde{x})$.

Traditionell wird eine Funktion differenzierbar in U genannt, wenn sie in jedem $\tilde{x} \in U$ differenzierbar ist. Wenn $F'(\cdot)$ zudem stetig ist, dann heisst sie stetig differenzierbar in U . Die beiden Definitionen der stetigen Differenzierbarkeit einer Funktion auf einem offenen Intervall stimmen \u00fcberein, wie wir sehen werden.

In der Integrationstheorie (der Analysis II) haben wir einen Begriff der Stammfunktion kennen gelernt, welcher den klassischen Begriff verallgemeinert, n\u00e4mlich

Definition 1.2. Eine stetige Funktion auf einem offenen Intervall $U = (a, b)$ heisst absolutstetig, wenn eine lokal integrable Funktion $f(\cdot)$ existiert, sodass gilt

$$F(d) - F(c) = \int_c^d f(x) dx \quad \text{f\u00fcr alle kompakten Intervalle } [c, d] \subset (a, b).$$

Eine solche Funktion heisst eine Stammfunktion der lokal integrablen Funktion f .

Warnung: Der traditionelle Begriff der in U differenzierbaren Funktion ist ein Begriff, der sich als ausserordentlich ungl\u00fccklich und unbrauchbar erweist. Er erinnert m\u00f6glicherweise an den Begriff der Holomorphie bei komplexwertigen Funktionen einer komplexen Variablen. Im Gegensatz zum Begriff der holomorphen Funktion ist er aber ein Fremdk\u00f6rper in jeder vern\u00fcnftigen Theorie der Differenzierbarkeit. Ein Beispiel mag diese

These stützen: Die Funktion $F(x) = x^2 \cdot \sin \frac{1}{x^2}$ für $-1 < x < +1$ ist in allen Punkten differenzierbar; sie ist aber in keiner Umgebung des Nullpunkts absolutstetig.

These: Gute Begriffe sind einerseits der Begriff der stetigen Differenzierbarkeit einer Funktion, den wir unten für Funktionen mehrerer reeller Veränderlicher herausarbeiten werden, und andererseits der Begriff der Absolutstetigkeit (oder Totalstetigkeit), den in der Maßtheorie sehr weitgehend verallgemeinert wird.

Begriffe, die nicht weit führen, aber dennoch (seit Cauchy's Zeit) in den Anfängervorlesungen für Mathematiker quälend diskutiert werden, sind 'Differenzierbarkeit in jedem Punkt eines Intervalls' sowie 'Differenzierbarkeit in Lebesgue-fastallen Punkten eines Intervalls'.

Satz 1.1.1. *Wenn $F(\cdot)$ im Intervall $U = (a, b)$ stetig differenzierbar ist, dann ist sie in U auch absolutstetig.*

Beweis. *Es sei $f(\cdot) = F'(\cdot)$ und für ein beliebig gewähltes x_0*

$$G(x) = F(x) - F(x_0) - \int_{x_0}^x f(u) \, du.$$

Es gilt offenbar $G'(x) = 0$ für alle $x \in (a, b)$. Der Beweis ist erbracht, wenn wir zeigen, dass eine Funktion $G(\cdot)$, die überall in (a, b) die Ableitung 0 hat, eine Konstante ist.

Wir beweisen einen etwas schärferen Satz, welcher in Lehrbüchern der elementaren Differential- und Integralrechnung unter dem Namen 'Satz von Rolle' herausgestellt wird oder auch als 'Mittelwertsatz der Differentialrechnung'.

Satz 1.1.2. *Die stetige Funktion $F(\cdot)$ sei in jedem Punkt $\tilde{x} \in (a, b)$ differenzierbar. Wenn*

$$\frac{1}{d-c}(F(d) - F(c)) = \theta \quad \text{für ein Intervall } [c, d] \subset (a, b),$$

dann existiert ein \tilde{x} im Inneren (c, d) mit $F'(\tilde{x}) = \theta$.

Beweis. *Die Funktion $F(x) - \theta \cdot x$ hat in den Endpunkten des Intervalls $[c, d]$ den gleichen Wert. Wenn sie konstant ist, ist nichts zu beweisen. Sie nimmt in jedem Fall das Minimum oder das Maximum in einem inneren Punkt \tilde{x} an. Dort ist die Ableitung $= \theta$; denn es gilt $F(\tilde{x} + h) - F(\tilde{x}) = \tilde{a} \cdot h + o(|h|)$; und eine Funktion $H(\cdot)$ mit $H(\tilde{x}) = 0$, $H'(\tilde{x}) \neq 0$ nimmt in jeder Umgebung von \tilde{x} sowohl positive als auch negative Werte an.*

Wir werden den 'Mittelwertsatz der Differentialrechnung' nicht weiter verwenden, — weil wir den klassischen Begriff der 'Differenzierbarkeit in jedem Punkt' ablehnen.

Partielle Ableitungen

Bevor wir uns den Funktionenräumen $\mathcal{C}^1(\mathbf{U}, \mathbb{R})$ zuwenden, betrachten wir zunächst einmal einzelne stetig differenzierbare auf einer offenen Teilmenge \mathbf{U} des Spaltenraum $\mathbb{R}_{\mathfrak{S}_p}^2$. Die Koordinatenfunktionen seien mit \mathbf{x}, \mathbf{y} bezeichnet. Die Funktionen f auf \mathbf{U} haben die Gestalt $F(\mathbf{x}, \mathbf{y})$ mit $(\mathbf{x}, \mathbf{y}) \in \mathbf{U}$. Für jedes feste $\tilde{\mathbf{y}} \in \mathbb{R}$ ist $\mathbf{U}_{\tilde{\mathbf{y}}} = \{\mathbf{x} : (\mathbf{x}, \tilde{\mathbf{y}}) \in \mathbf{U}\}$ eine (möglicherweise leere) offene Menge von \mathbb{R} , also eine disjunkte Vereinigung offener Intervalle. Wenn die Einschränkung der Funktion f auf alle diese Intervalle stetig differenzierbar ist, dann sagt man, f sei stetig nach \mathbf{x} differenzierbar. Die (auf ganz \mathbf{U} definierte) stetige Ableitung wird mit $f_1 = \frac{\partial}{\partial \mathbf{x}} f$ bezeichnet. Entsprechend ist $f_2 = \frac{\partial}{\partial \mathbf{y}} f$ definiert, wenn die Einschränkung von f auf alle ‘Schnitte’ $\mathbf{U}_{\tilde{\mathbf{x}}} = \{\mathbf{y} : (\tilde{\mathbf{x}}, \mathbf{y}) \in \mathbf{U}\}$ stetig differenzierbar ist.

Definition 1.3 (Die Funktionenalgebren $\mathcal{C}^r(\mathbf{U}, \mathbb{R})$).

Eine Funktion f auf einer offenen Teilmenge $\mathbf{U} \subseteq \mathbb{R}_{\mathfrak{S}_p}^2$ heisst stetig (partiell) differenzierbar, wenn die partiellen Ableitungen $f_1 = \frac{\partial}{\partial \mathbf{x}} f$ und $f_2 = \frac{\partial}{\partial \mathbf{y}} f$ stetige Funktionen sind. Sie heisst zweimal stetig differenzierbar, wenn f_1 und f_2 stetig (partiell) differenzierbar sind. Sie heisst r -mal stetig differenzierbar, wenn die partiellen Ableitungen $(r - 1)$ -mal stetig differenzierbar sind. Die Algebra der r -mal stetig differenzierbaren Funktionen wird mit $\mathcal{C}^r(\mathbf{U}, \mathbb{R})$ bezeichnet.

Wir bemerken, dass jede der \mathfrak{p} partiellen Ableitungen den Funktionenraum $\mathcal{C}^r(\mathbf{U}, \mathbb{R})$ in den Funktionenraum $\mathcal{C}^{r-1}(\mathbf{U}, \mathbb{R})$ abbildet.

Bevor wir uns weiter mit dem Begriff der stetigen (partiellen) Differenzierbarkeit auf \mathbf{U} beschäftigen, betrachten wir den Begriff der Differenzierbarkeit in einem Punkt.

Definition 1.4 (Totale Differenzierbarkeit im Punkt $\tilde{\mathbf{P}}$).

Eine Funktion f auf der offenen Menge $\mathbf{U} \subseteq \mathbb{R}_{\mathfrak{S}_p}^2$ heisst total differenzierbar im Punkt $\tilde{\mathbf{P}} \in \mathbf{U}$, wenn eine affine Funktion $\tilde{\mathbf{a}}$ existiert, sodass gilt

$$f(\mathbf{P}) - \tilde{\mathbf{a}}(\mathbf{P}) = o(\|\mathbf{P} - \tilde{\mathbf{P}}\|) \quad \text{für } \|\mathbf{P} - \tilde{\mathbf{P}}\| \rightarrow 0$$

d. h., wenn eine Linearform $\tilde{\ell} = d\tilde{\mathbf{a}}$ existiert, sodass

$$f(\tilde{\mathbf{P}} + \mathbf{v}) = f(\tilde{\mathbf{P}}) + \tilde{\ell}(\mathbf{v}) + o(\|\mathbf{v}\|) \quad \text{für } \|\mathbf{v}\| \rightarrow 0$$

Statt $\tilde{\ell}(\mathbf{v})$ notiert man auch $\langle \tilde{\ell}, \mathbf{v} \rangle$ oder $\langle d\tilde{\mathbf{a}}, \mathbf{v} \rangle$. Wenn f im Punkt $\tilde{\mathbf{P}}$ total differenzierbar ist, dann ist natürlich die Linearform durch das ‘lokale’ Verhalten der Funktion $f(\cdot)$ bei $\tilde{\mathbf{P}}$ eindeutig bestimmt. Man notiert $d\tilde{\mathbf{a}} = df|_{\tilde{\mathbf{P}}}$ und nennt diese Linearform den Anstieg der glatten Funktion im Punkt $\tilde{\mathbf{P}}$.

Tupel und Matrizen. Wir betrachten jetzt Tupel stetiger Funktionen auf einer offenen Teilmenge \mathbf{U} des \mathfrak{p} -dimensionalen affinen Raums der reellen J -Spalten.

Zunächst haben wir das J -Tupel der Koordinatenfunktionen $\{\mathbf{x}^j : j \in J\}$. Dieses denken wir uns gelegentlich angeordnet; und wir notieren dann $\{\mathbf{x}^j : j = 1, \dots, \mathfrak{p}\}$. Der hochgestellte Index darf da natürlich nicht als eine Potenz interpretiert werden: das Quadrat der

Funktion x^j müsste mit $(x^j)^2$ bezeichnet werden; $x^j x^k$ bezeichnet das punktweise Produkt der Funktionen x^j und x^k .

Der Begriff der totalen Differenzierbarkeit in einem Punkt \tilde{P} stellt Verbindungen zur linearen Algebra her. Eine Funktion f , die im Punkt $\tilde{P} \in U$ total differenzierbar ist, liefert eine Linearform auf dem p -dimensionalen Vektorraum $T_{\tilde{P}}$ der Tangentialvektoren. Die Tangentialvektoren verstehen wir als Spalten. Den Anstieg von f im Punkt \tilde{P} verstehen wir als eine p -Zeile, und wir bezeichnen sie $\langle df|_{\tilde{P}}, \cdot \rangle$ oder auch einfach $df|_{\tilde{P}}$. Es handelt sich um die J -Zeile mit den Einträgen $f_j(\tilde{P})$, den partiellen Ableitungen nach Koordinatenvariablen.

(Die Bezeichnung $f_j(\tilde{P}) = \frac{\partial}{\partial x^j} f(\tilde{P})$ ist üblich, aber nicht unproblematisch, wie wir später sehen werden.)

Die Familie der Anstiege der Koordinatenfunktionen ist ein linear unabhängiges p -Tupel; sie ist eine Basis $\{dx^j|_{\tilde{P}} : j \in J\}$ des Raums aller Linearformen auf dem Tangentialraum $T_{\tilde{P}}$; man sagt auch, dass sie eine Basis des Cotangentialraums $T_{\tilde{P}}^*$ bilden.— Mit den Begriffen Tangentialraum und Cotangentialraum im Punkt \tilde{P} werden wir uns später ausführlich zu befassen haben.

Die Koeffizienten in der Darstellung des Anstiegs einer Funktion f in dieser Basis sind offenbar die partiellen Ableitungen von f ; es gilt

$$df|_{\tilde{P}} = \sum_{j \in J} f_j(\tilde{P}) \cdot dx^j|_{\tilde{P}}.$$

Betrachten wir jetzt ein I -Tupel von Funktionen, die im Punkt \tilde{P} total differenzierbar sind $\{f^i : i \in I\}$. Die Anstiege versammeln wir als die Einträge einer $I \times J$ Matrix, der Funktionalmatrix oder Jacobi-Matrix der Familie:

$$f^i_j(\tilde{P}) = \frac{\partial f^i}{\partial x^j}(\tilde{P}). \quad \text{Es gilt} \quad df^i|_{\tilde{P}} = \sum_{j \in J} f^i_j(\tilde{P}) \cdot dx^j|_{\tilde{P}}$$

Das I -Tupel $F(\cdot) = \{f^i(\cdot) : i \in I\}$ wird auch eine in \tilde{P} total differenzierbare Abbildung in den Spaltenraum $\mathbb{R}_{S_p}^I$ genannt. Die Matrix $(f^i_j(\tilde{P}))_{ij}$ liefert dazu eine lineare Abbildung des Raums der J -Spalten in den Raum der I -Spalten. Man bezeichnet sie mit $F'(\tilde{P})$ und man nennt sie die Funktionalmatrix oder Jacobi-Matrix des I -Tupels. .

Fazit: Die totale Differenzierbarkeit der Abbildung $\mathbb{R}^J \supseteq U \xrightarrow{F(\cdot)} \mathbb{R}^I$ im Punkt \tilde{P} bedeutet die Existenz einer $I \times J$ -Matrix $F'(\tilde{P})$ sodass gilt

$$F(\tilde{P} + v) = F(\tilde{P}) + F'(\tilde{P}) \cdot v + o(\|v\|) \quad \text{für } \|v\| \rightarrow 0$$

Die Einträge der Matrix $F'(\tilde{P})$ sind die partiellen Ableitungen der f^i nach den x^j .

Der folgende Satz liegt auf der Hand

Satz 1.1.3. *Wenn f und g im Punkt \tilde{P} total differenzierbar sind, dann auch die punktweise Summe und das punktweise Produkt. Es gilt*

$$d(f + g)|_{\tilde{P}} = df|_{\tilde{P}} + dg|_{\tilde{P}}; \quad d(f \cdot g)|_{\tilde{P}} = g(\tilde{P}) \cdot df|_{\tilde{P}} + f(\tilde{P}) \cdot dg|_{\tilde{P}}.$$

Satz 1.1.4 (Kettenregel).

Sei $F(\cdot)$ eine in \tilde{P} totaldifferenzierbare Abbildung und $G(\cdot)$ eine in $\tilde{Q} = F(\tilde{P})$ total differenzierbare Abbildung. Für die (in einer Umgebung von \tilde{P} definierte zusammengesetzte Abbildung $H(\cdot) = G(F(\cdot))$) gilt dann

$$H'(\tilde{P}) = G'(F(\tilde{P})) \cdot F'(\tilde{P}) \quad \text{Matrizenprodukt.}$$

Beweis. Mit $w = F(\tilde{P} + v) - F(\tilde{P}) = F'(\tilde{P}) \cdot v + o(\|v\|)$ haben wir

$$\begin{aligned} H(\tilde{P} + v) - H(\tilde{P}) &= G(\tilde{Q} + w) - G(\tilde{Q}) = \\ &= G'(\tilde{Q}) \cdot w + o(\|w\|) = G'(\tilde{Q}) \cdot (F'(\tilde{P}) \cdot v + o(\|v\|)) + o(\|w\|) \end{aligned}$$

Wegen der totalen Differenzierbarkeit von $F(\cdot)$ ist der Rest $o(\|w\|)$ ein $o(\|v\|)$.

Stetige Differenzierbarkeit

Definition 1.5. Eine Funktion f heisst stetig differenzierbar auf der offenen Menge $U \subseteq \mathbb{R}_{\mathbb{S}^p}^p$, wenn sie in jedem Punkt total differenzierbar ist und zudem die Linearform df stetig vom Fußpunkt abhängt.

Satz 1.1.5. Die Funktion f ist genau dann stetig differenzierbar auf U , wenn die partiellen Ableitungen auf U stetig sind.

Beweis. Wir führen den Beweis im Falle $p = 2$. Es sei e_x, e_y die Basis zum affinen Koordinatensystem (x, y) . Wenn $f = F(x, y)$ im Punkt \tilde{P} total differenzierbar ist, dann gilt für $v = \lambda e_x + \mu e_y$ und $h \rightarrow 0$

$$f(\tilde{P} + h \cdot v) = f(\tilde{P}) + h \cdot \tilde{\ell}(\lambda e_x + \mu e_y) + o(\|h\|)$$

Für die speziellen Verschiebungsvektoren e_x, e_y erhalten wir $\tilde{\ell}(e_x) = f_1(\tilde{P}), \tilde{\ell}(e_y) = f_2(\tilde{P})$. Die stetige Differenzierbarkeit von f impliziert die Stetigkeit der partiellen Ableitungen.

Es seien umgekehrt die partiellen Ableitungen f_1, f_2 stetig, und es sei eine Rechtecks-umgebung von \tilde{P} so klein gewählt, dass in dieser Umgebung gilt $|f_j(P) - f_j(\tilde{P})| < \varepsilon$. Für ein $v = \lambda e_x + \mu e_y$ in dieser Umgebung gilt dann mit der Linearform $\tilde{\ell}(\lambda e_x + \mu e_y) = \lambda f_1(\tilde{P}) + \mu f_2(\tilde{P})$

$$\begin{aligned} f(\tilde{P} + v) - f(\tilde{P}) - \tilde{\ell}(v) &= \\ &= \left(f(\tilde{P} + v) - f(\tilde{P} + \lambda e_x) - \mu f_2(\tilde{P}) \right) + \left(f(\tilde{P} + \lambda e_x) - f(\tilde{P}) - \lambda f_1(\tilde{P}) \right) \end{aligned}$$

Der zweite Summand ist $o(\|v\|)$ wegen der partiellen Differenzierbarkeit nach x im Punkt \tilde{P} . Der erste lässt sich durch $\varepsilon \cdot |\mu|$ abschätzen wegen der partiellen Differenzierbarkeit nach y im Punkt $\tilde{P} + \lambda e_x$ und $|f_2(\tilde{P} + \lambda e_x) - f_2(\tilde{P})| < \varepsilon$

Eine offene Menge U im \mathbb{R}^p heisst sternförmig in bezug auf den Punkt \tilde{P} , wenn für jedes $P \in U$ die Verbindungsstrecke zu \tilde{P} zu U gehört.

Satz 1.1.6. Wenn für f in einem sternförmigen Gebiet die partiellen Ableitungen stetig sind, dann existiert ein p -Tupel stetiger Funktionen $M_1(\cdot), \dots, M_p(\cdot)$ sodass

$$f(P) - f(\tilde{P}) = (x^1(P) - x^1(\tilde{P})) \cdot M_1(P) + \dots + (x^p(P) - x^p(\tilde{P})) \cdot M_p(P).$$

Beweis.

Wir betrachten für t in einer Umgebung von $[0, 1]$ die Funktion $F(t) = f(\tilde{P} + t(P - \tilde{P}))$. Sie ist stetig differenzierbar mit der stetigen Ableitung

$$F'(t) = \sum f_j(\tilde{P} + t(P - \tilde{P})) \cdot (x^j(P) - x^j(\tilde{P})).$$

Die Funktionen $M_j(P) = \int_0^1 f_j(\tilde{P} + t(P - \tilde{P})) dt$ leisten das Verlangte.

Die für den obigen Beweis eingeführten Notationen sind auch nützlich für den Beweis des folgenden wichtigen Satzes.

Satz 1.1.7 (Satz von H. A. Schwarz).

Die Funktion f besitze in einer Umgebung von \tilde{P} stetige partielle Ableitungen f_1, f_2 , die selbst entlang der Koordinatenachsen totalstetig sind, die also partielle Ableitungen $f_{11}, f_{12}, f_{21}, f_{22}$ besitzen. Wenn diese Funktionen im Punkt \tilde{P} stetig sind dann gilt $f_{12}(\tilde{P}) = f_{21}(\tilde{P})$.

Beweis.

Für den Beweis benötigen wir die Absolutstetigkeit der Funktionen $f_1(P + \mu e_y)$ als Funktionen von μ in einer Umgebung der 0 und die Absolutstetigkeit der Funktionen $f_2(P + \lambda e_x)$ als Funktionen von λ in einer Umgebung der 0. Zusammen mit der Absolutstetigkeit von $f(\tilde{P} + \lambda e_x + \mu e_y)$ als Funktion von μ ergibt sich

$$\begin{aligned} f(\tilde{P} + \lambda e_x + \mu e_y) - f(\tilde{P} + \lambda e_x) &= \int_0^\mu f_2(\tilde{P} + \lambda e_x + t e_y) dt \\ f_2(\tilde{P} + \lambda e_x + t e_y) - f_2(\tilde{P} + t e_y) &= \int_0^\lambda f_{21}(\tilde{P} + s e_x + t e_y) ds \end{aligned}$$

$$f(\tilde{P} + \lambda e_x + \mu e_y) - f(\tilde{P} + \lambda e_x) = \int_0^\mu f_2(\tilde{P} + t e_y) dt + \int_0^\mu \int_0^\lambda [f_{21}(\tilde{P} + s e_x + t e_y)] ds dt$$

Der erste Term rechts ist $\int_0^\mu f_2(\tilde{P} + t e_y) dt = f(\tilde{P} + \mu e_y) - f(\tilde{P})$. Somit haben wir die ‘gemischte Differenz’ als Doppelintegral dargestellt

$$f(\tilde{P} + \lambda e_x + \mu e_y) - f(\tilde{P} + \lambda e_x) - f(\tilde{P} + \mu e_y) + f(\tilde{P}) = \int_0^\mu \int_0^\lambda [f_{21}(\tilde{P} + s e_x + t e_y)] ds dt.$$

Wenn wir die Rolle der beiden Achsen vertauschen, dann liefert die Stetigkeit der Funktionen f_{12} und f_{12} die Behauptung.

Der Satz sollte so memoriert werden:

Satz 1.1.8 (Satz von der Hesse-Matrix).

Wenn f zweimal stetig differenzierbar ist, dann ist die Matrix der zweiten Ableitungen eine symmetrische Matrix, die in stetiger Weise vom Fußpunkt abhängt. Zu jedem Fußpunkt P existiert eine symmetrische Bilinearform $\mathbf{b}_P(\cdot, \cdot)$, sodass

$$f(P + \mathbf{v}) = f(P) + \langle \mathbf{d}f|_P, \mathbf{v} \rangle + \frac{1}{2} \mathbf{b}_P(\mathbf{v}, \mathbf{v}) + o(\|\mathbf{v}\|^2) \quad \text{für } \mathbf{v} \rightarrow 0$$

Die Bilinearform, als 2×2 -Matrix notiert, ist die Hesse-Matrix. Der Fußpunkt sei $(0, 0)$. Die zweiten partiellen Ableitungen liefern für (x, y) in einem sternförmigen Gebiet

$$\begin{aligned} f(x, y) - f(0, 0) - x \cdot f_1(0, 0) - y \cdot f_2(0, 0) &= \\ &= \int_0^1 [x^2 f_{11}(sx, sy) + 2xy f_{12}(sx, sy) + y^2 f_{22}(sx, sy)] \cdot (1 - s) \, ds. \end{aligned}$$

Beweis. Es bleibt nur noch die letzte Formel zu zeigen: Wegen der Totalstetigkeit der Funktion $f_j(tx, ty)$ im Intervall $[0, 1]$ haben wir für alle $t \in [0, 1]$

$$\begin{aligned} f_j(tx, ty) &= \int_0^t [x f_{j1}(sx, sy) + y f_{j2}(sx, sy)] \, ds \\ f(x, y) &= \int_0^1 [x f_1(tx, ty) + y f_2(tx, ty)] \, dt = \\ &= \int_0^1 \int_0^t [x^2 f_{11}(sx, sy) + 2xy f_{12}(sx, sy) + y^2 f_{22}(sx, sy)] \, ds \, dt = \\ &= \int_0^1 [x^2 f_{11}(sx, sy) + 2xy f_{12}(sx, sy) + y^2 f_{22}(sx, sy)] \cdot (1 - s) \, ds. \end{aligned}$$

Bemerkung zur Totalstetigkeit im \mathbb{R}^2 : Es sei $d\mu(x, y)$ ein endliches signiertes Borel-Maß auf dem \mathbb{R}^2 oder auf einem Rechteck im \mathbb{R}^2 . Man definiert dann die Verteilungsfunktion $F(x, y) = \mu((-\infty, x] \times (-\infty, y])$. Man gewinnt daraus leicht das Maß eines beliebigen Rechtecks

$$\mu((x_1, x_2] \times (y_1, y_2]) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1).$$

Das signierte Maß μ ist genau dann totalstetig bzgl. des zweidimensionalen Lebesgue-Maßes, wenn eine integrierbare Funktion $p(\cdot, \cdot)$ existiert, sodass

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y p(s, t) \, ds \, dt.$$

Unser Beweis des Satzes von Schwarz zeigt, dass das Maß zu einem $F(\cdot, \cdot)$, dessen partielle Ableitungen überall die genannten Stetigkeitseigenschaften besitzen, totalstetig ist mit der stetigen Dichte $p(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y)$.

Didaktische These: Die Beweise von Aussagen, welche in traditioneller Weise den sog. Mittelwertsatz der Differentialrechnung heranziehen, führen durchweg in Sackgasen. Die Pendanten dieser Sätze, die sich auf den Begriff der Absolutstetigkeit (d. h. der Totalstetigkeit bzgl. des Lebesgue-Maßes) stützen, sind überlegen. Als Beispiel kann der Satz von H. A. Schwarz dienen oder die folgende Variante des Satzes von der Differentiation unter dem Integral-Zeichen. ('Parameterabhängiges Integral') Dabei geht es um die Gültigkeit der Formel

$$\frac{d}{dx} \int F(x, \omega) \, d\mu(\omega) = \int \frac{d}{dx} F(x, \omega) \, d\mu(\omega).$$

Ein konkreter Fall wäre etwa die Frage nach der Ableitung des zweiten Euler'schen Integrals (als einer Darstellung der Gamma-Funktion auf \mathbb{R}_+):

$$\frac{d}{d\alpha} \int_0^\infty x^{\alpha-1} e^{-x} \, dx = \int_0^\infty \ln x \cdot x^{\alpha-1} e^{-x} \, dx \quad \text{für alle } \alpha > 0.$$

Satz 1.1.9. *Es sei $F(x, \omega)$ produktmessbar über $(\mathbb{R}, \mathfrak{B}) \times (\Omega, \mathfrak{A})$ mit*

1. $F(\cdot, \omega)$ ist absolutstetig für alle ω ; d. h. es existiert $f(\cdot, \cdot)$, sodass

$$F(x, \omega) - F(x_0, \omega) = \int_{x_0}^x f(y, \omega) \, dy \quad \text{für alle } \omega$$

($f(\cdot, \cdot)$ kann offenbar produktmessbar gewählt werden.)

2. Für ein Maß $d\mu(\omega)$ sei

$$G(x) = \int F(x, \omega) \, d\mu(\omega) \quad \text{und} \quad \iint |f(y, \omega)| \, d\mu(\omega) \, dy < \infty.$$

Es gilt dann

$$G(x) - G(x_0) = \int [F(x, \omega) - F(x_0, \omega)] \, d\mu(\omega) = \int_{x_0}^x \left[\int f(y, \omega) \, d\mu(\omega) \right] \, dy.$$

Anders gesagt: Das 'parameterabhängige Integral' $G(x)$ ist als Funktion des Parameters absolutstetig, wobei sich die Ableitung als Integral über den differenzierten Integranden ergibt: $G'(x) = \int f(x, \omega) \, d\mu(\omega)$.

Der Satz erweist sich als eine Anwendung des Satzes von Fubini.

Bevor wir uns tiefergehend mit den stetig differenzierbaren Fragen befassen, werfen wir einen Blick auf die (nicht notwendigerweise glatten) konvexen Funktionen. Hier treten in manchen Punkten sog. Tangentialkegel (an den Graphen der Funktion) an die Stelle von tangentialen Hyperebenen.

Richtungsableitungen und konvexe Funktionen

Es sei f eine stetige Funktion auf einer offenen Teilmenge des \mathbb{R}^p . Man sagt, dass sie im Punkt \tilde{P} eine Richtungsableitung in der Richtung des Verschiebungsvektors v besitzt, wenn der folgende Limes existiert:

$$\lim_{h \searrow 0} \frac{1}{h} (f(\tilde{P} + h \cdot v) - f(\tilde{P})) = \nabla'(\tilde{P}, v)f.$$

Bemerkung: Wenn $\nabla'(\tilde{P}, v)f$ existiert, dann auch $\nabla'(\tilde{P}, \lambda v)f$ für alle $\lambda > 0$, und es gilt

$$\nabla'(\tilde{P}, \lambda v)f = \lambda \cdot \nabla'(\tilde{P}, v)f.$$

Wenn in einem Punkt \tilde{P} alle Richtungsableitungen existieren, dann bilden sie eine positivhomogene Funktion $\nabla'(\tilde{P}, \cdot)f$. Genau dann, wenn es sich hierbei um eine lineare Funktion handelt, ist die Funktion im Punkt \tilde{P} total differenzierbar. $\nabla'(\tilde{P}, \cdot)f = df|_{\tilde{P}}$.

Zur Erinnerung: Bei den Funktionen auf einer Teilmenge der reellen Achse ($p = 1$) kennt man den Begriff der rechtsseitigen- und der linksseitigen Ableitung. Der Begriff der Richtungsableitung ist damit eng verwandt; er bezieht sich allerdings nicht auf die Besonderheiten der reellen Achse (mit ihrer Anordnung und ihrer Normierung durch den Punkt 1); und das bringt es mit sich, dass die Richtungsableitung eine positivhomogene Funktion ist, und nicht ein Zahlenpaar, wie die 'einseitigen' Ableitungen. Die rechtsseitige Ableitung ist die Geschwindigkeit, mit welcher sich die Funktion ändert, wenn das Argument mit Geschwindigkeit 1 nach rechts geht.

Satz 1.1.10. *Sei U eine offene konvexe Menge und $k(\cdot)$ eine konvexe Funktion, die auf U endlichwertig ist. In jedem Punkt $\tilde{P} \in U$ existieren dann Richtungsableitungen in allen Richtungen v und die Funktion $\nabla'(\tilde{P}, \cdot)k$ ist in jedem Punkt \tilde{P} ein Minkowski-Funktional, d. h. eine positivhomogene sublineare Funktion.*

Der Beweis ist eine einfache Übungsaufgabe.

Eine konvexe zweimal stetig differenzierbare Funktion ist offenbar dadurch ausgezeichnet, dass sie in allen Punkten ihres Endlichkeitsbereichs eine positivsemidefinite Hesse-Matrix besitzt. Man kann leicht zeigen: Wenn die Hesse-Matrix überall im Endlichkeitsbereich positivdefinit ist, dann hat auch die Legendre-Transformierte diese Eigenschaft, und die Hesse-Matrizen in korrespondierenden Punkten sind zueinander invers. Die konvexen Funktionen dieser speziellen Art haben wir oben die runden konvexen Funktionen genannt. Auf die glatten Funktionen dieser Art wollen wir hier nicht weiter eingehen; denn die konvexe Geometrie tendiert in eine ganz andere Richtung als die Geometrie der glatten Mannigfaltigkeiten, welcher wir uns im Folgenden zuwenden wollen.

1.2 Krummlinige Koordinaten; Tangentialvektoren

Stetig differenzierbare reellwertige Funktionen auf einer offenen Teilmenge U des \mathbb{R}^n bezeichnen wir mit Buchstaben wie f, g, h, \dots oder auch u, v, \dots . Die Gesamtheit aller dieser Funktionen wird mit $\mathcal{C}^1(U)$ bezeichnet; wenn wir deutlich machen wollen, dass es um reellwertige Funktionen geht, notieren wir $\mathcal{C}^1(U, \mathbb{R})$. Es handelt sich um eine Funktionenalgebra. Jede der partiellen Ableitungen D_j bildet $\mathcal{C}^1(U)$ in den Raum $\mathcal{C}^0(U)$ der stetigen Funktionen ab. Es gilt

$$D_j(f + g) = D_j f + D_j g; \quad D_j(f \cdot g) = g \cdot D_j f + f \cdot D_j g.$$

Eine Funktion heisst r -mal stetig differenzierbar, wenn die partiellen Ableitungen $D_j f$ ($r-1$)-mal stetig differenzierbar sind. Die Algebra der r -mal stetig differenzierbaren Funktionen wird mit $\mathcal{C}^r(U)$ bezeichnet. Für einen Multi-Index $\alpha = (\alpha_1, \dots, \alpha_n)$ mit Gewicht $\alpha_1 + \dots + \alpha_n \leq r$ bezeichnet man mit $D_\alpha f$ die Funktion, die man erhält, wenn man f α_1 -mal nach der ersten Koordinate partiell ableitet, α_2 -mal nach der zweiten, \dots . Der Satz von H. A. Schwarz garantiert, dass es gleichgültig ist, in welcher Reihenfolge man die partiellen Differentiationen durchführt.

Die einfachsten Funktionen dieser Art sind die (auf U eingeschränkten) affinen Funktionen und speziell die ‘natürlichen’ Koordinatenvariablen x^j . Besonders einfache Beispiele sind auch die quadratischen Formen $\sum_{jk} a_{jk} \cdot x^j \cdot x^k$, sowie allgemeiner die polynomialen Funktionen vom Grad m in den n Koordinatenvariablen. Diese haben die Form

$$f = a + \sum_j a_j \cdot x^j + \sum_{jk} a_{jk} \cdot x^j \cdot x^k + \dots + \sum_{j_1 j_2 \dots j_m} a_{j_1 j_2 \dots j_m} \cdot x^{j_1} \cdot x^{j_2} \cdot \dots \cdot x^{j_m}.$$

(wo man in den üblichen Darstellungen annimmt, dass die Koeffizientensysteme $a_{j_1 j_2 \dots j_l}$ für alle $l = 2, 3, \dots, m$ symmetrisch sind.

Natürlich kann man nicht alle $f \in \mathcal{C}^1(U)$ mit einfachen Operationen aus ‘elementaren’ Funktionen gewinnen; die f werden bei uns nur selten einmal durch Formeln spezifiziert sein. Wenn man zum Ausdruck bringen will, dass man sich $f(P)$ (für jedes $P \in U$) aus dem Tupel der Koordinaten $x^1(P), \dots, x^n(P)$ gewonnen denkt, schreibt man $f(\cdot) = F(x^1(\cdot), \dots, x^n(\cdot))$ oder auch $f = F(x^1, \dots, x^n)$. Jedes $f \in \mathcal{C}^1(U)$ besitzt eine solche Darstellung. Man notiert lieber einfach f , wenn man Distanz halten will zu irgendwelchen (‘natürlichen’ oder nicht so natürlichen) Koordinatensystemen.

Seitenblick auf Konventionen in Linearer Algebra bzw. affiner Geometrie:

In der Lehre von den endlichdimensionalen Vektorräumen V sagt man, eine Basis sei ein linear unabhängiges System maximaler Länge $\{v_j : j \in J\}$; jeder Vektor v lässt sich in eindeutiger Weise als Linearkombination der Basisvektoren gewinnen $v = \sum_{j \in J} v_j \cdot x^j$. Die x^j heissen die Koordinaten des Vektors (bzgl. der gegebenen Basis); sie sind lineare Funktionen auf V , wenn man sie in Abhängigkeit von v versteht. Das Tupel $\{x^j(\cdot) : j \in J\}$ ist die zu $\{v_j : j \in J\}$ duale Basis.

Die Techniken des Basiswechsels und (dual dazu) des Koordinatenwechsels sollten bekannt sein: Zu einem Basiswechsel gehört eine invertierbare $I \times J$ -Matrix A

$$\sum_{j \in J} \mathbf{v}_j \cdot \mathbf{x}^j = \mathbf{v} = \sum_{i \in I} \mathbf{u}_i \cdot \mathbf{y}^i,$$

$$\mathbf{y}^i = \sum_{j \in J} \mathbf{a}^i_j \cdot \mathbf{x}^j \quad \text{für } i \in I, \quad \sum_{i \in I} \mathbf{u}_i \cdot \mathbf{a}^i_j = \mathbf{v}_j \quad \text{für } j \in J.$$

In der Theorie der krummlinigen Koordinaten finden wir analoge Verhältnisse. Die Matrix zu einem Koordinatenwechsel heisst die Jacobi-Matrix zum Koordinatenwechsel. Der Koordinatenwechsel liefert in jedem Punkt eine linear veränderte Darstellung der Tangentialvektoren im Sinne eines Basiswechsel im Tangentialraums $T_{\mathbf{p}}$. (Eine formale Definition des Begriffs Tangentialraums gegen wir unten.) Die Einträge der Jacobi-Matrix ergeben sich als partielle Ableitungen; sie sind also \mathcal{C}^{r-1} -glatt, wenn der Koordinatenwechsel \mathcal{C}^r -glatt ist.

Warnung: Häufig nimmt man (ohne ein Wort darüber zu verlieren), dass die Indexmenge J zu einer Basis in einer speziellen Weise angeordnet oder sogar durchgezählt ist, sodass man also vom ersten, zweiten, ... Basisvektor sprechen kann. Die Koordinaten \mathbf{x}^j schreibt man von oben nach unten in eine Spalte; die Basisvektoren werden von links nach rechts aufgezählt. Die Matrix eines Basis-(bzw. Koordinatenwechsels) wirkt auf die Spalte der Koordinaten von links und auf die Zeile der Basisvektoren von rechts. In den folgenden Überlegungen wird die Aufzählung der Basisvektoren (und Koordinatenfunktionen) vorerst keine Rolle spielen; wir können sorglos damit umgehen. Erst später werden wir gleichgerichtete und gegengerichtete Basen zu unterscheiden haben. Der Wechsel zu einer gleichgerichteten Basis ist durch eine Matrix mit positiver Determinante gekennzeichnet.

Definition 1.6 (\mathcal{C}^1 -Koordinatisierung).

Es sei U eine offene Teilmenge des $\mathbb{R}_{\mathbb{S}^p}^n$. Ein n -Tupel stetig differenzierbarer Funktionen $\{\mathbf{y}^i : i \in I\}$ auf U heisst eine \mathcal{C}^1 -Koordinatisierung von U , wenn gilt

1. Die Punkte $P \in U$ sind durch die Zahlentupel $\mathbf{y}^i(P)$ eindeutig bestimmt
2. Die Differentiale $d\mathbf{y}^i$ sind in jedem Punkt $P \in U$ linear unabhängig.

Beispiel 1.2.1 (Polarkoordinaten).

Es sei U die entlang der negativen reellen Achse geschlitzte komplexe Ebene, (verstanden als offene Teilmenge des Raums $\mathbb{R}_{\mathbb{S}^p}^2$). Jeder Punkt $z \in U$ hat eine eindeutige Darstellung

$$z = r \cdot e^{i\phi} \quad \text{mit } r > 0, \quad \phi \in (-\pi, +\pi)$$

Das Funktionenpaar $\{r, \phi\}$ ist eine \mathcal{C}^1 -Koordinatisierung von U . Der Anstieg dieser beiden Funktionen ergibt sich folgendermaßen aus dem Anstieg der cartesischen Koordinaten

$$\begin{aligned} x &= r \cdot \cos \phi & dx &= \cos \phi \cdot dr - r \sin \phi \cdot d\phi, \\ y &= r \cdot \sin \phi & dy &= \sin \phi \cdot dr + r \cos \phi \cdot d\phi, \end{aligned}$$

Die Jacobi-Matrix hat übrigens die Determinante $r(\cos^2 \phi + \sin^2 \phi) = r \neq 0$.

Die Jacobi-Matrix zu einem Koordinatenwechsel hat (mindestens) zwei Interpretationen. Zunächst einmal bemerken wir: Jedes $f \in \mathcal{C}^1(\mathcal{U}, \mathbb{R})$ liefert in jedem $P \in \mathcal{U}$ einen Anstieg $df|_P$. Die Menge der möglichen Anstiege in P ist ein n -dimensionaler Vektorraum, man nennt ihn den Cotangentialraum in P und notiert ihn T_P^* . Eine Basis ist z. B. $\{dx^j|_P : j \in J\}$, wo $\{x^j : j \in J\}$ das ‘natürliche’ Koordinatensystem bezeichnet. Ist $\{y^i : i \in I\}$ eine weitere \mathcal{C}^1 -Koordinatisierung, so ist $\{dy^i|_P : i \in I\}$ eine weitere Basis. Die Basistransformation ist durch eine Jacobi-Matrix gegeben

$$y^i = f^{(i)}(x^1, \dots, x^n) \quad dy^i = \sum_{j \in J} f^i_j \cdot dx^j.$$

Man bemerke, dass die $I \times J$ -Matrix (f^i_j) zu einem Koordinatenwechsel in jedem P nicht-singulär ist und dass die Einträge $f^i_j(P)$ als die partiellen Ableitungen der Funktionen $f^{(i)}$ stetige Funktionen sind. Man notiert gelegentlich

$$(J^i_j)(\cdot) = \frac{\partial(y^i)}{\partial(x^j)}(\cdot) \quad \text{oder kurz} \quad J(\cdot) = \frac{\partial(y^I)}{\partial(x^J)}(\cdot)$$

Die Jacobimatrix transformiert das Tupel der Anstiege (als Spalte aufgelistet) in das Tupel der Anstiege in dem neuen Koordinatensystem.

$$\begin{pmatrix} dy^1 \\ \vdots \\ dy^n \end{pmatrix} = J \cdot \begin{pmatrix} dx^1 \\ \vdots \\ dx^n \end{pmatrix}; \quad dy^i = \sum J^i_j \cdot dx^j.$$

Eine zweite Interpretation der Jacobi-Matrix zu einem Koordinatenwechsel fußt auf dem Begriff des Tangentialvektors, den wir jetzt entwickeln wollen.

Tangentialvektoren

Der Dualraum zu T_P^* heisst der Tangentialraum im Punkt P ; er wird mit T_P bezeichnet. Seine Elemente, die Tangentialvektoren im Punkt P , wollen wir nun auch etwas konkreter beschreiben, zuerst geometrisch-intuitiv und dann formal-algebraisch.

Ein glatt parametrisiertes Kurvenstück, welches zur Zeit 0 durch den Punkt $P_0 = P$ läuft, ist eine Familie von Punkten $\{P_t : t \sim 0\}$, für welche die Funktion $\{f(P_t) : t \sim 0\}$ stetig differenzierbar ist für alle $f \in \mathcal{C}^1$. (Es genügt offenbar zu fordern, dass $\{y^i(P_t) : t \sim 0\}$ stetig differenzierbar ist für irgend ein n -Tupel von Koordinatenfunktionen y^i . (Der Familienparameter t variiert in einer kleinen Umgebung, die wir hier nicht spezifizieren müssen.)

Es sei ein glatt parametrisiertes Kurvenstück vorgegeben. Die Ableitung der Funktion $f(P_t)$ bezeichnen wir mit $\dot{f}(t)$ oder kurz \dot{f} . Für $g = G(x^1, \dots, x^n)$ gilt $\dot{g} = \sum G_j(P_t) \cdot \dot{x}^j$. Ist $\{y^i = f^{(i)}(x^1, \dots, x^n) : i \in I\}$ eine Koordinatisierung, so haben wir $\dot{y}^i = \sum_j f^i_j(P_t) \cdot \dot{x}^j$. Die J -Spalte mit den Einträgen $\dot{x}^j(0)$ beschreibt in der x -Koordinatisierung die Geschwindigkeit, mit welcher die Kurve zur Zeit 0 durch den Punkt $P = P_0$ geht. Andererseits beschreibt die I -Spalte mit den Einträgen $\dot{y}^i(0)$ diesen ‘Tangentialvektor der Kurve’ in

der \mathbf{y} -Koordinatisierung. Man definiert nun formal den Tangentialvektor der Kurve (ohne irgendein Koordinatensystem zu bemühen) als die lineare Abbildung

$$X_P : C^1 \ni f \longmapsto \lim_{t \rightarrow 0} \frac{1}{t} (f(P_t) - P(0)).$$

Jedes glatt parametrisierte Kurvenstück durch P liefert einen Tangentialvektor X_P . Wenn zwei Kurven mit derselben Geschwindigkeit durch den Punkt P laufen, dann liefern sie denselben Tangentialvektor X_P . Man kann also sagen, ein Tangentialvektor mit dem Fußpunkt P ist eine Äquivalenzklasse glatt parametrisierter Kurven durch den Punkt P .— Wir wollen manchmal der (etwas rätselhaften) Tradition folgen, nach welcher Tangentialvektoren mit dem Fußpunkt P mit Buchstaben wie X_P, Y_P, \dots bezeichnet werden. (Eine etwas näherliegende Bezeichnung wäre vielleicht v_P, w_P, \dots)

Die Geschwindigkeit \dot{f} , mit welcher sich f beim Durchgang durch P ändert, hängt offenbar nur vom Anstieg der Funktion im Punkt P ab; der Wert $X_P(f)$ wird daher auch $\langle df, X_P \rangle$ notiert. Somit ist klar, dass die X_P die Elemente des Dualraums von T_P^* sind. Weiter ist klar, dass X_P eine Eigenschaft besitzt, welche die Algebraiker die Eigenschaft einer Derivation nennen. Es gilt

$$X_P(f \cdot g) = f(P) \cdot X_P(g) + g(P) \cdot X_P(f).$$

Didaktische Anmerkung Wir haben hier ein sehr wichtiges Muster vor uns, nach welchem mathematische Objekte und Strukturen konstruiert werden.

1. Man besorgt sich Gegenstände durch Operationen an den bekannten Objekten
2. Man legt fest, wann die so konstruierten Gegenstände als gleich gelten,
3. Man legt fest, nach welchen Regeln mit den Äquivalenzklassen zu rechnen ist.

Der folgende Satz zeigt, dass die algebraische Eigenschaft der Derivation charakteristisch ist für die Tangentialvektoren (wenn man hier einmal die Fragen der Glattheit überspielt).

Satz 1.2.1 (Derivationen im Sinne der Algebra). *Ist D eine lineare Funktion auf der Algebra $C^\infty(\mathbf{U}, \mathbb{R})$ der unendlich oft differenzierbaren Funktion auf $\mathbf{U} \ni P_0$ mit der Eigenschaft*

$$D(f \cdot g) = f(P_0) \cdot D(g) + g(P_0) \cdot D(f),$$

so existiert dann genau ein Tangentialvektor im Punkt P_0 , sodass $D(f) = \langle df, X_{P_0} \rangle$ für alle $f \in C^\infty(\mathbf{U}, \mathbb{R})$.

Beweis. *Wir wählen eine Koordinatisierung $\{x^j : j \in J\}$ mit $x^j(P_0) = 0$ für alle j . Zu jedem f gibt es ein \mathbf{n} -Tupel von Funktionen M_j , sodass $f(\cdot) - f(P_0) = \sum x^j(\cdot) \cdot M_j(\cdot)$; dabei sind die Werte $M_j(P_0)$ die partiellen Ableitungen der Funktion f im Punkt P_0 . Eine Derivation annulliert die Konstante 1 und damit alle Konstanten.*

Mit $\alpha^j = D(x^j)$ gilt $D(f) = \sum [D(x^j) \cdot M_j(P_0) + x^j(P_0) \cdot D(M_j)] = \sum \alpha^j \cdot f_j(P_0)$. $D(\cdot)$ ist also eine Linearkombination der partiellen Ableitungen von f im Punkt P_0 .

Tangentialvektorfelder.

Wenn jedem $P \in U$ ein Tangentialvektor mit dem Fußpunkt P zugeordnet ist, dann spricht man von einem Tangentialvektorfeld $X = \{X_P : P \in U\}$. Wenn das Vektorfeld X so geartet ist, dass für jedes glatte f ($f \in C^\infty(U, \mathbb{R})$) die ‘Ableitung entlang X ’ eine glatte Funktion ist $X(f) : P \mapsto X_P(f)$, dann spricht man von einem glatten Vektorfeld.

Die Gesamtheit \mathfrak{X} aller glatten Vektorfelder ist ein hochinteressantes mathematisches Objekt, mit welchem wir uns später näher beschäftigen werden.

Im Abschnitt über Differenzialgleichungen werden wir uns auch mit konkreten Tangentialvektorfeldern befassen. Hier betrachten nur ein einfaches Beispiel:

Beispiel (Das Geschwindigkeitsfeld zur gleichförmigen Drehung der euklidischen Ebene). Betrachten wir die gleichförmige Drehung mit der Winkelgeschwindigkeit ω . Der Punkt P_0 mit den cartesischen Koordinaten $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ gelangt nach der Zeit t in den Punkt P_t mit den cartesischen Koordinaten $\begin{pmatrix} x_t \\ y_t \end{pmatrix}$.

$$\phi^t : \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} r \cdot \cos(\phi_0) \\ r \cdot \sin(\phi_0) \end{pmatrix} \mapsto \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} r \cdot \cos(\phi_0 + \omega t) \\ r \cdot \sin(\phi_0 + \omega t) \end{pmatrix}$$

Der Tangentialvektor im Punkt P_0 ist in cartesianischen Koordinaten die Spalte

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}(P_0) = \omega \cdot \begin{pmatrix} -r \cdot \sin(\phi_0) \\ r \cdot \cos(\phi_0) \end{pmatrix} = \omega \cdot \begin{pmatrix} -y \\ x \end{pmatrix}(P_0)$$

In Polarkoordinaten ausgedrückt ist der Geschwindigkeitsvektor konstant gleich $\begin{pmatrix} 0 \\ \omega \end{pmatrix}$. Die Koordinatentransformation des Tangentialvektorfelds ergibt sich andererseits auch mittels der Jacobi-Matrix

$$J = \frac{\partial(x, y)}{\partial(r, \phi)} = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \quad J \cdot \begin{pmatrix} 0 \\ \omega \end{pmatrix} = \omega \cdot \begin{pmatrix} -r \cdot \sin \phi \\ r \cdot \cos \phi \end{pmatrix} = \omega \cdot \begin{pmatrix} -y \\ x \end{pmatrix}$$

Cotangentialvektorfelder.

Wenn jedem $P \in U$ ein Cotangentialvektor mit dem Fußpunkt P zugeordnet ist, dann spricht man von einem Cotangentialvektorfeld. Die einfachsten Cotangentialvektorfelder sind die Differentiale df : Jedem ‘Fußpunkt’ P wird eine Linearform $df|_P$ auf dem Raum T_P der Tangentialvektoren zugeordnet. Wenn $\{x^j : j \in J\}$ und $\{y^i : i \in I\}$ C^1 -Koordinatisierungen sind, dann sind die Anstiege $\{dx^j : j \in J\}$ ebenso wie die Anstiege $\{dy^i : i \in I\}$ in jedem Fußpunkt P Basen des Cotangentialraums T_P^* . Die Cotangentialfelder ω (die man auch Pfaff’sche Formen nennt) haben Darstellungen

$$\sum_j a_j dx^j = \omega = \sum_i b_i dy^i$$

mit ortsabhängigen Koeffizienten $a_j(\cdot)$ und $b_i(\cdot)$, von welchen wir zumindest Stetigkeit voraussetzen. Diese Tupel werden als Zeilen notiert. Die I-Zeile \mathbf{b} wird durch die Jacobi-matrix (f^i_j) vom Format $I \times J$ in die J-Zeile \mathbf{a} transformiert. In der Tat gilt

$$\sum_i b_i dy^i = \sum_i b_i \left(\sum_j f^i_j dx^j \right) = \sum_j \left(\sum_i b_i f^i_j \right) dx^j = \sum_j a_j dx^j.$$

Vorausschau: Die Cotangentialvektorfelder sind mathematisch und in den Anwendungen etwas ganz anderes als die Tangentialfelder.

Die Tangentialvektorfelder sind häufig als Geschwindigkeitsfelder zu interpretieren; sie stehen meistens in Verbindung mit einparametrischen Scharen von Transformationen. Eine wichtige mathematische Operation im Raum der Vektorfelder ist neben der Addition (‘Überlagerung’) das sog. Lie-Produkt, auf welches wir später zu sprechen kommen: $(\mathbf{X}, \mathbf{Y}) \mapsto [\mathbf{X}, \mathbf{Y}] \in \mathfrak{X}$.

Cotangentialvektorfelder auf der anderen Seite treten in der Physik häufig als Kraftfelder in Erscheinung; mit Geschwindigkeitsfeldern haben sie nichts zu tun. Die Cotangentialvektorfelder heissen auch 1-Formen oder Differentialformen vom Grad 1 oder auch Pfaff’sche Formen. 1-Formen sind Objekte, die entlang von Kurven integriert werden.

Rückschau: Kurven, entlang welcher 1-Formen wie etwa $f \cdot d\mathbf{g}$ integriert werden, sind uns bereits früher begegnet. In der Theorie der Mannigfaltigkeiten heissen solche Kurven auch ‘singuläre’ 1-Simplices. Man gewinnt sie als Äquivalenzklassen von stetig differenzierbaren Abbildungen des kompakten Einheitsintervalls in die Mannigfaltigkeit. Als äquivalent gelten glatt parametrisierte Kurven, wenn die Parametrisierungen durch eine glatte Bijektion auseinander hervorgehen.— Man fordert nicht die Injektivität der Abbildungen in die Mannigfaltigkeit; so kann beispielsweise die ‘Spur’ einer Kurve mehrfach durchlaufen werden, was einer durchaus nichtäquivalenten Kurve entspricht.

Es gibt auch k -Formen von den Graden $k = 2, 3, \dots, n$. Sie sind geeignet, über Bereiche integriert zu werden, die man k -dimensional nennen könnte; es sind aber i. Allg keine Mannigfaltigkeiten. Das muss später genauer gesagt werden. Die elementarsten Bereiche, über welche man eine k -Form integriert, sind die sog. singulären k -Simplices (oder die singulären k -Würfel). Man gewinnt sie als Äquivalenzklassen von glatten Abbildungen des k -dimensionalen Standard-Simplex (bzw. des k -dimensionalen Standard-Einheitswürfels) in die Mannigfaltigkeit.

Die Gesamtheit aller k -Formen ω^k auf einer n -dimensionalen Mannigfaltigkeit N ($k = 1, \dots, n$) ist ein algebraisch hochinteressantes Objekt $\bigoplus_{k=0}^n \Omega^k N$. k -Formen werden mittels des sog. Dachprodukts miteinander multipliziert; und k -Formen werden mit der Cartan-Ableitung (auch äussere Ableitung genannt) differenziert.

Es sollte nochmals gesagt sein, dass die Theorie der Differentialformen zunächst einmal sehr wenig zu tun hat mit der Theorie der Tangentialvektorfelder.

Anhang: Zum Umgang mit den 1-Formen in der Thermodynamik.

Die Begriffe um die Mannigfaltigkeiten haben sich als geeignet erwiesen, die mathematische Thermodynamik durchsichtiger zu machen. R. Feynman beginnt in seinen berühmten Lecture Notes das Kapitel 45 „Erläuterungen zur Thermodynamik“ mit dem Satz:

„Die Thermodynamik ist ein ziemlich kompliziertes und komplexes Gebiet, wenn man sich daran macht, sie anzuwenden“. Er sagt dann weiter: „Das Gebiet der Thermodynamik ist deshalb kompliziert, weil es so viele verschiedene Arten der Beschreibung ein und derselben Sache gibt. Wenn wir das Verhalten eines Gases beschreiben wollen, können wir sagen, daß der Druck von der Temperatur und dem Volumen abhängt oder wir können sagen, daß das Volumen von der Temperatur und dem Druck abhängt. Oder was die innere Energie U betrifft, könnten wir sagen, daß sie von der Temperatur und dem Volumen abhängt, ...“.

Die hier beklagte Schwierigkeit ist keine physikalische, sondern eine mathematische. Sie betrifft die mathematische Sprache, in der die Zusammenhänge beschrieben werden. Die Schwierigkeiten ergeben sich (nach meiner Einschätzung) nur daraus, dass man nicht klar unterscheidet zwischen den Funktionen (auf einer Mannigfaltigkeit) und ihren Darstellungen in ausgezeichneten Koordinatensystemen. In den elementaren Lehrbüchern gebraucht man (leider!) bis heute den Begriff 'Funktion mehrerer Veränderlicher' synonym mit dem Begriff 'Funktionen auf einem mehrdimensionalen Raum'. Über die daraus (möglicherweise) entstehenden Schwierigkeiten wollen wir hier nun bei der Gelegenheit eines (stark vereinfachten) 'Anwendungsfall' sinnieren. Dabei müssen wir allerdings gelegentlich Begriffe aus der Welt der Mannigfaltigkeiten benutzen, die wir erst später genauer erläutern. Was die Anwendungssituation betrifft, wollen wir etwas weiter ausholen.

Die Thermodynamiker kennen den Begriff eines Einkomponentensystems. Als Beispiel nehmen wir N Mol Wasser, welche in einen Behälter eingesperrt ist. Wasser gibt es bekanntlich in drei 'Aggregatzuständen': Eis, Wasserdampf und flüssiges Wasser; unter gewissen Bedingungen können sie koexistieren. Wir denken im Folgenden an Zustände, wo alle Substanz als Dampf vorliegt.

Für das System gibt es eine Reihe bedeutsamer 'Zustandsvariabler' wie z. B. Volumen, Druck, Temperatur ... , deren Werte V, p, T in jedem Zustand verhältnismäßig einfach abgelesen werden können. Es gibt weitere Zustandgrößen, die der Messung nicht unmittelbar zugänglich sind, wie z. B. die Entropie S und die innere Energie U . Leicht gemessen werden können jedoch die Veränderungen dieser Zustandsvariablen bei Zustandsänderungen. (Wir denken hier nur an Zustandsänderungen, die langsam über sog. Gleichgewichtszustände führen.) Dem System wird mechanische Energie zugeführt, indem man (beim Druck p) komprimiert, also das Volumen verringert. Dem System wird durch Heizen thermische Energie hinzugefügt.

Energieerhaltung:

Wir erinnern an dieser Stelle daran, dass die Physiker erst im 19. Jahrhundert verstanden haben, dass Wärme kein Stoff ist sondern eine Energieform. In das Naturgesetz von der Energieerhaltung musste die Wärme als Energieform einbezogen werden.

Die dem System 'infinitesimal' zugeführte mechanische Energie wird naheliegenderweise mit $\alpha = -p \cdot dV$ bezeichnet. Die (ebenfalls leicht messbare) zugeführte thermische Energie wird heutzutage mit $\omega = T \cdot dS$ bezeichnet. (Früher benutzten Physiker und Chemiker die höchst problematische Bezeichnung ΔQ .) In einem Einkomponentensystem gibt es keine chemischen Umwandlungen, keine elektrischen Aufladung oder dergleichen. Wärme und mechanische Energie sind die einzigen Energieformen, die zu berücksichtigen sind, wenn es um die Änderungen der inneren Energie geht. Der Energieerhaltungssatz hat somit die Form

$$dU = \alpha + \omega = -p \cdot dV + T \cdot dS.$$

Wenn man schreibt $\omega = dU + p \cdot dV$, dann soll das so interpretiert werden: die durch Heizen zugeführte thermische Energie wird zum Teil zur Veränderung der inneren Energie eingesetzt und zum anderen Teil als mechanische Arbeit abgeführt. Wie die Aufteilung erfolgt, kann man wählen. Beispiele sind die isobare Zustandsänderung (der Druck wird konstant gehalten) oder die isochore Zustandsänderung (das Volumen wird konstant gehalten). Zustandsänderungen, bei welchen keine thermische Energie auf das System übertragen wird, heißen adiabatische Zustandsänderungen; man kann sie (wegen $dS = \frac{1}{T}\omega$) auch als Zustandsänderungen bei konstanter Entropie verstehen.

Die Entropie und die innere Energie:

Es ist eine Grundtatsache (oder ein 'Axiom') der Gleichgewichtsthermodynamik, dass die Entropie eine positiv homogene konkave Funktion von Volumen und innerer Energie ist:

$$S = F(V, U) \quad (\text{'Zustandsgleichung'})$$

Es ist natürlich völlig unhistorisch, wenn man den Begriff der Entropie an den Anfang der thermodynamischen Theorie stellt. Es war ein langwieriger Prozess, welcher die Naturwissenschaftler im 19. Jahrhundert zu der Erkenntnis führte, dass es Zustandsgrößen U und S gibt, die, obwohl man sie nicht direkt messen kann, wesentlichen Erklärungswert für die Phänomene der Wärmelehre haben. Über die Entdeckung des mechanischen Wärmeäquivalents (R. Mayer, J.P. Joule) gelangte man nach vielem terminologischen Hin und Her zum Begriff der (inneren) Energie, welche ein System in einem gewissen Zustand besitzt (William Thomson (Lord Kelvin), 1851). Man entdeckte den Begriff der absoluten Temperatur. R. Clausius erfand den Begriff der Entropie (1854). Mit diesen Begriffen formulierte man Hauptsätze mit dem Ziel einer axiomatisch aufgebauten Theorie.

Die ersten Einsichten zum zweiten Hauptsatz stammen aus den Überlegungen von Sadi Carnot, wie man eine Wärmekraftmaschine mit maximalem Wirkungsgrad bauen kann. Feynman meint (S. 618), es handle sich um einen der wenigen Fälle, wo Ingenieursarbeit in fundamentaler Weise zur physikalischen Theorie beigetragen hat. Wie hoch Carnot's Programm gesteckt war, zeigt das folgende Zitat aus dem Buch

Carnot, Sadi: *Betrachtungen über die bewegende Kraft des Feuers und die zur Entwicklung dieser Kraft geeigneten Maschinen*, 1824.

(Übersetzung in Ostwalds Klassikern der Exakten Wissenschaften, Bd.37)

„Das Phänomen der Erzeugung von Bewegung durch Wärme ist nicht unter einem hinlänglich allgemeinen Gesichtspunkt betrachtet worden. Man hat es nur an Maschinen untersucht, deren Wirkungsweise ihm nicht die ganze Entwicklung gestattet hat, deren es fähig ist. An derartigen Maschinen zeigt sich das Phänomen sozusagen verstümmelt und unvollständig, so dass es schwierig wird, seine Grundlagen zu erkennen und seine Gesetze zu studieren.

Um das Princip der Erzeugung von Bewegung durch Wärme in seiner ganzen Allgemeinheit zu betrachten, muss man es sich unabhängig von jedem Mechanismus und jedem besonderen Agens vorstellen; man muss Ueberlegungen durchführen, welche ihre Anwendungen nicht nur auf Dampfmaschinen haben, sondern auf jede denkbare Wärmemaschine, welches auch der angewandte Stoff sei, und in welcher Art man auf ihn einwirkt.

Die Maschinen, welche ihre Bewegung nicht von der Wärme empfangen, wie die, deren Antrieb die Kraft der Menschen oder Thiere, der Fall des Wassers, die Strömung der Luft etc., ist, können bis in ihre kleinsten Einzelheiten mittels der theoretischen Mechanik studirt werden. Alle Fälle sind vorgesehen, alle möglichen Bewegungen sind allgemeinen festbegründeten Principien unterworfen, welche unter allen Umständen Anwendung finden. Das ist das Kennzeichen einer vollständigen Theorie. Offenbar fehlt eine ähnliche Theorie bei den Wärmemaschinen. Man wird sie erst besitzen, wenn die Gesetze der Physik genügend ausgedehnt und verallgemeinert sind, um von vornherein alle Wirkungen festzustellen, welche die Wärme ausübt, wenn sie sich in bestimmter Weise an einem beliebigen Körper bethätigt.“

Carnot erläutert die Phänomene am Beispiel der gewöhnlichen Dampfmaschine und kommt dann zu einem allgemeinen Prinzip:

„Bei den geschilderten Vorgängen erkennt man leicht die Herstellung des Gleichgewichts des Wärmestoffs, seinen Uebergang von einem mehr oder weniger erhitzten Körper auf einen kälteren. . . .

Die Erzeugung von bewegender Kraft ist daher bei den Dampfmaschinen nicht sowohl auf einen wirklichen Verbrauch des Wärmestoffs zurückzuführen, sondern auf seinen Uebergang von einem heissen Körper zu einem kalten, d.h. auf die Herstellung seines Gleichgewichtes, welches durch irgend eine Ursache, eine chemische Wirkung, wie die Verbrennung, oder irgend eine andere, gestört worden war.“

Da man die zugeführte Wärme und die zugeführte mechanische Arbeit messen kann, kann man auch den Zuwachs von U und S (entlang einer Kurve) messen.

$$U_2 - U_1 = \int_{\mathfrak{c}} (\omega + \alpha); \quad S_2 - S_1 = \int_{\mathfrak{c}} \frac{1}{T} \omega.$$

Damit werden die Zustandsgrößen U und S bis auf eine additive Konstante meßbar. Es bleibt die Bestimmung der Konstanten. Feynman schreibt dazu (Seite 632): *Lange Zeit wurde geglaubt, daß die absolute Entropie nichts bedeutet - daß nur Entropiedifferenzen definiert werden könnten - aber schließlich schlug W.Nernst das vor, was wir mit Wärme-Theorem bezeichnen und was auch der dritte Hauptsatz der Thermodynamik genannt wird. . . .*

Wenn man die Existenz einer Zustandsgleichung postuliert (bei uns hier erst einmal nur für ein Einkomponentensystem), dann finden weitere Zustandsgrößen einen gesicherten Platz in der Theorie. Da die Entropie auch monoton mit der inneren Energie anwächst, kann man die Zustandsgleichung nach U auflösen; und man erhält

$$U = k(S, V) \quad \text{mit } k(\cdot) \text{ konvex und } \frac{\partial k}{\partial S} > 0 \text{ stetig.}$$

Es stellt sich heraus, dass die partiellen Ableitungen, d. h. die Koeffizienten im Differential als der negative Druck und die Temperatur zu deuten sind.

$$dU = T \cdot dS - p \cdot dV = \omega + \alpha; \quad dS = \frac{1}{T} \cdot dU + \frac{p}{T} \cdot dV.$$

Diese Formel zeigt, dass die reziproke Temperatur der Eulersche Faktor ist, welcher die 1-Form der Wärmezufuhr $\omega = dU + p \cdot dV$ zu einem Differential macht: $\frac{1}{T}\omega = dS$. In alten Zeiten, als die 1-Formen noch nicht auf dem Lehrplan standen, schrieb man $\frac{1}{T}\Delta Q = dS$.

In unserer Notation haben wir die Abhängigkeit der Quantitäten V, U, S von der Molzahl N unterdrückt; sie ist trivial, da man es hier mit sog. extensiven Größen zu tun hat. Die extensiven Quantitäten teilen sich auf, wenn man das System aufteilt. Die Zustandsgrößen p und T auf der anderen Seite sind sog. intensive Größen. Wenn man ein System mit N Mol (nach sauberer Durchmischung) aufteilt in N Systeme mit einem Mol, dann findet man in jedem Teilsystem denselben Druck und dieselbe Temperatur. Wenn wir uns also auf ein Mol beziehen, dann müssen wir (streng genommen) S als das spezifische Volumen ansprechen, U als die spezifische innere Energie und S als die spezifische Entropie.

Eine Grundannahme der mathematischen Thermodynamik, die wir noch nicht erwähnt haben, besagt, dass der Zustandsraum für ein Mol unserer Substanz eine zweidimensionale C^2 -Mannigfaltigkeit ist, wobei die Paare (V, U) , (U, S) , oder (S, V) als Koordinatensysteme fungieren können. In der Umgebung mancher Punkte der Mannigfaltigkeit können auch die Paare (V, T) oder (p, T) als lokale Koordinatensysteme dienen; es kommt bekanntlich neben der Glattheit darauf an, dass die Differentiale dort linear unabhängig sind.

(In Zuständen, in welchen diese Bedingung verletzt ist, gibt es sog. Phasenübergänge.)

Bei den verschiedenen Koordinatisierungen ergeben sich in den Differentialen ‘Koeffizienten’, die manchmal leicht interpretiert werden können, z. B. :

$$\omega = c_v \cdot dT + \Lambda_v \cdot dV \quad \text{oder} \quad \omega = c_p \cdot dT + \Lambda_p \cdot dp$$

c_v heisst die spezifische Wärme bei konstantem Volumen, c_p heisst die spezifische Wärme bei konstantem Druck. c_v gibt Auskunft, wieviel Wärme man dem System bei konstant gehaltenem Volumen zuführen muss, wenn man die Temperatur erhöhen will. Auch für den Koeffizienten Λ_v hatte man früher einen Namen: ‘latente Wärme in Bezug auf das Volumen’. Entsprechende Namen hatte man für die zweite Aufspaltung von ω .

Die Volumenänderung, die (infinitesimal!) durch Erhöhung des Drucks hervorgerufen wird, wird durch den Koeffizienten der Kompressibilität beschrieben. Dabei kommt

es darauf an, auf welchem Wege die Zustandsänderung herbeigeführt wird. Man muß z. B. die adiabatische Kompressibilität von der isothermen Kompressibilität unterscheiden. Es besteht ein interessanter Zusammenhang mit den spezifischen Wärmen, genau gesagt mit dem Quotienten $\gamma = \frac{c_p}{c_v}$.

Satz 1.2.2. *In einem Einkomponentenmodell gilt für die Kompressibilitäten*

$$\left(\frac{\partial p}{\partial V}\right)_{\text{adiab}} \cdot \left(\frac{\partial p}{\partial V}\right)_{\text{isoth}}^{-1} = \gamma = \frac{c_p}{c_v}.$$

Beweis. *Wir gehen davon aus, dass in den Punkten des betrachteten Bereich des Zustandsraum die Differentiale (dV, dT) eine Basis des Cotangentialraums bilden. Wir stellen die übrigen interessanten 1-Formen in dieser Basis dar, vorerst ohne Interpretation in der Sprache der Physik.*

$$\begin{aligned} dp &= a \cdot dV + b \cdot dT, \\ dU &= c \cdot dV + d \cdot dT, \quad \omega = (c + p) \cdot dV + d \cdot dT \\ dS &= e \cdot dV + f \cdot dT. \end{aligned}$$

Offenbar gilt $c_v = d$ und $\left(\frac{\partial p}{\partial V}\right)_T = a$. Weiter haben wir

$$1. \quad \omega = (c + p) \cdot \frac{1}{a}(dp - b \cdot dT) + d \cdot dT = \frac{c+p}{a} dp + \left(d - \frac{b(c+p)}{a}\right) dT,$$

$$\text{also} \quad \gamma = \frac{c_p}{c_v} = 1 - \frac{b(c+p)}{ad},$$

$$2. \quad dp = a \cdot dV + \frac{b}{f} \cdot (dS - e \cdot dV) = \left(a - \frac{be}{f}\right) dV + \frac{b}{f} \cdot dS,$$

$$\text{also} \quad \left(\frac{\partial p}{\partial V}\right)_S \cdot \left(\frac{\partial p}{\partial V}\right)_T^{-1} = 1 - \frac{b}{a} \cdot \frac{e}{f}$$

$$3. \quad dU = T \cdot dS - p \cdot dV = T(e \cdot de + f \cdot dT) - p \cdot dV = (Te - p) \cdot dV + Tf \cdot dT.$$

$$\text{also} \quad c = Te - p, \quad d = Tf, \quad \frac{c+p}{d} = \frac{e}{f}.$$

Damit ist die Gleichheit bewiesen. Die Kompressibilitäten stehen im Verhältnis γ .

Hinweis: Bei vielen Gasen herrscht in großen Bereichen des Zustandsraums annähernd die Gleichung des idealen Gases: $p \cdot V = R \cdot T$, wo R die berühmte universelle Gas-konstante ist. In diesem Falle gilt

$$\begin{aligned} \ln p &= -\ln V + \ln(RT); \quad d(\ln p) = -\frac{1}{V}dV + \frac{1}{T}dT, \quad \left(\frac{\partial \ln p}{\partial V}\right)_T = -\frac{1}{V} \\ \left(\frac{\partial \ln p}{\partial V}\right)_{\text{adiab}} &= \frac{1}{p} \cdot \gamma \left(\frac{\partial \ln p}{\partial V}\right)_{\text{isoth}} = -\frac{\gamma}{V}. \end{aligned}$$

In Bereichen, wo γ konstant ist, haben die Adiabaten die Gestalt $p \cdot V^\gamma = \text{const.}$

Bei manchen Substanzen ist in weiten Bereichen die innere Energie nur von der Temperatur abhängig. Es gilt also die einfache 'calorische Zustandsgleichung' :

$$dU = c_V \cdot dT + 0 \cdot dV.$$

In den Bereichen der Mannigfaltigkeit, wo dies und die Gleichung des idealen Gases gilt, haben wir $c_V = c_p + R$; denn

$$c_p dT + \Lambda_p dp = \omega = dU - p dV = c_V dT - d(pV) + V dp = (c_V - R) dT + V dp.$$

Zur Geschichte: Wenn man die Dichte $\rho = \frac{\text{Masse}}{\text{Volumen}}$ anstelle des Volumens einführt, dann wird aus der Gleichung für die Kompressibilitäten eine Beziehung, die in der frühen Geschichte der Physik große Aufmerksamkeit gefunden hat:

Newton hatte für die Schallgeschwindigkeit c in einem Gas die Formel $c^2 = \frac{dp}{d\rho}$ abgeleitet. Der bekannte Ansatz für das ideale Gas

$$p \cdot V = \text{const} \quad \text{oder} \quad p = \text{const} \cdot \rho \quad ,$$

führte nun aber zu einem Wert für $\frac{dp}{d\rho}$, welcher in der Luft nicht mit den Experimenten übereinstimmte. Die Schallgeschwindigkeit war in Wirklichkeit größer. Laplace argumentierte dann (1816), daß die Schwankungen des Luftdrucks bei der Schallfortpflanzung viel schneller sind als der Temperatúrausgleich in der Luft. Laplace kam zum richtigen Schluß, daß für die Druckschwankung infolge der Dichteschwankung nicht die isotherme Änderung des Drucks, sondern die adiabatische Druckänderung maßgeblich ist.

$$c^2 = \left(\frac{\partial p}{\partial \rho} \right)_{\text{adiab}} = \gamma \cdot \left(\frac{\partial p}{\partial \rho} \right)_{\text{isoth}} \quad \text{mit} \quad \gamma = \frac{c_p}{c_V} \quad .$$

1.3 Extremwert-Aufgaben

Das Bestimmen von Extremwerten ist ein uraltes Thema der Mathematik, fast ebenso alt wie das ‘Lösen von Gleichungen’. Die erste Veröffentlichung der Leibnizschen Form der Differential- und Integralrechnung erfolgte 1684 in einem sechsseitigen Artikel in einer von Leibniz selbst gegründeten mathematischen Zeitschrift, den „Acta Eruditorum“. Der Titel der Arbeit war (hier in deutscher Übersetzung): „Eine neue Methode für Maxima und Minima sowie für Tangenten, die durch gebrochene und irrationale Werte nicht beeinträchtigt wird, und eine merkwürdige Art des Kalküls dafür.“ Sie enthielt die Symbole dx , dy und die Differentiationsregeln einschließlich $d(uv) = u dv + v du$ sowie die Bedingung $dy = 0$ für Extremwerte. Die Leibnizsche Erklärung der Grundlagen des neuen Kalküls litt unter derselben Unbestimmtheit wie die Newtonsche. Manchmal waren seine dx , dy endliche Größen, manchmal aber Größen, die kleiner als jede angebbare Zahl aber doch nicht Null waren. — Die Differentiale gerieten unter Verdikt. Sie erschienen über 200 Jahre nur noch in Zusammensetzungen wie $\frac{dy}{dx}$ oder $\int f(x) dx$.

Der Kalkül der Differentiale df konnte sich erst im 20. Jahrhundert durchsetzen, nachdem einerseits der Raum \mathbb{R} der reellen Zahlen verstanden war und andererseits die lineare Geometrie der n -dimensionalen Räume im Sinne der Theorie von H. Graßmann (1809 - 1877) („Lineale Ausdehnungslehre“). Heute ist die Theorie der Differentialformen auf Mannigfaltigkeiten ein Herzstück von Geometrie und Analysis. Die Grundzüge dieser Theorie werden wir in dieser Anfängerveranstaltung nach und nach entwickeln. Was wir im folgenden Abschnitt sagen, kann gewiss auch mit einfacheren Worten gesagt werden; der Wert der elaborierten Sprache wird in späteren Entwicklungen deutlich werden..

Ganz im Sinne der Idee von Leibniz können wir sagen: Ist f eine differenzierbare Funktion auf einer Teilmenge des \mathbb{R}^n , dann kommen als Extremalpunkte im Inneren des Definitionsbereichs nur solche P in Betracht, in welchen das Differential verschwindet, $df|_P = 0$. Ist nämlich das Differential in einem inneren Punkt P nicht Null, so gibt es Tangentialvektoren v im Tangentialraum T_P mit $\langle df|_P, v \rangle \neq 0$; es gibt dann glatte Kurvenstücke durch P , entlang welcher die Funktion f mit positiver Geschwindigkeit ansteigt und absteigt. Diese Argumentation ist nicht nur für Zielfunktionen auf dem \mathbb{R}^n passend, sondern auch dann, wenn der Definitionsbereich von $f(\cdot)$ eine Mannigfaltigkeit ist, oder eine durch ‘Nebenbedingungen’ definierte Untermannigfaltigkeit.

Etwas komplizierter ist das Problem der Extremwertbestimmung auf den sog. ‘Mannigfaltigkeiten mit Rand’. Hier muss man sagen: Ein Randpunkt des Bereichs B kommt nur dann als Maximalpunkt in Betracht kommt, wenn

$$\langle df|_P, v \rangle \leq 0 \quad \text{für alle ins Innere von } B \text{ gerichteten Tangentialvektoren } v.$$

Wir bemerken: Die Extremalstellen am Rand des Definitionsbereichs der Zielfunktion spielen bekanntlich die entscheidende Rolle in der Theorie der linearen Optimierung, die dem Leser möglicherweise aus der Linearen Algebra bekannt ist. Dort spielen ähnliche Konstruktionen eine Rolle wie die Lagrange-Multiplikatoren, die wir im Folgenden für Extremalstellen im Inneren einer Untermannigfaltigkeit in den Blick bringen wollen.

Extremalstellen mit Nebenbedingungen; Lagrange-Multiplikatoren.

Beispiel 1.3.1.

Es sei A eine symmetrische reelle $n \times n$ -Matrix und $f(x) = x^T A x$ die dazugehörige quadratische Form. Wir suchen nach den Extremalstellen der auf die Einheitskugel $M = \{x : x^T x = 1\}$ eingeschränkten Funktion f . Wir werden sehen, dass der Maximalwert $\sup\{x^T A x : x^T x = 1\}$ der größte Eigenwert der Matrix A ist.

Der Anstieg df (auf dem $\mathbb{R}_{S^1}^n$) wird wie üblich durch einen ‘Zeilenvektor’ dargestellt; seine Einträge sind die partiellen Ableitungen, also die Richtungsableitungen in den Richtungen der ‘natürlichen’ Basisvektoren v_j .

$$\lim_{h \rightarrow 0} \frac{1}{h} [(x + hv_j)^T A (x + hv_j) - x^T A x] = v_j^T A x + x^T A v_j = 2 \cdot x^T A v_j$$

Das Differential $df|_x$ ist also durch die Zeile $2 \cdot x^T A$ dargestellt.

Wir bestimmen nun andererseits die Tangentialräume an M in den Punkten x . Diese Tangentialvektoren v sind dadurch gekennzeichnet, dass entlang von ihnen die Funktion $g(x) = x^T x$ stationär ist. Das Differential in x ist durch die Zeile $2 \cdot x^T$ gegeben. Wir suchen die Punkte x , für welche gilt

$$x^T A v = 0 \quad \text{für alle } v \text{ mit } x^T v = 0, \quad \text{d. h.} \quad \{v : x^T v = 0\} \subseteq \{v : x^T A v = 0\}$$

Die Gleichheit der Nullräume der beiden Linearformen bedeutet, dass $x^T A$ ein Vielfaches von x^T ist, d. h. dass ein λ existiert, sodass $Ax = \lambda x$. Uns interessieren die normierten x mit dieser Eigenschaft; das sind also die normierten Eigenvektoren. Der Maximalwert von f auf der Einheitskugel ist der größte Eigenwert der Matrix A .

Warnung: Die obigen Überlegungen sind ein theoretischer aber praktisch unnützer Beitrag zu dem praktisch wichtigen Problem, für eine symmetrische Matrix A den größten Eigenwert λ zu ermitteln. Es wäre unerträglich umständlich, das charakteristische Polynom zu berechnen und daraus die Eigenwerte. In der Praxis findet man λ dadurch, dass man (durch ein Iterationsverfahren) eine Maximalstelle mit der Nebenbedingung $\|x\| = 1$ ermittelt. Wenn man sich für die weiteren Eigenwerten interessiert, dann leistet die Methode von Courant gute Dienste: man sucht Maximalstellen im Raum der bisher ermittelten Eigenvektoren senkrechten Einheitsvektoren.

Die Einheitskugel ist eine besonders einfache Untermannigfaltigkeit des \mathbb{R}^n . Sie ist das Nullstellengebilde einer glatten Funktion. Auch sonst sind die Nebenbedingungen für ein Maximierungsproblem oft von der Art, dass gefordert wird, dass gewisse Funktionen einen vorgegebenen Wert annehmen. Wenn das Schnittgebilde der Niveau-Flächen dieser Funktionen genügend regulär ist, dann ist die Methode der Lagrange-Multiplikatoren in der Tat anwendbar.

Satz 1.3.1 (Lagrange-Multiplikatoren).

Es seien f, f^1, \dots, f^m glatte Funktionen auf einer n -dimensionalen Mannigfaltigkeit N und \tilde{P} ein Punkt, in welchem die Anstiege $\{df^i|_{\tilde{P}} : i = 1, \dots, m\}$ linear unabhängig sind.

Wenn die auf die Menge $\tilde{M} = \{P : f^i(P) = f^i(\tilde{P}) \text{ für } i = 1, \dots, m\}$ eingeschränkte Funktion f im Punkt \tilde{P} extremal ist, dann ist $df|_{\tilde{P}}$ eine Linearkombination der $df^i|_{\tilde{P}}$. Es existieren Zahlen λ_i (sog. Lagrange-Multiplikatoren), sodass

$$\sum \lambda_i \cdot df^i|_{\tilde{P}} = df|_{\tilde{P}}.$$

Beweis. Ist v ein Tangentialvektor in \tilde{P} , welcher alle Anstiege $df^{(i)}|_{\tilde{P}}$ annulliert, den Anstieg $df|_{\tilde{P}}$ aber nicht, dann existiert in der Untermannigfaltigkeit \tilde{M} eine Kurve durch \tilde{P} , entlang welcher f ansteigt. Wenn f unter den Nebenbedingungen in \tilde{P} extremal ist, dann ist der Covektor $df|_{\tilde{P}}$ linear abhängig von den Covektoren $df^{(i)}|_{\tilde{P}}$.

Die Anwendung des Satzes sollte man sich so vorstellen. Man suche die Punkte \tilde{P} , in welchen das lineare Gleichungssystem $\sum_i \lambda_i \cdot df^i|_{\tilde{P}} = df|_{\tilde{P}}$ eine Lösung besitzt. Wenn ein solcher Punkt die Nebenbedingungen erfüllt, dann ist er ein Kandidat für eine Extremalstelle.

Beispiel 1.3.2. Den Punkten ω in einer (endlichen) Menge Ω seien Zahlen $u(\omega)$ im Intervall $[a, b]$ zugeordnet. (a sei der minimale, b der maximale Wert der Funktion). $u(\omega)$ heisst die Energie der Konfiguration ω . Die Wahrscheinlichkeitsgewichtungen p auf Ω heissen die Zustände des Systems: $(p_\omega \geq 0, \quad G(p) = \sum p_\omega = 1.)$

Aufgabe. Zu vorgegebener Energie $U(p) = \sum u(\omega)p_\omega = u^* \in (a, b)$ soll der Zustand maximaler Entropie $S(p) = - \sum p_\omega \cdot \ln p_\omega$ ermittelt werden.

Wir verstehen das als ein Maximierungsproblem mit zwei Nebenbedingungen im $|\Omega|$ -dimensionalen Raum der Tupel $(p_\omega)_{\omega \in \Omega}$: $G(p) = 1, U(p) = u^*$. Die Anstiege der beteiligten Funktionen sind

$$dG = \sum dp_\omega; \quad dU = \sum u(\omega) dp_\omega; \quad dS = - \sum (\ln p_\omega + 1) dp_\omega.$$

Gesucht sind Lagrange-Multiplikatoren α, β : $dS|_{\tilde{p}} = \alpha \cdot dG|_{\tilde{p}} + \beta \cdot dU|_{\tilde{p}}$.

$$\begin{aligned} - \sum (\ln p_\omega + 1) dp_\omega &= \sum (\alpha + \beta u(\omega)) dp_\omega \\ \ln p_\omega + 1 &= -\alpha - \beta u(\omega) \quad \text{für alle } \omega \in \Omega \\ \tilde{p}_\omega &= \exp(1 - \alpha) \cdot \exp(-\beta u(\omega)) \end{aligned}$$

Wir müssen noch die Werte justieren: Aus $G(p) = 1$ ergibt sich

$$\tilde{p}_\omega = \frac{1}{Z(\beta)} \exp(-\beta u(\omega)), \quad \text{mit} \quad Z(\beta) = \sum \exp(-\beta u(\omega)).$$

Mit einem passenden Wert von β müssen wir jetzt versuchen, die Energie zu justieren. Die Funktion $Z(\cdot)$ heisst die Zustandssumme zur Funktion $u(\cdot)$. Sie ist bekanntlich logarithmisch konvex; und sie hat die (in der Variablen β monotone) logarithmische Ableitung

$$\frac{d}{d\beta} \ln Z(\beta) = - \sum u(\omega) \frac{1}{Z(\beta)} \exp(-\beta u(\omega)) = - \sum u(\omega) \tilde{p}_\omega = -U(\tilde{p}).$$

Was sind nun die interessanten Werte der logarithmischen Ableitung der Zustandssumme? Die möglichen Werte sind offenbar die Punkte im Intervall (\mathbf{a}, \mathbf{b}) . Wenn wir einen Wert $\mathbf{u}^* > \frac{1}{|\Omega|} \sum \mathbf{u}(\omega)$ vorgeben, dann hat das Maximierungsproblem eine positive Lösung $\beta > 0$. In diesem Fall nennt man den reziproken Wert $\frac{1}{\beta}$ die Temperatur des (im thermischen Gleichgewicht befindlichen) Systems mit der Energie \mathbf{u}^* . Es existiert genau eine Lösung des Maximierungsproblems und diese hat die Form $\tilde{p}_\omega = \frac{1}{Z(\beta)} \exp(-\beta \mathbf{u}(\omega))$.

. Die Wahrscheinlichkeitsgewichtungen dieser Form heissen die Gibbs-Verteilungen.

Bemerke: Die Rechnung hat nicht gezeigt, ob die ermittelte Gibbs-Verteilung eine Maximalstelle, eine Minimalstelle oder sonst eine stationäre Stelle des Entropie-Funktionalis ist. Das Beispiel ist wichtig genug, um es nochmals (in verallgemeinerter Form) zu behandeln. Wir wollen nun die durch die Methode der Lagrange-Multiplikatoren im Spezialfall nahegelegte Vermutung bestätigen, dass die Gibbs-Verteilung in bezug auf die Entropie eine besondere Stellung haben.—Die Rechnungen geben uns Gelegenheit, nochmals an wichtige Eigenschaften der runden konvexen Funktionen zu erinnern.

Minimale relative Entropie; ein direkter Zugang

Mit der Methode der Lagrange-Multiplikatoren konnten wir zeigen, dass die Entropiefunktion bei einer linearen Nebenbedingung nur in den Gibbs-Verteilungen stationär ist. Wir zeigen jetzt in einer direkten Herangehensweise, die sich auf Konvexitätsargumente stützt, dass die Gibbs-Verteilungen tatsächlich Stellen des globalen Extremums sind.

Satz 1.3.2 (Minimale relative Entropie).

Es sei $\{\mathbf{q}_\omega : \omega \in \Omega\}$ eine überall positive Wahrscheinlichkeitsgewichtung auf der abzählbaren Menge Ω . $\mathbf{u}^1, \dots, \mathbf{u}^m$ sei ein m -Tupel reellwertiger Funktionen auf Ω , (als eine m -Spalte $\mathbf{U}(\omega)$ notiert), sodass für eine offene Umgebung V des Nullpunkts (im Zeilenraum \mathbb{R}_{Sp}^m) die Funktion $M(\theta) = \sum_{\omega} e^{\langle \theta, \mathbf{U}(\omega) \rangle} \cdot \mathbf{q}_\omega$ endliche Werte hat.

Zu dieser logarithmisch konvexen Funktion definieren wir

$$\psi(\theta) = \ln M(\theta), \quad \text{und} \quad K(\mathbf{w}) = \sup\{\langle \theta, \mathbf{w} \rangle - \psi(\theta) : \theta \in V\}$$

Die Ableitungen $\psi'(\cdot)$ und $K'(\cdot)$ sind dann zueinander inverse Abbildungen

$$\psi'(\cdot) : V \rightarrow W \subseteq \mathbb{R}_{\text{Sp}}^m; \quad K'(\cdot) : W \rightarrow V \subseteq \mathbb{R}_{\text{Z}}^m$$

Für jedes $\mathbf{w} \in W$ gibt es genau eine Gibbsverteilung

$$\tilde{p}_\omega = \mathbf{q}_\omega \cdot e^{\langle \tilde{\theta}, \mathbf{U}(\omega) \rangle} \cdot e^{-\psi(\tilde{\theta})} \quad \text{mit} \quad \sum_{\omega} \mathbf{U}(\omega) \cdot \tilde{p}_\omega = \mathbf{w}.$$

Auf dem Simplex der Wahrscheinlichkeitsgewichtungen $\mathbf{p} = \{\mathbf{p}_\omega : \omega \in \Omega\}$ betrachten wir die ‘relative Entropie’ $R(\cdot ; \mathbf{q})$

$$R(\mathbf{p} ; \mathbf{q}) = \sum_{\omega} \mathbf{p}_\omega \cdot \ln \left(\frac{\mathbf{p}_\omega}{\mathbf{q}_\omega} \right) = \sum_{\omega} \mathbf{q}_\omega \cdot k \left(\frac{\mathbf{p}_\omega}{\mathbf{q}_\omega} \right) \quad \text{mit } k(x) = x \cdot \ln x$$

Es gilt dann $R(\mathbf{p} ; \mathbf{q}) \geq R(\tilde{\mathbf{p}} ; \mathbf{q}) = K(\mathbf{w})$ für alle \mathbf{p} mit $\sum_{\omega} \mathbf{U}(\omega) \cdot \mathbf{p}_\omega = \mathbf{w}$.

Beweis. Nach Jensens Ungleichung gilt $R(p; q) > 0$; für alle $p \neq q$; denn

$$\sum_{\omega} q_{\omega} \cdot k\left(\frac{p_{\omega}}{q_{\omega}}\right) \geq k\left(\sum_{\omega} q_{\omega} \frac{p_{\omega}}{q_{\omega}}\right) = 0.$$

Für die Gibbs-Verteilung \tilde{p} zu $\theta = K'(w)$ gilt

$$\sum U(\omega) \cdot \tilde{p}_{\omega} = e^{-\psi(\theta)} \sum q_{\omega} \frac{d}{d\theta} e^{\langle \theta, U(\omega) \rangle} = \frac{M'(\theta)}{M(\theta)} = \psi'(\theta).$$

$$R(\tilde{p}; q) = \sum \tilde{p}_{\omega} \ln\left(\frac{\tilde{p}_{\omega}}{q_{\omega}}\right) = \sum \tilde{p}_{\omega} (\langle \theta, U(\omega) \rangle - \psi(\theta)) = \langle \theta, w \rangle = K(w).$$

Für jedes p mit $\sum p_{\omega} U(\omega) = w = \psi'(\tilde{\theta})$ gilt aber

$$\begin{aligned} R(p; q) &= \sum p_{\omega} \cdot \ln\left(\frac{p_{\omega}}{p_{\omega}} \cdot \frac{\tilde{p}_{\omega}}{q_{\omega}}\right) = R(p; \tilde{p}) + \sum p_{\omega} \cdot \langle \tilde{\theta}, U(\omega) \rangle - \psi(\tilde{\theta}) \\ &= R(p; \tilde{p}) + \langle \tilde{\theta}, \psi'(\tilde{\theta}) \rangle - \psi(\tilde{\theta}) = R(p; \tilde{p}) + K(\tilde{w}) = R(p; \tilde{p}) + R(\tilde{p}; q). \end{aligned}$$

Die Idee der Variationsrechnung

Wir wenden uns in heuristischer Manier einem klassischen Extremalwertproblem in einem Funktionenraum zu. Genauer gesagt, geht es um optimale parametrisierte Kurven. Die Idee, dass ein Funktional in einer Extremposition verschwindende Richtungsableitungen hat, führt hier auf ein System von Differentialgleichungen. Ob die Lösungen der Differentialgleichungen das Extremwertproblem lösen, muss dann auf anderen Wegen diskutiert werden.

Wir betrachten glatt parametrisierte Kurven, die in einem 'Zeitintervall' $[t_0, t_1]$ von einem vorgegebenen Punkt q_0 auf einer n -dimensionalen Mannigfaltigkeit N zu einem vorgegebenen Punkt q_1 führen. Die Mannigfaltigkeit darf man sich vorerst gerne als eine Mannigfaltigkeit mit einer globalen Koordinatisierung vorstellen. Wir wollen uns aber nicht an ein bestimmtes Koordinatensystem binden, und gehen daher nicht davon aus, dass N eine Teilmenge des \mathbb{R}^n ist. Eine glatt parametrisierte Kurve auf einer Mannigfaltigkeit $\gamma(\cdot)$ hat zu jedem Zeitpunkt eine Geschwindigkeit. Wir wollen sie als parametrisierte Kurve im Tangentialbündel betrachten. Die betrachteten Objekte sind also von der Gestalt

$$\{(\gamma(t), \dot{\gamma}(t)) : t \in [t_0, t_1]\} \quad \text{mit} \quad \gamma(t_0) = Q_0, \gamma(t_1) = Q_1.$$

Wir werden für gewisse Funktionale \mathfrak{S} ('Wirkungsfunktionale' genannt) Extrempunkte bestimmen. Es sei $L(t, q, \dot{q})$ eine 'zeitabhängige' Funktion auf dem Tangentialbündel und

$$\mathfrak{S}(\gamma(\cdot)) = \int_{t_0}^{t_1} L(t, \gamma(t), \dot{\gamma}(t)) dt.$$

In einer rigorosen Theorie müsste man Glattheitsforderungen spezifizieren; solche zu präzisieren, wollen wir hier aber nicht versuchen; wir werden uns unten auf eine sehr spezielle Situation konzentrieren, in welcher die Glattheitsfragen nicht vordringlich sind.

Eine Kurve $\tilde{\gamma}(\cdot)$ kann allenfalls dann ein Extrempunkt für das Wirkungsfunktional sein, wenn die Richtungsableitungen verschwinden. Wenn wir die Kurve $\tilde{\gamma}(\cdot)$ ein wenig stören, dann notieren wir (ziemlich unpräzise) die gestörte Kurve $\gamma_\varepsilon(\cdot) = \tilde{\gamma}(\cdot) + \varepsilon \cdot \mathbf{h}(\cdot)$, wobei $\mathbf{h}(\mathbf{t}_0) = \mathbf{0} = \mathbf{h}(\mathbf{t}_1)$. Die Richtungsableitung in der durch $\mathbf{h}(\cdot)$ gegebenen Richtung ergibt sich als Grenzwert für $\varepsilon \rightarrow 0$

$$\begin{aligned} & \frac{1}{\varepsilon} [\mathfrak{S}(\gamma(\cdot) + \varepsilon \cdot \mathbf{h}(\cdot)) - \mathfrak{S}(\gamma(\cdot))] = \\ &= \frac{1}{\varepsilon} \int_{\mathbf{t}_0}^{\mathbf{t}_1} \left[\mathbf{L}(\mathbf{t}, \gamma(\mathbf{t}) + \varepsilon \cdot \mathbf{h}(\cdot), \dot{\gamma}(\mathbf{t}) + \varepsilon \cdot \dot{\mathbf{h}}(\cdot)) - \mathbf{L}(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) \right] dt \\ &\longrightarrow \int_{\mathbf{t}_0}^{\mathbf{t}_1} \left[\mathbf{L}_1(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) \cdot \mathbf{h}(\mathbf{t}) + \mathbf{L}_2(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) \cdot \dot{\mathbf{h}}(\mathbf{t}) \right] dt \\ &= \int_{\mathbf{t}_0}^{\mathbf{t}_1} \left[\mathbf{L}_1(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) - \frac{d}{dt} \mathbf{L}_2(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) \right] \cdot \mathbf{h}(\mathbf{t}) dt. \end{aligned}$$

Die letzte Umformung ergibt sich durch partielle Integration. Der Grenzübergang beruht auf dem Prinzip der Differentiation unter dem Integralzeichen; dieses bedarf im konkreten Fall einer Begründung.

Wenn nun diese Integrale für eine genügend große Kollektion von zulässigen Störfunktionen $\mathbf{h}(\cdot)$ verschwindet, dann impliziert das das Verschwinden des Faktors in eckigen Klammern. Man beachte: Wenn es um Kurven auf einer n -dimensionalen Mannigfaltigkeit geht, dann sind die $\mathbf{h}(\cdot)$ n -Tupel von Funktionen; und das Verschwinden der Klammer ist ein n -Tupel von Differentialgleichungen. Diese n Forderungen an eine Extremalkurve $\gamma(\cdot)$ heissen die Euler'schen Differentialgleichungen der Variationsrechnung, oder auch die Gleichungen von Euler-Lagrange.

$$\frac{d}{dt} \mathbf{L}_2(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})) = \mathbf{L}_1(\mathbf{t}, \gamma(\mathbf{t}), \dot{\gamma}(\mathbf{t})).$$

Die berühmtesten Anwendungen hat man in der sog. Lagrange'schen Mechanik. Die Lagrange'sche Mechanik befasst sich mit konservativen mechanischen Systemen. Ein Lagrange'sches System ist gegeben durch eine glatte Mannigfaltigkeit $(\mathbf{N}, \mathcal{D}_{\mathbf{N}})$ und eine Funktion $\mathbf{L}(\mathbf{t}, \mathbf{q}, \dot{\mathbf{q}})$ auf dem Tangentialbündel. Das Hamilton'sche Prinzip der kleinsten Wirkung besagt nun

Die nach den Gesetzen der Mechanik ablaufende Bewegung eines Systems zeichnet sich dadurch aus, dass das zwischen den Lagen $(\mathbf{t}_0, \mathbf{q}_0)$ und $(\mathbf{t}_1, \mathbf{q}_1)$ des Systems gewonnene Bahnintegral der Lagrange-Funktion, das Wirkungsintegral, einen Extremwert annimmt.

Die einfachsten Fälle sind die sog. Newton'schen Systeme; dort ist die Lagrange-funktion die Differenz von kinetischer und potentieller Energie.

Beispiel 1.3.3 (Ebene Bewegung in einem zentralen Kraftfeld).

Wir beginnen mit der Vorstellung: Die Sonne mit der großen Masse M steht unbeweglich im Zentrum des cartesischen Koordinatensystems (x, y) ; sie übt auf die Erde mit der kleinen Masse m eine Kraft aus gemäß dem Gravitationspotential $U(r) = -\gamma \cdot m \cdot M \cdot r^{-1}$. (γ ist die Gravitationskonstante.) Man sagt, dass es sich um ein ‘autonomes’ Newton’sches System handelt, weil die Lagrangefunktion nicht von der Zeit abhängt:

$$L(q, \dot{q}) = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - U(r) = \frac{1}{2}(\dot{r}^2 + r^2 \cdot \dot{\phi}^2) - U(r).$$

Wir wählen auf dem Tangentialbündel das Koordinatensystem $(r, \phi, \dot{r}, \dot{\phi})$

$$\begin{aligned} dL(r, \phi, \dot{r}, \dot{\phi}) &= L_1(r, \phi, \dot{r}, \dot{\phi}) \cdot \begin{pmatrix} dr \\ d\phi \end{pmatrix} + L_2(r, \phi, \dot{r}, \dot{\phi}) \cdot \begin{pmatrix} d\dot{r} \\ d\dot{\phi} \end{pmatrix} \\ &= (r \cdot \dot{\phi}^2 - U'(r), 0) \cdot \begin{pmatrix} dr \\ d\phi \end{pmatrix} + (\dot{r}, r^2 \dot{\phi}) \cdot \begin{pmatrix} d\dot{r} \\ d\dot{\phi} \end{pmatrix}. \end{aligned}$$

Die Euler-Lagrange-Gleichungen lauten $\frac{d}{dt}(\dot{r}, r^2 \dot{\phi}) = (r \cdot \dot{\phi}^2 - U'(r), 0)$.

Die zweite Gleichung liefert die zeitliche Invarianz des Drehimpulses: $r^2 \cdot \dot{\phi} = M$.

Damit erhält die erste Gleichung die Gestalt

$$\ddot{r} = r \cdot \dot{\phi}^2 - U'(r) = M^2 \cdot r^{-3} - U'(r) = -V'(r)$$

mit $V(r) = U(r) + \frac{M^2}{2r^2}$, dem sog. effektiven Potential. Man sagt: Die radiale Beschleunigung ist gleich der vom effektiven Potential erzeugten Kraft: $\ddot{r} = -V'(r)$.

Differentialgleichungen dieser Art werden wir später diskutieren.

Der Fall $V(r) = -c \frac{1}{r} + \frac{M^2}{2r^2}$ heisst Kepler’s Problem; wir werden die Lösungskurven später ‘explizit’ berechnen. Wir werden sehen, dass der Radius zeitlich periodisch ist; er schwankt in monotoner Weise hin und her zwischen einem maximalen und einem minimalen Wert. Interessant ist natürlich auch die Bahn; insbesondere ist interessant, wie der Winkel sich ändert zwischen den Punkten extremaler Entfernung, dem Perihel und dem Aphel.

Beispiel 1.3.4 (Die Brachystochrone).

Johann Bernoulli (1667 -1748), der Bruder von Jacob Bernoulli (1654 -1705) und Lehrer von L. Euler (1707-1783) war neben seinem Bruder ein überragender Pionier der Analysis im Stile von Leibniz. Auf Grund seiner Arbeit zum Problem der Brachystochrone gilt er u.a. als Entdecker der Variationsrechnung. In diesem Problem sucht man die (ebene) Kurve der kürzesten Fallzeit T für einen Massepunkt, der sich unter dem Einfluss des Schwerfelds bewegt. Der Massepunkt gleitet aus der Höhe \tilde{z} entlang der Kurve zu einem Punkt in der horizontalen Entfernung \tilde{x} auf der Höhe 0 . Wenn man die horizontale Entfernung von der Ausgangslage als Funktion der Höhe z betrachtet, dann sucht man also eine monotone Funktion $f(z)$ auf dem Intervall $[0, h]$ mit Randwerten $x(0) = 0$, $x(\tilde{z}) = \tilde{x}$. Die Integrationsvariable ist nicht wie in der Lagrange’schen Mechanik die Zeit, sondern der Höhenverlust. Die Geschwindigkeit $v(z)$ des Massepunkts ergibt sich aus dem Energie-Erhaltungssatz: nach dem Höhenverlust z ist die potentielle Energie mgz in kinetische

Energie $\frac{1}{2}mv(z)^2$ verwandelt. (g ist die Erdbeschleunigung. Die Ableitung der Kurve bestimmt das Verhältnis von Horizontalgeschwindigkeit zur Sinkgeschwindigkeit

$$\begin{aligned} \frac{x'(s)}{z'(s)} &= f'(z(s)) & v(z)^2 &= z'(s)^2 [1 + f'(z)^2] \\ \frac{dz}{ds} &= \frac{v(s)}{\sqrt{1 + f'^2(z(s))}}, & T &= \int_0^h \frac{\sqrt{1 + f'^2(z)}}{v(z)} dz. \end{aligned}$$

Das Zielfunktional ist also das Integral einer z -abhängigen Funktion auf dem Tangentialraum einer eindimensionalen Mannigfaltigkeit

$$L(z, x, \dot{x}) = \frac{1}{\sqrt{2g}} \cdot \frac{\sqrt{1 + \dot{x}^2}}{\sqrt{z}}, \quad T = \int_0^h L(z, x(z), \dot{x}(z)) dz.$$

Wegen $L_1(\cdot, \cdot, \cdot) = 0$ besagen hier die Euler-Lagrange-Gleichungen

$$\text{const} = L_2(z, x(z), \dot{x}(z)) = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2} \cdot \sqrt{z}}.$$

Dies ist (für jede der Konstanten) eine Differentialgleichung erster Ordnung für die Funktion $x(z)$ mit den Randbedingungen $x(0) = 0$, $x(h) = \tilde{x}$. Welche der Lösungskurven für das Brachistochronenproblem in Betracht kommen, muß noch diskutiert werden.

Wir werden in den Übungen das Resultat von Johann Bernoulli bestätigen, dass es sich um sog. Zykloiden handelt; das sind elementar darstellbare Funktionen, die im 17.-ten Jahrhundert aus verschiedenen Zusammenhängen gut bekannt waren. Insbesondere Descartes (1596-1650) und B. Pascal (1623-1662) hatten sie untersucht (siehe Struik, S. 114). Es wird übrigens berichtet, dass Newton (1643 - 1727) das im Jahr 1696 von Johann Bernoulli formulierte Brachystochronenproblem mit seinen Methoden umgehend lösen konnte.

Didaktische Anmerkung:

Es ist charakteristisch für die Frühzeit der Differential- und Integralrechnung, daß Problemlösungen darin bestanden, aus dem (ziemlich beschränkten) Vorrat an etablierten Kurven die richtige herauszufinden. Nach diesem Prinzip verlaufen heutzutage noch viele Übungsaufgaben zur elementaren Analysis. Der Charme dieser Art von Analysis liegt darin, daß man immer wieder alte Bekannte trifft. Dagegen ist zu bedenken, daß sich angesichts des allgemeinen Funktionsbegriffs die Anliegen der modernen Analysis verschoben haben. Man erwartet nicht, daß die Probleme durch altbekannte spezielle Funktionen gelöst werden. Man denkt gleichzeitig allgemeiner und praktischer. Man interessiert sich einerseits für die allgemeinen Prinzipien der mathematischen Konstruktion; andererseits entwickelt man effiziente Rechenverfahren, die für die konkreten praktischen Probleme realistische Näherungslösungen liefern.

Zur Vorgeschichte

Physikalische Minimalprinzipien haben eine Geschichte, die bis vor Leibniz zurückreicht. Fermat (um 1650) stellte fest („Fermats Prinzip“)

Die Ausbreitung des Lichts erfolgt stets so, daß die optische Weglänge (= geometrische Weglänge \times Brechungsindex des Mediums) gegenüber benachbarten, nicht den Gesetzen der Optik gehorchenden Wegen, ein Extremum hat.

Man bemerke: In den sog. isotropen Medien hängt der Brechungsindex nur vom Ort ab. Aber auch in anisotropen Medien, wo der Brechungsindex von der Richtung abhängt, hat der Begriff der optischen Weglänge einen guten Sinn. Man könnte die Weglänge als ein Kurvenintegral schreiben (welches in diesem speziellen Fall nicht von der Parametrisierung der Kurve abhängt). In der elementaren geometrischen Optik ist das aber nicht üblich. Man denkt beim Fermat'schen Prinzip nämlich zunächst einmal nur an stückweise homogene Medien, die in glatten Grenzflächen aneinandergrenzen. Der Brechungsindex ist dann entlang der Strahlen stückweise konstant und die Strahlen der geometrischen Optik setzen sich aus Geradenstücken zusammen. Man kann das auf den Fall eines stetig variierenden Brechungsindex verallgemeinern. Es ist technisch aufwendig und unübersichtlich, wenn man versucht, die stückweise lineare Minimierungstheorie mit der glatten Theorie der Lagrangefunktionen zu vereinigen.

(In den obigen Überlegungen haben wir angenommen dass $L(\cdot)$ zweimal stetig differenzierbar ist.)

Kürzeste krumme Kurven haben schon Archimedes beschäftigt. Archimedes (287?-212) setzte sich über die These von Aristoteles (384-322) hinweg, daß man Gerades und Krummes nicht vergleichen könne und entwickelte in seinen Schriften „Kreismessung“ und „Kugel und Zylinder“ eine Theorie der Bogenlänge. Er entwickelte einige Prinzipien, wie man Bogenlängen abschätzen kann. Archimedes definierte schon ganz allgemein, was konvexe Kurven und konvexe Flächen sind; und er erkannte Prinzipien, wie z.B. das folgende: Seien P und Q Punkte auf einer konkaven Kurve; dann sind alle Verbindungen oberhalb dieser Kurve länger als das konkave Kurvenstück selber (siehe Walter II, S. 152).

Der allgemeine Kurvenbegriff geriet dann aber in Vergessenheit. Erst die Theorien von Huygens (1629-1695), Newton, Leibniz und den Brüdern Bernoulli brachten im 17. Jahrhundert den großen Aufschwung für eine Theorie der glatten Kurven. Die großen Meister waren dann Euler und Lagrange, bis die Begeisterung für die speziellen ‘expliziten’ Kurven schliesslich um 1800 stark nachliess.

1.4 Der Begriff der Mannigfaltigkeit

Eine offene Teilmenge U des \mathbb{R}^n ist eine n -dimensionale Mannigfaltigkeit, die insofern besonders einfach erscheint, dass sie eine ‘natürliche’ globale Koordinatisierung aufweist. Bei allgemeineren Mannigfaltigkeiten braucht man eine Fülle von sog. lokalen Koordinatisierungen, um festzulegen, was die r -mal stetig differenzierbaren Funktionen auf der Mannigfaltigkeit sind. Man kreierte dafür den Begriff des Atlas mit lokalen Karten. In konkreten Fällen ist die Frage der Glattheit allerdings oftmals gänzlich unproblematisch.

Beispiel 1.4.1 (3-Sphäre).

Die Menge der reellen Quadrupel (x^0, x^1, x^2, x^3) mit $\sum (x^k)^2 = 1$ ist die 3-Sphäre, eingebettet in den euklidischen \mathbb{R}^4 . Als lokale Koordinaten in einer Umgebung U eines Punktes \tilde{P} kommen Tripel der x^k in Betracht, wo darauf zu achten ist, dass keine dieser drei Variablen in U den Extremalwert 1 annimmt. — Es sollte klar sein, was es bedeutet, dass eine Funktion f auf der 3-Sphäre eine glatte Funktion ist.

Beispiel 1.4.2 (2-Torus).

Den 2-Torus kann man präsentieren als das Quadrat $\{(\phi, \psi) : |\phi| \leq \pi, |\psi| \leq \pi\}$, in welchem die gegenüberliegenden Seitenlinien identifiziert sind. Die glatten Funktionen auf dem 2-Torus kann man verstehen als die doppelt- 2π -periodischen Funktionen $f = F(\phi, \psi)$. Beispiele sind die Realteile der trigonometrischen Polynome $F(\phi, \psi) = \sum_{mn} c_{mn} e^{im\phi + in\psi}$. (Die Summe ist über endlich viele Paare ganzer Zahlen (m, n) zu erstrecken.) Die glatten Funktionen auf dem Torus kann man durch solche ‘trigonometrischen Polynome approximieren.— Wir bemerken, dass man zwar hinreichende Bedingungen kennt, welche ‘Koeffizientenfolgen’ (c_{mn}) zu \mathcal{C}^r -Funktionen führen. Eine genaue Abgrenzung erscheint aber nicht in Reichweite. Man kann nicht einmal kurz und knapp sagen, welche Koeffizientenfolgen in einer trigonometrischen Reihe zu einer stetigen Funktion führen.

Wir geben nun eine formale Definition des Begriff der Mannigfaltigkeit nach dem Muster, welches wir bereits früher benützt haben. Beispielsweise:

1. Eine Menge S wird zu einem metrischen Raum, indem man eine Metrik $d(\cdot, \cdot)$ auszeichnet. Von einer Metrik ist zu fordern: ...
2. Eine Menge S ist ein Hausdorff-Raum mit abzählbarer Basis (HRaB), wenn ein Mengensystem \mathcal{U} als das System der offenen Mengen ausgezeichnet ist, wobei gefordert ist, dass \mathcal{U} die folgenden Eigenschaften besitzt: ...

Definition 1.7 (\mathcal{C}^r -Mannigfaltigkeit).

Ein HRaB M trägt die Struktur einer n -dimensionalen \mathcal{C}^r Mannigfaltigkeit, wenn ein System \mathcal{D} stetiger reellwertiger Funktionen als das System der glatten Funktionen ausgezeichnet ist. Von \mathcal{D} ist zu fordern:

- (i) (Lokalität) Wenn für eine Funktion f auf M zu jedem $P \in M$ eine Umgebung existiert, auf welcher sie mit einer glatten Funktion übereinstimmt, dann gehört sie selbst zu \mathcal{D} .

- (ii) (Komponierbarkeit) Sind $g^{(1)}, \dots, g^{(N)} \in \mathcal{D}$, so gilt $f = F(g^{(1)}, \dots, g^{(N)}) \in \mathcal{D}$ für jede \mathcal{C}^r -Funktion $F(\cdot, \dots, \cdot)$ auf dem \mathbb{R}^N .
- (iii) (Existenz von Karten) Zu jedem $P \in M$ existiert eine Umgebung U und ein n -Tupel von glatten Funktionen $\{x^j : j = 1, \dots, n\}$, welches U bijektiv auf eine offene Menge des \mathbb{R}^n abbildet, sodass gilt: Eine Funktion g auf U ist genau dann die Einschränkung einer glatten Funktion auf U , wenn eine \mathcal{C}^r -Funktion $G(\cdot, \dots, \cdot)$ existiert, sodass $g(\cdot) = G(x^1(\cdot), \dots, x^n(\cdot))$ auf U .

Sprechweise 1.4.1. Ein Tupel $\{x^j : j = 1, \dots, n\}$ wie in (iii) nennt man ein lokales Koordinatensystem auf U . Man versteht es als eine Abbildung $\varphi : U \rightarrow \varphi(U) \subseteq \mathbb{R}_{\text{Sp}}^n$ und bezeichnet die Umkehrabbildung mit φ^{-1} . Das Paar (U, φ) heisst eine Karte von U .

Eine Familie von Karten $\{(U_\alpha, \varphi_\alpha) : \alpha \in I\}$ heisst ein Atlas für die Mannigfaltigkeit, wenn $M = \bigcup U_\alpha$.

(Die Karten heissen im Englischen ‘charts’; das Wort ‘map’ bezeichnet Abbildung in einem allgemeinen Sinn: z. B. *a chart (U, φ) (in an atlas for an n -dimensional real manifold M) is a diffeomorphic map of an open set $U \subseteq M$ onto an open subset of \mathbb{R}^n .*)

Eine Karte (U, φ) werden wir häufig als eine lokale Koordinatisierung auf U verstehen: $\{x^j : j \in J\}$ auf U und $\{y^i : i \in I\}$ auf V sind die bevorzugten Bezeichnungen für Karten.

Man braucht die Karten, um festzulegen, was die glatten Funktionen auf einer glatten Mannigfaltigkeit sind. \mathcal{C}^r -Glattheit (oder einfach r -Glattheit) meint r -malige stetige Differenzierbarkeit. Die unendlich oft differenzierbaren Funktionen auf einer \mathcal{C}^∞ -Mannigfaltigkeit heissen auch die \mathcal{C}^∞ -Funktionen.

Die folgenden Sätze formulieren wir für n -dimensionale \mathcal{C}^1 -Mannigfaltigkeiten (N, \mathcal{D}_N) . Die Übertragung auf \mathcal{C}^r -Mannigfaltigkeiten liegt auf der Hand.

Satz 1.4.1 (Koordinatenwechsel).

Es sei $\varphi = \{x^j : j \in J\}$ ein lokales Koordinatensystem auf $U \ni \tilde{P}$, und $F(\cdot) = \{f^{(i)}(\cdot) : i \in I\}$ bei $\tilde{x} = \varphi(\tilde{P})$ stetig differenzierbar mit nichtsingulärer Jacobi-Matrix $(f^i_j)(\tilde{P})$. Dann ist auch die Familie $\{y^i : i \in I\} = \{f^{(i)}(\varphi(\cdot)) : i \in I\}$ ein lokales Koordinatensystem bei \tilde{P} (auf einer evtl. verkleinerten Umgebung \tilde{U} .)

Beweis. Es gilt zu zeigen, dass man in einer Umgebung von \tilde{x} die Abbildung $F(\cdot) = \{f^{(i)}(\cdot) : i \in I\}$ invertieren kann mit einer stetig differenzierbaren Umkehrabbildung $G(\cdot) = \{g^{(j)}(\cdot) : j \in J\}$. Wenn man nämlich die Funktionen x^j mit den y^i darstellen will, dann benötigt man $g^{(j)}(\cdot)$, sodass $x^j = g^{(j)}(y^1, \dots, y^n)$ lokal bei $\tilde{y} = F(\tilde{x})$. Eine auf U glatte Funktion $h = H(x^1, \dots, x^n)$ besitzt dann auch eine Darstellung mit dem Tupel der Koordinatenfunktionen $\{y^i : i \in I\}$: $h = H(g^1(y), \dots, g^n(y)) = K(y^1, \dots, y^n)$.

Der Beweis ergibt sich aus dem berühmten Satz von der stetig differenzierbaren Umkehrabbildung, welchen wir bereits bei dem verallgemeinerten Newton-Verfahren (und dem Banach’schen Fixpunktsatz) diskutiert haben.

Satz 1.4.2 (Lokale Umkehrbarkeit einer glatten Abbildung).

In einer Umgebung von $\tilde{x} \in \mathbb{R}^n$ sei $F(\cdot)$ stetig differenzierbar mit einer nichtsingulären Jacobi-Matrix $F'(\tilde{x})$. Mit $\tilde{A} := F'(\tilde{x})$ und $\tilde{y} = F(\tilde{x})$ gilt dann:

Es existiert eine Umgebung V von \tilde{y} und dort eine surjektive Abbildung $G(\cdot)$ auf eine Umgebung \tilde{U} von \tilde{x} , sodass gilt

$$F(G(y)) = y \quad \text{für alle } y \in V, \quad G(F(x)) = x \quad \text{für alle } x \in U.$$

Beweis. Wir erinnern an den Beweis: Man konstruiert eine Funktionenfolge $G^{(n)}$ in einer (noch einzugrenzenden) Umgebung \tilde{V} von \tilde{y} :

$$G^{(0)}(y) \equiv \tilde{x}, \quad G^{(n+1)}(y) = G^{(n)}(y) - \tilde{A}^{-1} \cdot (F(G^{(n)}(y)) - y) = \Psi(G^{(n)}(y)).$$

Die Abbildung Ψ wirkt kontrahierend bzgl. der Supremumsnorm auf dem Raum der stetigen Funktionen über \tilde{V} , genauer gesagt, auf dem Raum derjenigen stetigen Abbildungen von \tilde{V} , die im Punkt \tilde{y} den Wert \tilde{x} annehmen. Die Funktionenfolge $G^{(n)} = \Psi^n(G^{(0)})$ ist eine Cauchy-Folge bzgl. der gleichmäßigen Konvergenz auf einer Umgebung $V \subseteq \tilde{V}$ von \tilde{y} . Der Limes ist ein 'Fixpunkt' G . Für diesen Fixpunkt gilt $F(G(y)) - y = 0$ für alle $y \in \tilde{V}$. Nach der Kettenregel gilt nun $F'(G(y)) \cdot G'(y) = I$. Die Ableitung $G'(\cdot)$ ist eine stetige Funktion. Die Funktion $G(\cdot)$ selber ist stetig differenzierbar.

Der Satz vom Koordinatenwechsel ergänzt einen Satz aus dem vorigen Teilabschnitt. Dort war an ein n -Tupel von glatten Funktionen $\{y^i : i \in I\}$ auf $U \subseteq \mathbb{R}^n$ die Forderung gestellt, dass die Abbildung $y(\cdot)$ injektiv ist und das Tupel der Differentiale linear unabhängig. Hier haben wir nun (mit Hilfe des Satzes von der glatten Umkehrabbildung) festgestellt, dass die lineare Unabhängigkeit der Differentiale in einem Punkt, d. h. die Invertierbarkeit der Jacobi-Matrix in diesem Punkt, die Injektivität in einer (genügend kleinen) Umgebung impliziert und dass die Umkehrabbildung dort glatt ist. Sie ist \mathcal{C}^r -glatt, wenn $y(\cdot)$ \mathcal{C}^r -glatt ist.

Definition 1.8 (Glatte Abbildung).

Eine Abbildung χ einer \mathcal{C}^s -Mannigfaltigkeit (M, \mathcal{D}_M) in eine \mathcal{C}^r -Mannigfaltigkeit (N, \mathcal{D}_N)

$$\chi : (M, \mathcal{D}_M) \longrightarrow (N, \mathcal{D}_N)$$

heißt eine glatte Abbildung, wenn die Pullback-Abbildung glatte Funktionen in glatte Funktionen abbildet, d. h.

$$\chi^*(f) \in \mathcal{D}_M \quad \text{für alle } f \in \mathcal{D}_N.$$

Die Definition der Glattheit lässt $s \leq r$ zu und auch, dass die beiden Mannigfaltigkeiten verschiedene Dimensionen besitzen, etwa $\dim M = m$, $\dim N = n$. Wenn man zwei glatte Abbildungen hintereinanderschaltet, erhält man eine glatte Abbildung.

Um die Glattheit einer Abbildung

$$\chi : (M, \mathcal{D}_M) \ni Q \longmapsto \chi(Q) \in (N, \mathcal{D}_N)$$

einer s -glatten Mannigfaltigkeit in eine r -glatte Mannigfaltigkeit zu erweisen ($s \leq r$), genügt es, die s -Glattheit der zurückgenommenen Funktionen $\chi^*(x^j)$ für irgendein Tupel von Koordinatenfunktionen zu beweisen. Das bedeutet, dass die Funktionen, die die Koordinaten der Bildpunkte durch das m -Tupel der Koordinaten der Urbildpunkte ausdrücken, glatt sind. Für die Glattheit ist also zu fordern, dass (für irgendeine Koordinatisierung $\{y^i : i \in I\}$) gilt

$$x^j(\chi(Q)) = G^{(j)}(y^1(Q), \dots, y^m(Q)) \quad \text{mit } s\text{-glatten } G^j(\cdot).$$

Die $n \times m$ -Matrix mit den Einträgen $\frac{\partial G^j}{\partial y^i}(Q)$ heisst die Funktionalmatrix oder Jacobi-Matrix zur Abbildung χ bzgl. der betreffenden Koordinatensysteme. Die Einträge sind $(s-1)$ -glatte Funktionen. Warnung: Die Jacobi-Matrix einer Abbildung darf nicht verwechselt werden mit der Jacobi-Matrix zu einem Koordinatenwechsel.

Da die Glattheit eine lokale Eigenschaft ist, verhält sich vieles ebenso wie im \mathbb{R}^n . Wir sollten nur nicht den Fehler machen, dass wir irgendwelchen ‘natürlichen’ Koordinatensystemen einen über r hinausgehenden Glattheitsgrad zubilligen.

Sprechweise 1.4.2 (Diffeomorphismus). Eine bijektive r -glatte Abbildung einer C^r -Mannigfaltigkeit, deren Umkehrung ebenfalls r -glatt ist, heisst ein C^r -Diffeomorphismus. Zwei C^r -Mannigfaltigkeiten heissen r -diffeomorph oder ‘gleich als C^r -Mannigfaltigkeiten’, wenn ein r -Diffeomorphismus existiert.

Beispiel 1.4.3. Eine komplexe 2×2 -Matrix U heisst bekanntlich eine unitäre Matrix, wenn gilt $U^*U = I$ (Einheitsmatrix). Die Menge dieser Matrizen ist eine Gruppe, die man häufig mit $\mathcal{U}(2)$ bezeichnet. Eine interessante Untergruppe der unitären Gruppe $\mathcal{U}(2)$ ist die spezielle unitäre Gruppe $\mathcal{SU}(2)$. Diese ist andererseits eine uns bereits bekannte dreidimensionale reelle Mannigfaltigkeit. Die U haben nämlich die Gestalt

$$U = \begin{pmatrix} a & -b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{mit } a, b \in \mathbb{C}, \quad |a|^2 + |b|^2 = 1.$$

Wenn wir die Real- und Imaginärteile einführen, dann wird ersichtlich, dass wir es bei der Mannigfaltigkeit $\mathcal{SU}(2)$ um die in den \mathbb{R}^4 eingebettete 3-Sphäre S^3 handelt.

Die unitäre Gruppe $\mathcal{U}(2)$ selbst ist eine vierdimensionale Mannigfaltigkeit. Ihre Elemente gewinnt man aus den speziellen unitären Matrizen durch Multiplikation mit einer Zahl vom Betrag 1. Die Menge der komplexen Zahlen vom Betrag 1 ist die 1-Sphäre S^1 . $\mathcal{U}(2)$ ist als C^∞ -Mannigfaltigkeit gleich dem cartesischen Produkt $S^3 \times S^1$.

Beispiel 1.4.4. Eine weitere interessante Gruppe ist die Gruppe $\mathcal{SU}(1, 1)$. Sie besteht aus den komplexen 2×2 -Matrizen der Gestalt

$$V = \begin{pmatrix} a & b \\ \bar{b} & \bar{a} \end{pmatrix} \quad \text{mit } a, b \in \mathbb{C}, \quad |a|^2 - |b|^2 = 1.$$

Es handelt es sich um eine nichtkompakte dreidimensionale Mannigfaltigkeit. Real- und Imaginärteil von b unterliegen keinen Beschränkungen; der Absolutbetrag von a ist durch $|b|$ eindeutig bestimmt. $\mathcal{SU}(1, 1)$ ist C^∞ -diffeomorph zu $\mathbb{R}^2 \times S^1$.

Sprechweise. Es gibt nur zwei zusammenhängende eindimensionale Mannigfaltigkeiten, eine kompakte, das ist die Kreislinie und eine nichtkompakte, das ist die reelle Achse.

Eine glatte Abbildung einer zusammenhängenden eindimensionalen Mannigfaltigkeit, ($m = 1$), nennt man auch (informell!) eine glatte Kurve; eine glatte Abbildung einer Kreislinie nennt man in diesem Sinne eine geschlossene Kurve.

(Wir erinnern daran, dass wir bei früherer Gelegenheit glatte Kurven als Äquivalenzklassen glatt parametrisierter Kurven definiert haben; an dieser formellen Definition wollen wir festhalten. Wir erinnern auch daran, dass wir auch schon mit nichtglatten Kurven zu tun hatten; als es darum ging entlang eines Kurvenstücks eine ‘Form’ $\int h \, dk$ zu integrieren, haben wir nur gefordert, dass die Kurvenstücke stetige Abbildungen mit beschränkter Schwankung sind.)

Auch der Fall $n = 1$ gibt Anlass zu naheliegenden informellen Sprechweisen. Glatte Funktionen auf M sind bei uns formell(!) Elemente von \mathcal{D}_M . Das sind Objekte, die man miteinander multiplizieren kann, Elemente in einem kommutativen Ring mit Einselement.

Es ist aber auch üblich, eine glatte Funktion auf M als eine glatte Abbildung in die eindimensionale Mannigfaltigkeit \mathbb{R} zu bezeichnen; bei dieser Sprechweise denkt man dann eher an das Hintereinanderschaltens von Abbildungen. Entsprechend versteht man ein n -Tupel von glatten Funktionen auf M manchmal auch als eine glatte Abbildung von M in die Mannigfaltigkeit \mathbb{R}_{Sp}^n .

Glatte Kurvenstücke und Pushforward:

Ein stetig differenzierbares (oder ‘glattes’) Kurvenstück $\{\gamma(t) : t \sim 0\}$ auf einer \mathcal{C}^1 -Mannigfaltigkeit (M, \mathcal{D}_M) ist eine Abbildung einer Intervallumgebung des Nullpunkts in (M, \mathcal{D}_M) , sodass $k(\gamma(\cdot))$ stetig differenzierbar ist für alle $k \in \mathcal{D}$.

Wenn $\chi : (M, \mathcal{D}_M) \rightarrow (N, \mathcal{D}_N)$ eine glatte Abbildung ist, dann wird ein glattes Kurvenstück bei Q_0 in ein glattes Kurvenstück bei $P_0 = \chi(Q_0)$ abgebildet. Kurvenstücke mit demselben Tangentialvektor werden in Kurvenstücke mit demselben Tangentialvektor abgebildet. Wir haben also eine lineare Abbildung der Tangentialräume $\chi_* : T_{Q_0} \rightarrow T_{P_0}$. Diese nennt man die Pushforward-Abbildung im Punkt Q_0 .

Anmerkung: Man beachte, dass die Spur eines glatten Kurvenstücks i. Allg. keine Mannigfaltigkeit ist. Ein Beispiel, welches in vielen Lehrbüchern erwähnt wird, ist die Spur der Neil’schen Parabel. Die Neil’sche Parabel ist die glatt parametrisierte Kurve $\left\{ \begin{pmatrix} t^3 \\ t^2 \end{pmatrix} : |t| < 1 \right\}$. Die Spur ist die Punktmenge $\left\{ \begin{pmatrix} x \\ y \end{pmatrix} : -1 < x < 1, y = |x|^{2/3} \right\}$.

Wir wollen dabei bleiben, dass eine (glatte) Kurve eine Äquivalenzklasse von (glatt) parametrisierten Kurven ist, wobei die Parametermenge ein kompaktes Intervall ist, meistens das Einheitsintervall $[0, 1]$. Die Kurve darf keinesfalls mit ihrer Spur verwechselt werden. Wir werden später auch Äquivalenzklassen von glatten Abbildungen des Standardsimplexes (oder des Standardwürfels) benötigen. Diese ‘singulären Simplexe’ oder ‘singulären Würfel’ sind keine Mannigfaltigkeiten. Sie haben aber mehr Struktur als Punktmenge; beispielsweise hat ein k -dimensionales singuläre Simplex einen Rand, wobei der Rand eines singulären Simplexes etwas anderes ist als der topologische Rand der Spur dieses

singulären Simplexes. Der Rand ist eine ‘Summe’ von $k + 1$ singulären Simplexen der Dimension $k - 1$.

Die ‘Spitze’ in der Neil’schen Parabel ist offenbar der Tatsache geschuldet, dass der Tangentialvektor im Parameterpunkt 0 verschwindet. Die Spur einer injektiven glatten Abbildung sieht freundlicher aus, wenn der Pushforward in allen Punkten injektiv ist. Man spricht in diesem Fall von einer Immersion. Genauer über Immersionen findet man in tiefergehenden Lehrbüchern. Wir wollen hier schon einmal darauf hinweisen, dass die Spur einer Immersion nicht notwendigerweise eine Untermannigfaltigkeit ist im Sinne der unten entwickelten Begriffsbestimmung.

Pfaff’sche Formen

Die Cotangentialvektoren in einem festen Punkt P der n -dimensionalen Mannigfaltigkeit (N, \mathcal{D}_N) bilden einen n -dimensionalen Vektorraum. Wenn $\{x^j : j \in J\}$ ein lokales Koordinatensystem bei P ist, dann sind die Anstiege $dx^j|_P$ eine Basis. Wir wollen uns jetzt mit Cotangentialfeldern befassen; ein Cotangentialfeld ordnet jedem $P \in N$ ein Cotangentialvektor mit dem Fußpunkt P zu. Die Zuordnung soll zumindest stetig sein; wir müssen aber immer wieder einmal genauer sein.

Definition 1.9. Es sei (N, \mathcal{D}_N) eine C^r -Mannigfaltigkeit. Ein Covektorfeld von der Gestalt $\omega = \sum h_m \cdot dk_m$ heisst eine Pfaff’sche Form, wenn die k_m r -mal und die h_m $(r - 1)$ -mal stetig differenzierbar sind.

Die Pfaff’schen Formen heissen auch Differentialformen der Ordnung 1 oder 1-Formen.

Satz 1.4.3. Ist $\{x^j : j \in J\}$ ein lokales Koordinatensystem auf U , so besitzt für jede Pfaff’schen Form die Einschränkung eine Darstellung

$$\omega = \sum g_j \cdot dx^j \quad \text{mit eindeutig bestimmten } g_j. \quad (1)$$

Die ‘Koeffizientenfunktionen’ g_j sind $(r - 1)$ -mal stetig differenzierbar.

Beweis. Zunächst eine Notation: Die partielle Ableitung nach der j -ten Koordinate (bei festgehaltenen übrigen Koordinaten) entspricht einem Tangentialvektorfeld D_j über U , welches man gelegentlich mit $\frac{\partial}{\partial x^j}$ bezeichnet. Der Tangentialvektor, der dem Punkt P zugeordnet wird, wird mit $\frac{\partial}{\partial x^j}|_P$ bezeichnet. Wir identifizieren die Pfaff’schen Formen durch die Auswertung in diesen Tangentialvektoren: Für $\omega = dk$ mit $k = K(x^1, \dots, x^n)$ ist

$$\left\langle dk, \frac{\partial}{\partial x^j}|_P \right\rangle = \frac{\partial}{\partial x^j} K(x^1(P), \dots, x^n(P)) = D_j k(P)$$

die j -te partielle Ableitung von K im Punkt P . Für jedes $\omega = h \cdot dk$ ist die Auswertung als Funktion von P

$$\left\langle \omega, \frac{\partial}{\partial x^j}|_P \right\rangle = \left\langle h \cdot dk, \frac{\partial}{\partial x^j}|_P \right\rangle = h(P) \cdot D_j(k)(P)$$

$(r - 1)$ -mal stetig differenzierbar. Wir haben speziell

$$\left\langle dx^m, \frac{\partial}{\partial x^j} \Big|_P \right\rangle = 0 \quad \text{für } m \neq j \quad \text{und } = 1 \text{ für } m = j$$

Für jede Pfaff'sche Form ist $\langle \omega, \frac{\partial}{\partial x^j} \rangle$ eine $(r - 1)$ -mal stetig differenzierbare Funktion.

Wenn ω irgendeine Pfaff'sche Form ist und $\langle \omega, \frac{\partial}{\partial x^j} \Big|_P \rangle = g_j(P)$ für alle $P \in U$, dann gilt $\omega = \sum g_j \cdot dx^j$.

Definition 1.10 (Geschlossene 1-Formen). Eine 1-Form ω heisst geschlossen, wenn es zu jedem P eine Umgebung U und eine Funktion g auf U gibt, sodass $dg = \omega$ auf U .

Satz 1.4.4. Ist ω geschlossen auf dem Bereich U der Koordinatisierung $\{x^j : j \in J\}$

$$\omega = \sum g_j \cdot dx^j \quad \text{mit } g_j \text{ stetig differenzierbar,}$$

so gilt

$$\frac{\partial}{\partial x^j} g_k = \frac{\partial}{\partial x^k} g_j \quad \text{für alle } j, k \in J.$$

Beweis. Nach dem Satz von H. A. Schwarz kommt es bei den partiellen Differentiationen nicht auf die Reihenfolge an. Wir werden später, wenn wir das berühmte Lemma von Poincaré behandeln, sehen, dass die Bedingung an die Ableitungen auch hinreichend ist für die Geschlossenheit der Form.

Beispiel. Die folgenden Formen auf dem gepunkteten \mathbb{R}^2 sind geschlossen:

$$\omega = \frac{1}{x^2 + y^2} (x dx + y dy), \quad \omega = \frac{1}{x^2 + y^2} (-y dx + x dy).$$

Sie sind der Real- und der Imaginärteil der komplexen Form

$$d(\ln z) = \frac{1}{z} dz = \frac{x - iy}{z\bar{z}} (dx + idy).$$

Definition 1.11 (Pullback für 1-Formen).

Ist $\varphi : (M, \mathcal{D}_M) \rightarrow (N, \mathcal{D}_N)$ eine r -glatte Abbildung, und ω eine Pfaff'sche Form auf dem Zielraum N , dann bezeichnet $\varphi^*(\omega)$ oder $\varphi^*\omega$ die zurückgenommene Pfaff'sche Form.

$$\langle \varphi^*\omega, v_P \rangle = \langle \omega, \varphi_*(v_P) \rangle \quad \text{für alle } v_P \in T_P.$$

Bemerke: Wenn $\omega = h \cdot dk$ und $\bar{h} = \varphi^*(h)$, $\bar{k} = \varphi^*(k)$ die zurückgenommenen Funktionen sind, dann gilt $\varphi^*\omega = \bar{h} \cdot d\bar{k}$.

Wir müssen noch etwas zur algebraischen Struktur sagen: Die Algebraiker kennen den Begriff eines A -Moduls, wo A ein kommutativer Ring mit Einselement ist. In einem A -Modul kann man nicht nur so wie in einem Vektorraum linear kombinieren, man auch mit den Elementen von A multiplizieren.

Definition 1.12. Es sei $(A, +, \cdot)$ ein kommutativer Ring mit 1. Eine additive Gruppe $(M, +)$ hat die Struktur eines A -Moduls, wenn eine Operation definiert ist

$$A \times M \longrightarrow M; \quad (\alpha, m) \longmapsto \alpha \cdot m$$

mit den Eigenschaften

$$\begin{aligned} \text{i)} \quad & \alpha(x + y) = \alpha x + \alpha y \\ \text{ii)} \quad & (\alpha + \beta) \cdot x = \alpha x + \beta x \\ \text{iii)} \quad & \alpha(\beta \cdot x) = (\alpha\beta) \cdot x \\ \text{iv)} \quad & 1 \cdot x = x . \end{aligned}$$

Notation.

Die Menge der glatten Pfaff'schen Formen auf einer \mathcal{C}^r -Mannigfaltigkeit hat die Struktur eines A -Moduls, wenn A der Ring der $(r-1)$ -mal stetig differenzierbaren Funktionen ist. Die übliche Bezeichnung für diesen Modul ist Ω_N^1 .

Wir halten fest: Für eine r -glatte Abbildung $\varphi : (M, \mathcal{D}_M) \longrightarrow (N, \mathcal{D}_N)$ ist der Pullback für 1-Formen ein A -Modul-Homomorphismus $\varphi^* : \Omega_M^1 \longleftarrow \Omega_N^1$.

Hinweis: Wir werden später auch noch die Moduln der k -Formen Ω_N^k zu betrachten haben ($k = 0, 1, \dots, n$) sowie ihre direkte Summe $\bigoplus_{k=0}^n \Omega_N^k$. Der Pullback ist auch für die Moduln der k -Formen ein Modul-Homomorphismus.

Über die 1-Formen soll hier schon einmal gesagt sein, dass es ihre Bestimmung ist, über glatte Kurvenstücke integriert zu werden, so wie es die Bestimmung eines einzelnen Kovektors ist, mit einem Tangentialvektor im betreffenden Fußpunkt evaluiert zu werden. Das Integral einer zurückgenommenen 1-Form $\varphi^*\omega$ über eine Kurve \mathcal{C} ist das Integral von ω über die Bildkurve. Die Integration von Formen ist hier aber noch nicht unser Thema; wir bleiben hier noch bei der algebraischen Konstruktionen um den Begriff der Mannigfaltigkeit.

Konstruktion einer Mannigfaltigkeitsstruktur durch einen Atlas:

Wenn man einen HRaB N mit einer Mannigfaltigkeitsstruktur \mathcal{D}_N ausstatten will, dann geschieht das meistens dadurch, dass man einen Atlas angibt. Man spezifiziert eine Familie $\{\mathbf{U}_\alpha, \varphi_\alpha\}$, wo die \mathbf{U}_α eine offene Überdeckung von N bilden, und jedes φ_α eine bijektive Abbildung von \mathbf{U}_α auf eine offene Teilmenge des \mathbb{R}^n .

Diese Daten definieren genau dann die Struktur einer \mathcal{C}^r -Mannigfaltigkeit, wenn für alle α, β mit $\mathbf{U}_\alpha \cap \mathbf{U}_\beta \neq \emptyset$ gilt: Für die Einschränkungen von φ_α und φ_β auf diesen Durchschnitt ist $\varphi_\beta \circ \varphi_\alpha^{-1}(\cdot)$ r -mal stetig differenzierbar.

Eine (global auf N definierte) Funktion f soll nun zu \mathcal{D}_N gehören, wenn für jede Koordinatenumgebung \mathbf{U}_α eine r -mal stetig differenzierbare Funktion $F_\alpha(\cdot)$ auf dem Bildbereich $\varphi_\alpha(\mathbf{U}_\alpha)$ existiert, sodass $f(\cdot) = F_\alpha(\varphi_\alpha(\cdot))$.

Beispiel 1.4.5 (Produkt-Mannigfaltigkeit).

Es seien (M_1, \mathcal{D}_1) und (M_2, \mathcal{D}_2) \mathcal{C}^r -Mannigfaltigkeiten. Das cartesische Produkt $M_1 \times M_2$ macht man zu einer $(m_1 + m_2)$ -dimensionalen \mathcal{C}^r -Mannigfaltigkeit, indem man festlegt:

Wenn $\mathbf{x}_1 = \{\mathbf{x}^j : j \in J_1\}$ eine Karte bei $\tilde{Q}_1 \in M_1$ ist und $\mathbf{x}_2 = \{\mathbf{x}^j : j \in J_2\}$ eine Karte bei $\tilde{Q}_2 \in M_2$, dann soll $\binom{\mathbf{x}_1}{\mathbf{x}_2} = \{\mathbf{x}^j : j \in J_1 + J_2\}$ eine Karte bei $(\tilde{Q}_1, \tilde{Q}_2)$ sein. Die bei $(\tilde{Q}_1, \tilde{Q}_2)$ glatten Funktionen sind also in der Nähe dieses Punkts von der Gestalt $f(P_1, P_2) = F\left(\binom{\mathbf{x}_1(P_1)}{\mathbf{x}_2(P_2)}\right)$ mit einer r -mal stetig differenzierbaren Funktion F auf dem $\mathbb{R}^{m_1+m_2}$.

Beispiel 1.4.6 (Das Tangentialbündel einer Mannigfaltigkeit).

Es sei (M, \mathcal{D}) eine C^∞ -Mannigfaltigkeit. E sei die disjunkte Vereinigung aller Tangentialräume $E = \bigcup_{P \in M} T_P$. (Mit $\pi : E \rightarrow M$ bezeichnet man üblicherweise die Fußpunktabbildung.) Die topologische Struktur des ‘Tangentialbündels’ E liegt auf der Hand. Man macht E nun aber auch zu einer Mannigfaltigkeit: Ist $(U, \{\mathbf{x}^j : j \in J\})$ ein lokales Koordinatensystem für M , so bringt dieses mit sich eine lineare Koordinatisierung der Tangentialräume mit Fußpunkt in U . Wir haben also eine Karte für $\bigcup_{P \in U} T_P$. So gewinnen wir einen Atlas für das Tangentialbündel. (Die Wechsel zwischen den Karten brauchen die Jacobi-Matrizen.) Die so koordinatisierten ‘Punkte’ des Tangentialbündels E werden häufig mit $(\mathbf{q}^j, \dot{\mathbf{q}}^j)$ bezeichnet; \mathbf{q}^j ist die j -te Koordinate des Fußpunkts. Wir haben gesehen, dass man einen Tangentialvektor dadurch charakterisieren kann, dass man angibt, wie schnell sich die Koordinaten entlang des Tangentialvektors ändern; die Tangentialvektoren werden durch die J -Spalte $\dot{\mathbf{q}}^j$ beschrieben, wenn für die Mannigfaltigkeit der Fußpunkte die lokale Koordinatisierung $\{\mathbf{q}^j : j \in J\}$ gewählt wurde.

Untermannigfaltigkeiten:

Wir kommen jetzt zum Begriff der Untermannigfaltigkeit, einem Begriff, der etwas heikler ist als der (in der Linearen Algebra studierte) Begriff des affinen Teilraums. Wenn wir an die affinen Mannigfaltigkeiten denken, dann können wir feststellen:

1. Das affine Bild einer affinen Mannigfaltigkeit ist eine affine Teilmannigfaltigkeit des Zielraums
2. Das volle Urbild eines Punkts im Zielraum ist ein affiner Teilraum des Urbildraums. Anders gesagt: die Lösungsmenge eines inhomogenen linearen Gleichungssystems ist ein affiner Teilraum des Urbildraums.

Bei den Mannigfaltigkeiten sind die Verhältnisse komplizierter. Dabei folgt die Definition des Begriffs einer Untermannigfaltigkeit einem universellen Muster:

Definition 1.13 (Untermannigfaltigkeit).

Es sei (N, \mathcal{D}_N) eine C^r -Mannigfaltigkeit. Eine Teilmenge $K \subseteq N$ heisst eine k -dimensionale Untermannigfaltigkeit, wenn das System aller glatten f , eingeschränkt auf K die C^r -Struktur einer k -dimensionalen Mannigfaltigkeit definiert.

Beispiel 1.4.7 (Funktionsgraphen als Untermannigfaltigkeiten des Produktraums).

Es sei $F(\mathbf{x}^{m+1}, \dots, \mathbf{x}^n)$ eine r -mal stetig differenzierbare Abbildung in den Raum \mathbb{R}_{Sp}^m der reellen m -Spalten \mathbf{x}^l , definiert auf einer offenen Teilmenge U^l des Raums aller $(n - m)$ -Spalten \mathbf{x}^l : $F(\cdot) : U^l \ni \mathbf{x}^l \mapsto \mathbf{x}^l = F(\mathbf{x}^l)$.

Der Graph M dieser Abbildung ist dann eine $(n - m)$ -dimensionale C^r -Mannigfaltigkeit, und zwar eine Untermannigfaltigkeit der Spaltenmenge $N = \mathbb{R}_{\text{Sp}}^m \times U^{\parallel}$, (aufgefasst als eine C^r -Mannigfaltigkeit.)

$$M = \left\{ \begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix} : \mathbf{x}^{\parallel} \in U^{\parallel}, \quad \mathbf{x}^{\perp} = F(\mathbf{x}^{\parallel}) \right\}.$$

Wenn h eine glatte Funktion auf U ist, $h = H(x^1, \dots, x^m, x^{m+1}, \dots, x^n) = H(\begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix})$, dann ist die Einschränkung auf den Graphen gleich der Einschränkung der Funktion $H(\begin{pmatrix} F(\mathbf{x}^{\parallel}) \\ \mathbf{x}^{\parallel} \end{pmatrix})$. Sie lässt sich also als r -glatte Funktion des $(n - m)$ -Tupels \mathbf{x}^{\parallel} darstellen. Andererseits liefert jede r -glatte Funktion des $(n - m)$ -Tupels \mathbf{x}^{\parallel} auf U^{\parallel} eine Funktion auf M , die man als die Einschränkung einer glatten Funktion auf N auffassen kann. Somit ist \mathbf{x}^{\parallel} auf M eine C^r -Koordinatisierung. Man bemerke: Nicht nur $\begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix}$ ist eine Koordinatisierung der Mannigfaltigkeit $U = \mathbb{R}^m \cdot U^{\parallel}$, sondern auch $\begin{pmatrix} \mathbf{x}^{\perp} - F(\mathbf{x}^{\parallel}) \\ \mathbf{x}^{\parallel} \end{pmatrix}$. Bei dieser alternativen Koordinatisierung sind die Punkte der Untermannigfaltigkeit M dadurch gekennzeichnet, dass die ersten m -Koordinaten verschwinden.

Notation. Wenn $\mathbf{x} = \{x^j : j \in J\}$ eine J -Spalte ist, und $J = J^{\perp} + J^{\parallel}$ eine Partition der Indexmenge, dann notieren wir gelegentlich $\mathbf{x} = \begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix}$.

(Übrigens: Wenn $\mathbf{w} = \{w_j : j \in J\}$ eine J -Zeile ist, dann notieren wir $\mathbf{w} = (\mathbf{w}^{\perp}, \mathbf{w}^{\parallel})$. Diese Situation finden wir z. B., wenn wir aus Basen zweier Vektorräume W_1 und W_2 eine Basis der direkten Summe $W_1 \oplus W_2$ zusammenstellen.)

Die Situation, die wir bei den Funktionsgraphen analysiert haben, ist typisch für Untermannigfaltigkeiten.

Satz 1.4.5. *Es sei M eine Teilmenge der n -dimensionalen C^r -Mannigfaltigkeit (N, \mathcal{D}_N) . M ist genau dann eine $(n - m)$ -dimensionale Untermannigfaltigkeit, wenn für jedes $Q_0 \in M$ eine Karte (U, φ_U) für die große Mannigfaltigkeit N existiert,*

$$\varphi_U = \{x^1, \dots, x^m, x^{m+1}, \dots, x^n\}, \quad \text{sodass} \quad \varphi(U \cap M) = \varphi(U) \cap (\{0\}^m \times \mathbb{R}^{n-m}).$$

Beweis. *Es sei $M \subseteq N$ eine $(n - m)$ -dimensionale Untermannigfaltigkeit, $Q_0 \in M$ und $\mathbf{x}^{\parallel} = \{x^j : j \in J^{\parallel}\}$ ein $(n - m)$ -Tupel von Funktionen, welches leistet, was man für die Struktur einer Untermannigfaltigkeit benötigt: Zu jedem glatten f existiert eine Funktion F von $n - m$ Variablen sodass $f = F(\mathbf{x}^{\parallel})$ auf $M \cap U$.*

Wir ergänzen das $(n - m)$ -Tupel durch ein m -Tupel $\mathbf{x}^{\perp} = \{x^j : j \in J^{\perp}\}$ sodass $\begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix}$ eine Karte bei Q_0 ist. Es ist dazu nur zu fordern, dass die Differentiale $d\mathbf{x}^j|_{Q_0}$ mit $j \in J^{\perp}$ die gegebenen linear unabhängigen $d\mathbf{x}^j|_{Q_0}$ mit $j \in J^{\parallel}$ zu einer Basis des Cotangentialraums ergänzen. Sei nun $F = \{F^j : j \in J^{\perp}\}$ das Tupel der Funktionen zu den ergänzenden x^j . Es ist ein zulässiger Koordinatenwechsel, wenn wir von $\begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix}$ zu $\begin{pmatrix} \mathbf{x}^{\perp} - F(\mathbf{x}^{\parallel}) \\ \mathbf{x}^{\parallel} \end{pmatrix}$ übergehen. In diesem neuen Koordinatensystem haben wir (für eine Umgebung U von Q_0)

$$U \cap M = U \cap \left\{ \begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix} : \mathbf{x}^{\perp} - F(\mathbf{x}^{\parallel}) = 0 \right\} = \left\{ \begin{pmatrix} \mathbf{x}^{\perp} \\ \mathbf{x}^{\parallel} \end{pmatrix} : \mathbf{x}^{\perp} = F(\mathbf{x}^{\parallel}) \right\}.$$

Also ist M lokal der Graph der Abbildung $x^{\parallel} \mapsto F(x^{\parallel})$. Die umgekehrte Richtung haben wir bereits abgehandelt: Teilmengen der Mannigfaltigkeit, die lokal wie Graphen aussehen, sind Teilmannigfaltigkeiten.

Wir haben bereits erwähnt, dass das Bild unter einer Immersion nicht notwendigerweise eine Untermannigfaltigkeit ist. Der folgende Satz, den wir nicht beweisen wollen, formuliert eine hinreichende Zusatzbedingung.

Satz 1.4.6. *Es sei $\chi : (M, \mathcal{D}_M) \longrightarrow (N, \mathcal{D}_N)$ eine injektive glatte Abbildung mit der Eigenschaft, dass der Pushforward in jedem $Q \in M$ injektiv ist. Wenn es zu jedem Bildpunkt $P_0 = \chi(Q_0)$ eine Umgebung U gibt, sodass $\chi(M) \cap U$ in U abgeschlossen ist, dann ist die Bildmenge $\chi(M)$ eine Untermannigfaltigkeit.*

Bemerkung: Die Zusatzbedingung garantiert, grob gesprochen, dass es in keinem $P_0 = \chi(Q_0)$ passiert, dass in jeder Umgebung Bildpunkte gibt, die von Punkten Q weit weg von Q_0 herkommen. Wir können U so klein machen, dass jede Punktfolge in $\chi(M) \cap U$, die gegen einen Punkt in U konvergiert, zum Bild $\chi(M)$ gehört.

Sprechweise (Niveaufläche). Es sei f eine Funktion auf der Mannigfaltigkeit (N, \mathcal{D}_N) . Die Menge $\{P : f(P) = \mathbf{b}\}$ wird die Niveaufläche zum Niveau \mathbf{b} genannt.

Wenn $f = \{f^i : i \in I\}$ ein m -Tupel von Funktionen ist und \mathbf{b} eine I -Spalte, dann kann man das Lösungsgebilde des Gleichungssystems $L = \{P : f(P) = \mathbf{b}\}$ als die Schnittmenge der Niveauflächen verstehen. Wenn man f als eine Abbildung in den $\mathbb{R}_{\mathbb{S}_p}^m$ auffasst, dann erscheint L als das volle Urbild der einpunktigen Menge $\{\mathbf{b}\}$. Manchmal nennt man dieses volle Urbild $f^{-1}(\{\mathbf{b}\})$ auch die Faser der Abbildung f über dem Zielpunkt \mathbf{b} .

Wenn in einem ‘guten’ Punkt $P_0 \in L$ das m -Tupel der Differentiale $df^i|_{P_0}$ linear unabhängig ist, dann bilden die Tangentialvektoren, welche alle diese Covektoren annullieren einen $(n - m)$ -dimensionalen Teilraum von T_{P_0} . Dies ist der Nullraum für den Pushforward in P_0 zur Abbildung $P \mapsto \{f^i(P) : i \in I\} \in \mathbb{R}^I$. Der Pushforward hat also im ‘guten’ Punkt P_0 vollen Rang m .

Wir formulieren diese Idee vom ‘guten’ Punkt auf dem Schnittgebilde von Niveauflächen in der Sprache der Abbildungen von Mannigfaltigkeiten.

Sprechweise 1.4.3. Gegeben sei eine glatte Abbildung einer n -dimensionalen Mannigfaltigkeit in eine m -dimensionalen Mannigfaltigkeit mit $m < n$.

$$\psi : (N, \mathcal{D}_N) \longrightarrow (M, \mathcal{D}_M)$$

Man nennt einen Punkt $P_0 \in N$ einen regulären Urbildpunkt für ψ , wenn der Pushforward in P_0 surjektiv ist. Man nennt $Q \in M$ einen regulären Bildpunkt, wenn alle $P \in \psi^{-1}(Q_0)$ reguläre Punkte sind.

Satz 1.4.7. *Es sei Q_0 ein regulärer Bildpunkt für die glatte Abbildung*

$$\psi : (N, \mathcal{D}_N) \longrightarrow (M, \mathcal{D}_M), \quad n = \dim N < \dim M = m$$

Dann ist $L = \{P : \psi(P) = Q_0\}$ eine Untermannigfaltigkeit mit $\dim L = n - m$.

Der Beweis beruht auf dem Satz von der implizit gegebenen Funktion, einer wichtigen Variante des Satzes von der glatten Umkehrabbildung, den wir hier noch einmal kurz rekapitulieren wollen.

Satz 1.4.8 (Satz von der implizit gegebenen Funktion).

Der Satz handelt von stetig differenzierbaren Funktionen $F(x)$ auf einer Umgebung U eines Punktes $x_0 \in \mathbb{R}^n$ mit Werten im \mathbb{R}^m , $F(x_0) = y_0$, wo die Jacobimatrix $F'(x_0)$ den vollen Rang m besitzt.

a) Der Fall $m = n$: Es existiert eine Umgebung V von y_0 und dort eine Abbildung

$$H(y), \quad \text{sodass} \quad F(H(y)) = y \quad \text{für alle } y \in V.$$

Man kann die Umgebung $U \ni x_0$ so klein machen, dass das F -Bild in V enthalten ist und $H(F(x)) = x$ für alle $x \in U$.

b) Der Fall $m < n$: Das n -Tupel der Koordinaten sei so zerlegt: $x = \begin{pmatrix} x^l \\ x^{\parallel} \end{pmatrix}$, dass

$$F'(x_0) = \left(\frac{\partial F}{\partial x^l}, \frac{\partial F}{\partial x^{\parallel}} \right) (x_0) \quad \text{mit} \quad \frac{\partial F}{\partial x^l}(x_0) \quad \text{invertierbar.}$$

Es existieren Umgebungen W von y_0 und U^{\parallel} von x_0^{\parallel} , sowie eine Abbildung

$$H: \quad W \times U^{\parallel} \ni \begin{pmatrix} y \\ x^{\parallel} \end{pmatrix} \longmapsto H \begin{pmatrix} y \\ x^{\parallel} \end{pmatrix} \in U, \quad \text{sodass} \quad F \left(H \begin{pmatrix} y \\ x^{\parallel} \end{pmatrix} \right) = y,$$

Man kann $W \times U^{\parallel}$ so klein wählen, dass $H \begin{pmatrix} F(x) \\ x^{\parallel} \end{pmatrix} = \begin{pmatrix} x^l \\ x^{\parallel} \end{pmatrix}$ für alle $x \in U^l \times U^{\parallel}$.

c) Kurz gesagt: Es existiert eine Umgebung U^l von x_0^l , sodass

$$U^l \times W \times U^{\parallel} \cap \left\{ \begin{pmatrix} x^l \\ y \\ x^{\parallel} \end{pmatrix} : F \begin{pmatrix} x^l \\ x^{\parallel} \end{pmatrix} = y \right\} = U^l \times W \times U^{\parallel} \cap \left\{ \begin{pmatrix} x^l \\ y \\ x^{\parallel} \end{pmatrix} : H \begin{pmatrix} y \\ x^{\parallel} \end{pmatrix} = \begin{pmatrix} x^l \\ x^{\parallel} \end{pmatrix} \right\}$$

Beweis. Die Aussage a) ist der Satz von der glatten Umkehrabbildung, den wir mit Hilfe des Banach'schen Fixpunktsatzes bewiesen haben. Die Teile b) und c) ergeben sich, wenn wir das gegebene Funktionensystem $F = \{F^i : i \in I\}$ in 'trivialer' Weise durch das J^{\parallel} -Tupel der späteren Koordinatenfunktionen x^{\parallel} ergänzen. Wir erhalten so nämlich ein n -Tupel von Funktionen mit einer invertierbaren Jacobi-Matrix.

Der Satz beweist den Satz über das volle Urbild eines regulären Bildpunkts:

Satz 1.4.9. Ist Q_0 ein regulärer Bildpunkt für die Abbildung $\psi : (N, \mathcal{D}_N) \longrightarrow (M, \mathcal{D}_M)$, so ist sein volles Urbild $\{P : \psi(P) = Q_0\}$ eine Untermannigfaltigkeit.

Beweis. Das Urbild ist lokal der Graph einer glatten Abbildung $x^{\parallel} \mapsto x^{\perp} = F(x^{\parallel})$. In der Regel haben wir sogar eine Schar von Untermannigfaltigkeiten, dann nämlich, wenn die Bildpunkte Q in einer Umgebung von Q_0 regulär sind.

Beispiel 1.4.8. Der Kreis vom Radius R im euklidischen \mathbb{R}^2 wird schon in der Schule beschrieben als der ‘geometrische Ort’ aller Punkte P mit $x^2(P) + y^2(P) = R^2$.

Alternativ kann man diese Kreis K_R auch als die Spur einer glatt parametrisierten Kurve gewinnen. Eine naheliegende parametrisierte Kurve wäre z. B.

$$K_R = \left\{ \begin{pmatrix} R \cdot \cos t \\ R \cdot \sin t \end{pmatrix} : t \in \mathbb{R}/2\pi \right\}.$$

Bei der Beschreibung einer Mannigfaltigkeit darf man nur selten eine globale Beschreibung erwarten; es geht in der Regel um lokale Beschreibungen mit Atlanten. Bei unserem Kreis können wir in der Nähe der Punkte $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} \pm R \\ 0 \end{pmatrix}$ die Koordinate $x(\cdot)$ verwenden; in der Nähe der Punkte $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ \pm R \end{pmatrix}$ können wir die Koordinate $y(\cdot)$ verwenden.

$$K_R = \left\{ \begin{pmatrix} x \\ y(x) \end{pmatrix} : y(x) = \pm \sqrt{R^2 - x^2} \right\} \quad \text{bzw.} \quad = \left\{ \begin{pmatrix} x(y) \\ y \end{pmatrix} : x(y) = \pm \sqrt{R^2 - y^2} \right\}.$$

1.5 Umformungen für spezielle Differentialgleichungen

Man sagt: Eine Differentialgleichung zu lösen, bedeutet, eine unbekannte Funktion zu bestimmen aus einer Gleichung, in welcher neben der unbekanntem Funktion auch ihre Ableitungen vorkommen. — In einem avanciertem Verständnis kann man eine so ungenaue Auskunft nicht akzeptieren; es muss genauer gesagt werden, was man (im jeweiligen Kontext) mit dem ‘Bestimmen’ einer Funktion (oder einer Funktionenklasse) meint; nur in exceptionellen Fällen kann gemeint sein, dass man eine ‘explizite’ Formel für die gesuchte Funktion findet. — Im gegenwärtigen Abschnitt werden wir uns mit (mehr oder weniger) speziellen Differentialgleichungen befassen, bei welchen der Lösungsbegriff zunächst einmal einigermaßen oberflächlich behandelt werden kann.

Sprechweise 1.5.1. Eine gewöhnliche Differentialgleichung n -ter Ordnung für die skalare Funktion $x(t)$ ist eine Gleichung der Form

$$G(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t), x^{(n)}(t)) = 0 \quad \text{mit einem glatten } G(\cdot).$$

Man spricht auch von einer Gleichung in impliziter Form. Wenn die partielle Ableitung nach der letzten Variablen nicht verschwindet, dann kann man zu einer äquivalenten Gleichung in expliziter Form übergehen:

$$x^{(n)}(t) = g(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n-1)}(t)) \quad (2)$$

Mit einem Standardverfahren kann man eine solche Gleichung in ein System von Differentialgleichungen erster Ordnung verwandeln. Man nennt das System auch eine Differentialgleichung für eine vektorwertige Funktion.

Sprechweise 1.5.2. Eine gewöhnliche Differentialgleichung erster Ordnung für die ‘vektorwertige’ Funktion $\mathbf{x}(t) = \{x^j(t) : j \in J\}$ ist ein J -Tupel von Gleichungen der Form

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{V}(t, \mathbf{x}(t)) \quad \text{mit einem stetigen } \mathbf{V}(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}_{\text{Sp}}^J \longrightarrow \mathbb{R}_{\text{Sp}}^J. \quad (3)$$

Bemerke: Im eindimensionalen Fall hat eine solche Gleichung die Gestalt $\frac{dx}{dt} = V(t, x(t))$. Wenn man, wie in Schulbüchern üblich, von vornherein x als unabhängige und y als abhängige Variable versteht, dann schreibt man auch einfach $y' = V(x, y)$.

Wir müssen später genauer sein mit den Definitionsbereichen und den Zielbereichen der ‘Abbildungen’ $V(\cdot, \cdot)$. Wir wollen hier aber schon einmal festhalten: $V(t, \cdot)$ ist in Wirklichkeit für jedes t ein stetiges Tangentialvektorfeld auf einer offenen Teilmenge einer Mannigfaltigkeit (N, \mathcal{D}) ; $V(t, \cdot)$ ordnet jedem Punkt P einen Tangentialvektor mit dem ‘Fußpunkt’ P zu. Dies geschieht in stetiger Weise; eine schärfere Bedingung, die lokale Lipschitz-Stetigkeit wird uns im nächsten Abschnitt beschäftigen. Es sollte klar sein: Wenn man auf der Mannigfaltigkeit lokale Koordinaten einführt $\mathbf{x}(t) = \{x^j(t) : j \in J\}$, und auf den Tangentialräumen das dazugehörige Koordinatensystem, dann kommt man gerade auf die oben beschriebene Situation.

Beispiel 1.5.1. In der Mechanik stellt man sich eine Masse m vor, welche an einer Feder hängt, deren in die Nulllage rücktreibende Kraft proportional zur Auslenkung x ist. Zusätzlich lässt man zu, dass die Schwingung durch eine zur Geschwindigkeit \dot{x} proportionalen Reibungskraft gebremst wird. Das Newton'sche Gesetz *Kraft = Masse · Beschleunigung* führt dann zur Gleichung des harmonischen Oszillators

$$m \cdot \ddot{x} + r \cdot \dot{x} + k \cdot x = f(t) \quad \text{zur 'anregenden Kraft' } f(\cdot).$$

Dieselbe Gleichung passt auf den einfachen elektrischen Schwingkreis. Zu einem solchen gehört eine Spule mit der Induktivität L , ein Kondensator mit der Kapazität C und ein Ohm'scher Widerstand R . Wenn eine 'elektromotorische Kraft' $U(t)$ den Schwingkreis antreibt, dann erfüllt die Ladung auf dem Kondensator $Q(t)$ die Differentialgleichung

$$L \cdot \ddot{Q} + R \cdot \dot{Q} + \frac{1}{C} \cdot Q = U(t)$$

(gemäß dem Gesetz, dass der Spannungsabfall, über die Schaltelemente summiert, der angelegten Spannung die Waage hält.)

Bei dieser Gleichung handelt es sich um eine sog. inhomogene lineare Differentialgleichung (mit konstanten Koeffizienten). Der Lösungsraum ist ein zweidimensionaler affiner Raum. Die Gleichung zur äusseren Kraft $= 0$, heisst die dazugehörige homogene Gleichung; ihre Lösungsmenge ist ein Vektorraum. Die allgemeine Lösung der inhomogenen Gleichung erhält man, wenn man zu einer speziellen Lösung der inhomogenen Gleichung alle Lösungen der homogenen Gleichung dazuaddiert.

Die Gleichung kann man leicht in eine 'vektorwertige' Differentialgleichung verwandeln, indem man $\dot{x}(t)$ oder $p(t) = m \cdot \dot{x}(t)$ als zweite unbekannt Funktion neben $x(t)$ betrachtet. Wegen $\frac{d}{dt}p(t) = -k \cdot x(t) - r \cdot \dot{x}(t) + f(t)$ erhalten wir

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = \begin{pmatrix} 0 & 1/m \\ -k & -r/m \end{pmatrix} \cdot \begin{pmatrix} x(t) \\ p(t) \end{pmatrix} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix}. \quad (4)$$

Es ist leicht, eine Basis für die Lösungsmenge der homogenen Gleichung anzugeben, und auch (mit der Methode der Variation der Konstanten) eine partikuläre Lösung der inhomogenen Gleichung. — Dazu kommen wir aber später in allgemeinerem Zusammenhang.

Beispiel 1.5.2. Wir betrachten einen schwingenden Massenpunkt, welcher von einer Kraft in die Ausgangslage zurückgetrieben wird, die von einem Potential $U(x)$ herrührt.

$$\ddot{x}(t) = -U'(x(t)) \quad \text{oder} \quad \frac{d}{dt} \left(\frac{1}{2} \dot{x}^2(t) \right) = -U'(x(t)) \cdot \dot{x}(t) = -\frac{d}{dt} U(x(t)). \quad (5)$$

Es existiert eine Konstante E , sodass $\frac{1}{2} \dot{x}^2(t) = E - U(x(t))$. In Bereichen, in welchen die gesuchte Funktion monoton (steigend oder fallend) ist, erhalten wir eine skalare Differentialgleichung erster Ordnung $\frac{d}{dt}x(t) = \pm \sqrt{2(E - U(x(t)))}$.

Man schreibt kurz

$$\frac{dx}{dt} = \pm \sqrt{2(E - U(x))} \quad \text{oder auch} \quad \frac{dt}{dx} = \pm \frac{1}{\sqrt{2(E - U(x))}}.$$

Diese Wendung der Verhältnisse soll im Folgenden erörtert werden.

Wir haben die Differentialgleichungen in unseren Beispielen ‘umgeschrieben’. Wir müssen jetzt erörtern, was wir unter einer Lösung der Gleichung verstehen wollen, und wir müssen klären, inwiefern das Umschreiben der Gleichungen den Lösungs begriff respektiert.

Definition 1.14. Eine glatt parametrisierte Kurve $\{q(t) : t \in [a, b]\}$ heisst ein Lösungsstück für die Differentialgleichung $\frac{d}{dt}q(t) = V(t, q(t))$, wenn für alle in einer Umgebung der Spur der Kurve glatten Funktionen f gilt

$$\frac{d}{dt}f(q(t)) = \langle df, V(t, q(t)) \rangle \quad \text{für alle } t \in (a, b). \quad (6)$$

$$f(q(t_1)) - f(q(t_0)) = \int_{t_0}^{t_1} \langle df, V(t, q(t)) \rangle dt \quad \text{für alle } a < t_0 < t_1 < b. \quad (7)$$

Die Lösungskurven in Koordinatendarstellung. Es sei $x(\cdot) = \{x^j(\cdot) : j \in J\}$ eine Koordinatisierung einer Umgebung der Spur der Kurve. Die Tangentialvektorfeld $V(t, \cdot)$ ordnet dann jedem Ort eine J -Spalte zu.

Andererseits erhalten wir eine Darstellung der Kurve und ihrer Tangentialvektoren durch zeitabhängige J -Spalten. $x(t) = x(q(t))$, $\dot{x}(t) = \frac{d}{dt}x(q(t))$.

Die parametrisierte Kurve $\{x(t) : t \in [a, b]\}$ im Spaltenraum \mathbb{R}_{Sp}^J beschreibt genau dann ein Lösungsstück, wenn $\frac{d}{dt}x^j(q(t)) = v^j(t, q(t))$. Für ein glattes $f(q) = F(x(q))$ mit $df = \sum F_j(x(t)) \cdot dx^j$ gilt dann nämlich

$$\frac{d}{dt}F(x(t)) = \sum F_j(x(t)) \cdot \frac{d}{dt}x^j(t) = \sum F_j(x(t)) \cdot v^j(t, q(t)) = \langle df, V(t, q(t)) \rangle.$$

Satz 1.5.1 (Umparametrisierung). *Es sei $\{q(t) : t \in [a, b]\}$ ein Lösungsstück für die Differentialgleichung $\frac{d}{dt}q(t) = V(t, q(t))$. Es sei $S(t)$ monoton steigend auf $(a, b]$ mit $S'(\cdot) \neq 0$. $T(s)$ sei die Umkehrung. Das Kurvenstück $\{r(s) = q(T(s)) : S(a) < s < S(b)\}$ ist dann ein Lösungsstück der Differentialgleichung $\frac{d}{ds}r(s) = V(T(s), r(s)) \cdot T'(s)$.*

Beweis. $\frac{d}{ds}q(T(s)) = \dot{q}(T(s)) \cdot T'(s) = V(T(s), q(T(s))) \cdot T'(s) = V(T(s), r(s)) \cdot T'(s)$.
Man notiert auch kurz $\frac{dq(T(s))}{dT(s)} = V(T(s), q(T(s)))$,

Sprechweise. Es ist oft bequem aber nicht zwingend, die Variable t als die ‘Zeit’ zu bezeichnen. (Bei unserem Studium der Brachystochrone spielte beispielsweise die Fallhöhe die Rolle der ‘Zeit’.) Wenn in einer Differentialgleichung das Vektorfeld nicht explizit von der ‘Zeit’ abhängt, dann spricht man von einem autonomen System. Bei autonomen

Systemen gestaltet sich die Umparametrisierung besonders einfach. Es geht hier wirklich um Kurven auf der Mannigfaltigkeit N : in den verschiedenen Parametrisierungen werden sie mit verschiedenen Geschwindigkeiten durchlaufen. Bei zeitabhängigen Vektorfeldern betrachtet man manchmal die Lösungskurven als Kurven im 'erweiterten Phasenraum' $\mathbb{R} \times N$.

Beispiel 1.5.3 (Zweidimensionale autonome Systeme). Gegeben sei das autonome System

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} h(x, y) \\ k(x, y) \end{pmatrix},$$

Wenn $h(x, y)$ auf dem gesuchten Lösungsstück $\{(x(t), y(t)) : t \in (a, b)\}$ nicht verschwindet, dann kann man die Lösungskurve durch die 'unabhängige' Variable x parametrisieren. Für die abhängige Variable $y = y(x)$ erhält man eine skalare Differentialgleichung.

Wenn $k(x, y)$ auf dem gesuchten Lösungsstück nicht verschwindet, dann kann man die Lösungskurve durch die 'unabhängige' Variable y parametrisieren. Für die abhängige Variable $x = x(y)$ erhält man ebenfalls eine skalare Differentialgleichung

$$\frac{dy}{dx} = \frac{k(x, y)}{h(x, y)}, \quad \frac{dx}{dy} = \frac{h(x, y)}{k(x, y)}$$

Man formuliert die Suche nach Lösungen der gegebene autonomen Differentialgleichung folgendermaßen: Gesucht ist die Nullkurven für die Pfaff'sche Form

$$\omega = k(x, y) dx - h(x, y) dy;$$

das sind Kurvenstücke $\{\gamma(t) : t \in (a, b)\}$ sodass

$$\langle \omega, \dot{\gamma}(t) \rangle = 0 \quad \text{für alle } t.$$

Besonders einfach wird die Frage, wenn die Form ω ein Differential ist: $\omega = dg$. Man spricht in diesem Falle von einer exakten Differentialgleichung. In diesem Fall sind die Nullkurven die Kurven mit $g(\gamma(t)) = \text{const.}$ — Angenehme Parametrisierungen der Lösungskurven kann man als eine sekundäre Angelegenheit behandeln.

Im folgende Spezialfall sind die Lösungskurven die Kreise:

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}; \quad \frac{dx}{dy} = -\frac{y}{x}; \quad \frac{dy}{dx} = -\frac{x}{y}.$$

$$\omega = x dx + y dy = \frac{1}{2}d(x^2 + y^2) = \frac{1}{2}dr^2 : \quad r^2(\gamma(t)) = \text{const.}$$

Sprechweise 1.5.3 (Euler'scher Multiplikator, exakte Differentialgleichung).

Wenn $M(x, y)$ eine nichtverschwindende Funktion ist, sodass $M(x, y) \cdot \omega$ ein Differential ist, dann heisst M ein integrierender Faktor, oder auch ein Euler'scher Multiplikator für die Pfaff'sche Form ω . Der Euler'sche Multiplikator macht die Differentialgleichung

$$\langle \omega, \dot{\gamma}(t) \rangle = 0 \quad \text{für alle } t.$$

zu der äquivalenten 'exakten' Differentialgleichung $\langle dg, \dot{\gamma}(t) \rangle = 0$ mit der offensichtlichen 'Lösung' $g(\gamma(\cdot)) = \text{const.}$

Euler'sche Multiplikatoren kann man (in sehr speziellen Fällen) manchmal erraten. Es ist aber aussichtslos, einen integrierenden Faktor berechnen zu wollen, wenn man nicht schon gut über die Lösungen der Differentialgleichung Bescheid weiss.

Beispiel (Ein erratener Euler'scher Multiplikator). Die Differentialgleichung ist

$$2xy \cdot \frac{dy}{dx} + (xy^2 + y^2) = 0$$

ist nicht 'exakt', wohl aber die Gleichung

$$2xye^x \cdot \frac{dy}{dx} + (xy^2 + y^2)e^x = 0$$

Die Lösungskurven liegen auf den Niveaulinie der Funktion $g(x, y) = y^2 \cdot x \cdot e^x$.

Beispiel 1.5.4 ('Getrennte Variable'). Eine skalare Differentialgleichung der Form $y' = f(x) \cdot g(y)$ nannte man früher eine Differentialgleichung mit getrennten Variablen. Man lernte, wie man das 'Lösungsproblem' zurückführen kann auf das Problem des 'Auffindens' von Stammfunktionen. Wir sehen das als Beispiel für das Problem für das 'Bestimmen' von Nullkurven zu der geschlossenen Form $\omega = \frac{1}{g(y)} dy - f(x) \cdot dx$.

Wenn $G(y)$ eine Stammfunktion von $\frac{1}{g(y)}$ ist und $F(x)$ eine Stammfunktion von $f(x)$, dann gilt $\omega = d(G(y) - F(x))$. Die Funktion $H(x, y) = G(y) - F(x)$ ist auf den Lösungskurven konstant. Wer eine Lösungskurve unbedingt mit der 'unabhängigen Variable x und der 'abhängigen' Variablen ausdrücken will, der muß die implizite Gleichung $H(x, y) = c$ nach y auflösen. Das funktioniert bekanntlich in der Nähe des Kurvenpunkts (x_0, y_0) , wenn die partielle Ableitung nach y nicht verschwindet, d. h. wenn $0 \neq G'(y_0)$.

Ein konkretes Beispiel bietet das Räuber-Beute-Modell von Lotka-Volterra

$$\begin{aligned}\dot{x} &= kx - axy \\ \dot{y} &= -ly + bxy\end{aligned}$$

im positiven Quadranten $x > 0, y > 0$. Die Parameter a, b, k, l sind allesamt als positiv angenommen.

Wir bemerken vorab, dass das System einen Fixpunkt besitzt, nämlich $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} l/b \\ k/a \end{pmatrix}$.

Wenn man sich zunächst einmal nicht für die 'zeitliche' Parametrisierung der Lösungskurven im positiven Quadranten interessiert, dann kann man folgendermaßen umformen

$$y'(x) = \frac{\dot{y}}{\dot{x}} = \frac{-ly + bxy}{kx - axy} = \frac{-l + bx}{x} \cdot \frac{y}{k - ay}$$

Gesucht im positiven Quadranten sind die Nullkurven der geschlossenen Form

$$\omega = \frac{ay - k}{y} dy + \frac{bx - l}{x} dx = d(G(y) + F(x)).$$

Man überlegt sich leicht, dass die Stammfunktionen $G(y)$ und $F(x)$ für positive Argumente positiv sind. Die Lösungskurven zum Räuber-Beute-Modell sind also die Niveaulinien einer konvexen Funktion, die ihr Minimum im Fixpunkt $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ annimmt.

Beispiel 1.5.5 (Funktionen und Funktionsgraphen).

Wir beginnen mit einer didaktischen Vorbemerkung: In der Terminologie der sehr elementaren Lehrbücher haben die einfachsten Differentialgleichungen die Gestalt $\mathbf{y}' = f(\mathbf{x}, \mathbf{y})$; Lösungen sind Funktionen $\mathbf{y}(x)$ auf einem Intervall I mit $\mathbf{y}'(x) = f(x, \mathbf{y}(x))$.

Wenn man von einer Lösungskurve spricht, dann meint man meist den Funktionsgraphen in der (x, \mathbf{y}) -Ebene, welcher mit Einheitsgeschwindigkeit in x -Richtung durchlaufen wird.

Den Funktionsgraphen $\Gamma = \left\{ \begin{pmatrix} x \\ \mathbf{y} \end{pmatrix} : x \in I, \mathbf{y} = \mathbf{y}(x) \right\}$ kann man nun aber auch mit einer anderen Parametrisierung ('Geschwindigkeit') durchlaufen. Wenn man sich entschließen will, die Lösung mit der Geschwindigkeit $\mathbf{h}(x, \mathbf{y})$ in x -Richtung (und der entsprechenden Geschwindigkeit in \mathbf{y} -Richtung) zu durchlaufen, dann lautet die entsprechende, jetzt zweidimensionale Differentialgleichung

$$\frac{d}{dt} \begin{pmatrix} x \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} h(x, \mathbf{y}) \\ k(x, \mathbf{y}) \end{pmatrix} \quad \text{mit} \quad \frac{h(x, \mathbf{y})}{k(x, \mathbf{y})} = f(x, \mathbf{y}).$$

Ein konkretes Beispiel ist die Gleichung

$$\mathbf{y}' = -\frac{x}{\mathbf{y}} \quad \text{oder auch} \quad \frac{d}{dt} \begin{pmatrix} x \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} -\mathbf{y} \\ x \end{pmatrix}.$$

Die Lösungsmenge dieser homogenen linearen Differentialgleichung ist ein zweidimensionaler Vektorraum; seine Elemente kann man auf verschiedene Weisen schreiben, z. B. als Real- und Imaginärteil der komplexen Linearkombinationen von e^{it} und e^{-it} .

Man kann sich auch für ganz andere Parametrisierungen der Funktionsgraphen entscheiden. Der Funktionsgraph ist aber in jedem Fall eine Nullkurve der geschlossenen 1-Form $\omega = \mathbf{y} \cdot d\mathbf{y} + x \cdot dx = \frac{1}{2}d(x^2 + \mathbf{y}^2)$. Die Funktionsgraphen sind also Niveaulinien der Funktion $r^2 = x^2 + \mathbf{y}^2$, somit Kreisbögen.

Fazit: Die Lösungen der skalaren Differentialgleichung $\mathbf{y}' = -\frac{x}{\mathbf{y}}$ sind die Funktionen $\mathbf{y}(x) = \pm\sqrt{r^2 - x^2}$ in passenden Definitionsbereichen. Es sollte sich aber in Fällen wie dem vorliegenden lohnen, geschickt zu anderen Parametrisierungen des Funktionsgraphen überzugehen.

Beispiel 1.5.6 (Nochmals: Ebene Bewegung in einem zentralen Kraftfeld).

Wir haben oben hergeleitet: Wenn ein Körper sich in einem zentralen Kraftfeld zum Potenzial $\mathbf{U}(r)$ bewegt, dann genügt seine Bahn den Differentialgleichungen

$$r^2 \cdot \dot{\phi} = M (= \text{const}) \quad \text{und} \quad \ddot{r} = -\frac{d}{dr} V(r) \quad \text{mit} \quad V(r) = \mathbf{U}(r) + \frac{M^2}{2r^2}. \quad (8)$$

Wir wollen jetzt den Radius als Funktion der Zeit, $r(t)$, und den Winkel als Funktion des Radius, $\Phi(r)$, genauer studieren. Eine wichtige Rolle neben dem Drehimpuls M spielt natürlich die Gesamtenergie E , die ebenfalls zeitinvariant ist.

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}r^2\dot{\phi}^2 + \mathbf{U}(r) = \frac{1}{2}\dot{r}^2 + \frac{M^2}{2r^2} + \mathbf{U}(r) = E (= \text{const}), \text{ wegen}$$

$$\ddot{r} \cdot \dot{r} = -\frac{d}{dr}V(r) \cdot \dot{r} = -\frac{d}{dt}V(r(t)); \quad \frac{1}{2}\dot{r}^2 + V(r) = \text{const.}$$

Besonders interessant sind für uns die Parameter (M, E) , für welche $r(t)$ zwischen einem minimalen Wert r_p und einem maximalen Wert r_a hin- und herschwankt. r_p und r_a sind Nullstellen der Funktion $(E - U(r))$. Dazwischen haben wir

$$r^2 \cdot \frac{d\phi}{dt} = M, \quad \frac{dr}{dt} = \sqrt{2(E - V(r))} \quad \text{und daher} \quad \frac{d\phi}{dr} = f(r) = \pm \frac{M/r^2}{\sqrt{2(E - V(r))}}.$$

Der Winkel zwischen einer Position minimaler Entfernung und einer Position maximaler Entfernung ist das Integral $\int_{r_p}^{r_a} f(r) dr$. Wir bemerken: Wenn dieser Winkel ein rationales Vielfaches von 2π ist, dann sind die Bahnen zu (M, E) periodisch.

In Keplers Problem, d. h. zu $U(r) = -k/r$, $V(r) = -k/r + \frac{M^2}{2r^2}$, besitzt $f(r)$ eine elementare Stammfunktion $\Phi(r) + \text{const}$, nämlich

$$\Phi(r) + \text{const} = \arccos \frac{M/r - k/M}{2E + k^2/M^2}$$

Wir ersetzen die Parameter (M, E) durch die Parameter $p = M^2/k$, $e = \sqrt{1 + 2EM^2/k^2}$ und erhalten (mit $\text{const} = 0$)

$$\phi = \arccos\left(\frac{1}{e}(p/r - 1)\right); \quad r = \frac{p}{1 + e \cos \phi}.$$

Dies ist nun die Brennpunktgleichung eines Kegelschnitts mit dem ‘Parameter’ p und der ‘Exzentrizität’ e . Für $E < 0$, $e < 1$ handelt sich um eine Ellipsenbahn..

Didaktische These:

In der Frühzeit der Beschäftigung mit Differentialgleichungen galt es als großer Erfolg, wenn man gewisse ein- oder zweidimensionale Probleme auf das Auffinden von (bekannt-) Stammfunktionen zurückführen konnte. Man sagte dann, dass man eine Differentialgleichung explizit lösen konnte. Die Meister im 18. Jahrhundert, allen voran Euler, waren bei diesen Anstrengungen in atemberaubender Weise erfolgreich. Noch entscheidender für ihren Erfolg war aber ihre Kompetenz, die richtigen und praktisch physikalischen Modelle herauszuarbeiten. Heutzutage erscheint für jemanden, der die einschlägige Physik nicht kennt, das Nachrechnen der alten Erfolge nur dann als eine ergiebige Übung, wenn es Licht auf weiterführende Techniken wirft.

Anhang: Von der Lagrangefunktion zur Hamiltonfunktion

Ein Szenario der analytischen Mechanik kann man sowohl durch eine Lagrange-Funktion als auch durch eine Hamilton-Funktion vorgeben. Wenn man von der Idee des Wirkungsintegrals ausgeht, dann gelangt man mit der Konstruktionsmethode der Variationsrechnung zu Differentialgleichungen zweiter Ordnung für die Bahnkurven im n -dimensionale Konfigurationsraum. Die Hamilton-Funktion auf der anderen Seite führt zu Differentialgleichungen erster Ordnung für die Bahnen im $2n$ -dimensionalen Phasenraum. Beide Herangehensweisen sind von großer praktischer und theoretischer Bedeutung, worüber wir aber in einer Anfängervorlesung allenfalls Andeutungen machen können. Der Übergang zwischen den beiden Ansätzen soll uns hier eine Gelegenheit sein, den Tangentialraum und den Cotangentialraum einer Mannigfaltigkeit etwas genauer ins Auge zu fassen.

Das Szenario der Lagrange-Gleichungen sei hier nochmals kurz skizziert:

Für $t \in [t_0, t_1]$ sei $L(t, \cdot, \cdot)$ eine glatte Funktion auf dem Tangentialbündel einer n -dimensionalen Mannigfaltigkeit (N, \mathcal{D}_N) . Für N verwenden wir gelegentlich den Namen Konfigurationsraum; das Cotangentialbündel T_N^* nennen wir gelegentlich den Namen Phasenraum; für das Tangentialbündel T_N haben wir keinen an die analytische Mechanik erinnernden Namen. Bei L denken wir an eine Lagrangefunktion; nähere Bedingungen an die Funktion $L(t, \mathbf{q}, \dot{\mathbf{q}})$ formulieren wir später.

Ein Beispiel, welches nicht aus der analytischen Mechanik stammt, aber doch in den hier vorgesehenen Rahmen passt, haben wir oben beim Problem der Brachystochrone kennengelernt: $L(z, x, \dot{x}) = \frac{1}{\sqrt{z}} \cdot \sqrt{1 + \dot{x}^2}$ für $z \in (0, h]$. Hier spielt z die Rolle, die sonst in der Lagrange-Mechanik die Zeit spielt. Den eindimensionalen ‘Konfigurationsraum’ denken wir uns durch die Variable x koordinatisiert.

Das oben behandelte Kepler-Problem ist in mancherlei Hinsicht sehr speziell: die Lagrangefunktion L ist da unabhängig von der ‘nullten’ Variablen, dort die Zeit genannt. Es ist auch besonders einfach als die Abhängigkeit von der ‘zweiten’ Variablen in jedem (t, \mathbf{q}) eine positiv definite quadratische Form ist. Im Fall der Brachystochrone sind die Funktionen $L(z, x, \cdot)$ immerhin konvexe Funktionen; und diese Besonderheit spielt in der Tat eine wesentliche Rolle in den folgenden Konstruktionen.

Zur Erinnerung: Der Tangentialraum in einer Konfiguration \mathbf{q} ist ein n -dimensionaler Vektorraum $T_{\mathbf{q}}$, der Raum der Geschwindigkeiten der durch \mathbf{q} laufenden glatten Kurven. Wenn eine lokale Koordinatisierung $\{x^j : j \in J\}$ in einer Umgebung einer Konfiguration gewählt ist, dann werden die Geschwindigkeiten durch J -Spalten dargestellt; wenn man zu einer anderen lokalen Koordinatisierung übergeht, dann liefert das eine lineare Abbildung des Spaltenraums. Das hat u. a. die Konsequenz, dass die Eigenschaft, dass $L(t, \mathbf{q}, \cdot)$ eine positiv definite quadratische Form ist, in jeder Koordinatisierung sichtbar wird. Auch die Eigenschaft, dass alle $L(t, \mathbf{q}, \cdot)$ konvex sind, wir in jeder Koordinatisierung sichtbar.

Wir fordern für das Weitere, dass alle $L(t, \mathbf{q}, \cdot)$ runde konvexe Funktionen sind, und wir betrachten in jedem (t, \mathbf{q}) die Legendre-Transformierte

$$H(t, \mathbf{q}, \mathbf{p}) = \sup \{ \langle \mathbf{p}, \dot{\mathbf{q}} \rangle - L(t, \mathbf{q}, \dot{\mathbf{q}}) \} \quad \text{für } \mathbf{p} \in T_{\mathbf{q}}^*.$$

Damit haben hier (für jedes feste t) eine Funktion $H(t, \cdot, \cdot)$ auf dem Cotangentialbündel

T_N^* , dem Phasenraum des (durch L spezifizierten) Systems. Die Funktionen $H(t, q, \cdot)$ sind allesamt konvex. Wir haben früher gelernt, dass die Ableitung $H_2(t, q, \cdot) = \frac{\partial}{\partial p} H(t, q, \cdot)$ die Umkehrabbildung ist zur Abbildung $L_2(t, q, \cdot)$.

Genauer muß man sagen, dass L und H auch den Wert ∞ annehmen können, und dass $L_2(t, q, \cdot)$ einen gewissen konvexen offenen Bereich $U_{t,q} \subseteq T_q$ C^1 -diffeomorph abbildet auf einen gewissen konvexen offenen Bereich $V_{t,q} \subseteq T_q^*$.

$$p = L_2(t, q, \dot{q}) \in V_{t,q} \iff \dot{q} = H_2(t, q, p) \in U_{t,q}.$$

Wenn (für gegebenes (t, q)) gilt $p = L_2(t, q, \dot{q})$, dann nennt man p den Impuls zur Geschwindigkeit \dot{q} ; und man nennt \dot{q} die Geschwindigkeit zum Impuls p (zur Zeit t im Fußpunkt q). Das Supremum, welches die Hamiltonfunktion im Punkt $p = L_2(t, q, \dot{q}) \in V_{t,q}$ liefert, wird im Punkt $\dot{q} \in U_{t,q}$ angenommen; für $p \in V_{t,q}$ gilt

$$H(t, q, p) = \langle p, \dot{q} \rangle - L(t, q, \dot{q}) \quad \text{mit} \quad \dot{q} = \frac{\partial}{\partial p} H(t, q, p).$$

Wir müssen nun die Konstruktionen für die (bisher fixierten) (t, q) zusammenbringen. Die $(2n+1)$ -dimensionale Mannigfaltigkeit $(t_0, t_1) \times T_N^*$ nennt man den erweiterten Phasenraum. Unsere Konstruktion liefert eine Funktion h , die auf einem offenen Teilmenge dieser Mannigfaltigkeit endlichwertig ist. Diese Funktion $h = H(t, q, p)$ wollen wir nun genauer betrachten. Insbesondere interessiert uns das Differential. Dabei benützen wir die Notationen

$$dh = h_0 dt + h_1 dq + h_2 dp = H_0(t, q, p) dt + H_1(t, q, p) dq + H_2(t, q, p) dp.$$

Die Punkte des Endlichkeitsbereich können durch (t, q, \dot{q}) identifiziert werden— man könnte von einer Koordinatisierung des Endlichkeitsbereichs sprechen. Die Formel für h ergibt wegen $p = L_2(t, q, \dot{q})$ das Differential in (t, q, p) -Koordinatisierung

$$\begin{aligned} dh &= \langle p, d\dot{q} \rangle + \langle dp, \dot{q} \rangle - L_1(t, q, \dot{q}) dq - L_2(t, q, \dot{q}) d\dot{q} - L_0(t, q, \dot{q}) dt \\ &= \langle dp, \dot{q} \rangle - L_1(t, q, \dot{q}) dq - L_0(t, q, \dot{q}) dt. \end{aligned}$$

Die Koeffizienten sind eindeutig bestimmte Funktionen. Der Vergleich ergibt

$$H_2(t, q, p) = \dot{q}, \quad H_1(t, q, p) = -L_1(t, q, \dot{q}), \quad H_0(t, q, p) = -L_0(t, q, \dot{q}).$$

Oben haben wir gesagt, dass zu jeder Kurve im Konfigurationsraum eine Kurve im Tangentialbündel $\{(q(t), \dot{q}(t)) : t \in [t_0, t_1]\}$ gehört, mit $\frac{d}{dt} q(t) = \dot{q}(t)$.

Jetzt sagen wir, dass zu jeder Kurve im Konfigurationsraum eine Kurve im Phasenraum gehört. $\{(q(t), p(t)) : t \in [t_0, t_1]\}$. In der Koordinatisierung (t, q, p) haben wir wegen $\dot{q} = H_2(t, q, p)$ die sog. erste Hamilton-Gleichung

$$\frac{d}{dt} q(t) = \frac{\partial}{\partial p} H(t, q(t), p(t))$$

Eine Extremalkurve erfüllt ausserdem die Euler-Lagrange-Gleichungen.

$$\begin{aligned}\frac{d}{dt}L_2(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) &= L_1(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) \\ \frac{d}{dt}\mathbf{p}(t) &= -H_1(t, \mathbf{q}(t), \mathbf{p}(t)) = -\frac{\partial}{\partial \mathbf{q}}H(t, \mathbf{q}(t), \mathbf{p}(t))\end{aligned}$$

Die ist die zweite Hamiltongleichung, und wir haben

$$\frac{d}{dt}\begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix}(t) = \begin{pmatrix} \frac{\partial}{\partial \mathbf{p}} \\ -\frac{\partial}{\partial \mathbf{q}} \end{pmatrix} H(t, \mathbf{q}(t), \mathbf{p}(t))$$

Die augenfällige Symmetrie dieser Gleichung ist der Ansatzpunkt der sog. symplektischen Geometrie. Eine hervorragende Darstellung dieser Zusammenhänge findet man in

V.I. Arnold. *Mathematical Methods of Classical Mechanics*, Springer-Verlag, 1989.

Hinweis: Wenn es darum gehen sollte, Lagrange- oder Hamilton-Gleichungen ‘wirklich zu lösen’, dann kann man von der Umformung schwerlich einen Vorteil erwarten. Die Probleme werden lediglich(?) in verschiedene theoretische Zusammenhänge gestellt.

1.6 Lösungskurven; dynamische Systeme

Die speziellen Differentialgleichungen in unseren Beispielen hatten allesamt mehrere Lösungen, und wir haben uns selten Gedanken gemacht, für welche der Lösungen (über welchem Zeitintervall) wir uns besonders interessieren sollten. Eine wichtige Vorgehensweise, eine spezielle Lösungskurve auszusondern, ergibt sich bei den sog. Anfangswertaufgaben.

Denken wir z. B. an den frei (d.h. ohne äussere Krafteinwirkung) schwingenden harmonischen Oszillator. Wenn zu einer Zeit t_0 die Position und die Geschwindigkeit festgestellt ist, dann liegt die Lösungskurve über alle Zeiten fest, und zwar sowohl in der Zukunft als auch in der Vergangenheit.) Bei den vektorwertigen Differentialgleichungen erster Ordnung wird die Anfangswertaufgabe (AWA) folgendermaßen gestellt:

Sprechweise 1.6.1 (AWA). Gegeben ist ein d -Tupel stetiger Funktionen $V(t, y)$ in einer Umgebung eines Punkts $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}^d$. Es soll geklärt werden

1. Gibt es ein Kurvenstück $\{y(t) : t \sim t_0\}$ mit $\dot{y}(t) = V(t, y(t))$ und $y(t_0) = y_0$? ('Existenz')
2. Liefert die AWA zu (t_0, y_0) ein eindeutig bestimmtes Kurvenstück $Y((t_0, y_0).\cdot)$? ('Eindeutigkeit')
3. Im Falle, dass die Frage von Existenz und Eindeutigkeit für alle (t_0, y) in einer Umgebung von (t_0, y_0) positiv zu beantworten sind, stellt sich die weitere Frage: Ist die Schar der Kurvenstücke $\{Y((t_0, y).\cdot) : y \in U\}$ stetig in einer Umgebung U von y_0 (im Sinne der lokal gleichmäßigen Konvergenz.) ('Stetige Abhängigkeit von den Anfangsdaten')

Wenn auch die stetige Abhängigkeit von der Anfangsposition gewährleistet ist, dann spricht man von einer 'wellposed AWA'. (Deutsche Übersetzungen von 'wellposed' wie etwa 'korrekt gestellt' sind nicht sehr gebräuchlich.)

Didaktischer Exkurs: Allgemeines zu Existenz und Eindeutigkeit: Bei der Frage nach Existenz und Eindeutigkeit der Lösung eines mathematischen Problems muss man (streng genommen!) immer festhalten, in welchen Bereichen mathematischer Objekte man die Lösungen sucht. Die Frage nach der Existenz ist umso eher positiv zu beantworten, je größer der Objektbereich ist. (Die Gleichung $x^2 = 2$ besitzt im Bereich der rationalen Zahlen keine Lösung; im Bereich der reellen Zahlen besitzt sie zwei Lösungen.)

Andererseits: Wenn der Objektbereich sehr groß ist, dann kann es schwierig sein, zu beweisen, dass eine bekannte naheliegende Lösung die einzige Lösung des Problems ist.

Ein bekanntes Beispiel ist das folgende: Gesucht sind die additiven reellwertigen Funktionen $f(\cdot)$ mit $f(1) = 1$. ($(f(a+b) = f(a) + f(b))$ für alle $a, b \in \mathbb{R}$) Es ist klar, dass für jede Lösung f des Problems gilt $f(r) = r$ für alle rationalen r . Wie aber steht es mit den irrationalen Argumentwerten? Übrigens: Die modernen Mathematiker glauben, dass es unstetige additive Funktionen gibt, obwohl sie noch nie eine gesehen haben und auch nicht erwarten, jemals eine zu sehen.

Beispiel 1.6.1. Ein einfaches Anfangswertproblem, welches nicht wellposed ist, ist das folgende: Gesucht ist eine skalare Funktion $y(\cdot)$ mit $\dot{y}(t) = |y(t)|^{2/3}$, $y(s) = 0$, Anders ausgedrückt: Gesucht sind alle stetigen $y(\cdot)$ mit

$$y(t) = \int_s^t |y(u)|^{2/3} du \quad \text{für } t \sim s.$$

Eine naheliegende Lösung ist $y(t) \equiv 0$. Es gibt aber noch weitere stetig differenzierbare Lösungen. Jedes Paar von Zeitpunkten $t' \leq 0 \leq t''$ gibt Anlass zu einer Lösung.

$$y(t) = \begin{cases} \frac{1}{27}(t - t'')^3 & \text{für } t \geq t'' \\ 0 & \text{für } t \in [t', t''] \\ \frac{1}{27}(t - t')^3 & \text{für } t \leq t' \end{cases}$$

Diese Lösung kommt aus $-\infty$, mündet dann zum Zeitpunkt t' in die Abszissenachse ein, um sich bei t'' wieder als eine kubische Parabel zu erheben.

Wir werden uns dem Problem von Existenz und Eindeutigkeit für Anfangswertaufgaben hier noch nicht nachdrücklich stellen. Wir wollen zunächst einen berühmten Satz beweisen, welcher besagt, dass gewisse Bedingungen an $V(\cdot, \cdot)$ hinreichend sind für die 'Wellposedness' der AWA auf C^1 -Mannigfaltigkeiten.

Im Folgenden bevorzugen wir manchmal die folgende Notationen für die AWA zu (s, \tilde{y}) :

$$\begin{aligned} \dot{y} &= V(t, y(t)) & y(s) &= \tilde{y} & \text{oder} \\ y(t) &= \tilde{y} + \int_s^t V(u, y(u)) du \end{aligned}$$

Die zweite Formulierung zeigt, dass eine stetige Lösung notwendigerweise stetig differenzierbar ist, jedenfalls für jedes stetige $V(\cdot, \cdot)$. (Von $V(\cdot, \cdot)$ werden wir stets mehr Regularität als nur die Stetigkeit fordern.)

Satz 1.6.1 ((Satz von Picard-Lindelöf)).

Sei $V(\cdot, \cdot)$ stetig in einer Umgebung von $(s, \tilde{y}) \in \mathbb{R} \times \mathbb{R}^n$, und es existiere eine Zahl L , sodass in dieser Umgebung gilt

$$\|V(t, y_2) - V(t, y_1)\| \leq L \|y_2 - y_1\| \quad \text{für alle } t \sim s \text{ und alle } y_1, y_2 \text{ in einer Umgebung von } \tilde{x}.$$

(„Gleichmäßige Lipschitz-Stetigkeit in der y -Richtung“).

Es gilt dann

a) In einem Intervall $(t', t'') \ni s$ existiert eine stetig differenzierbare Lösung der AWA,

b) Jede absolutstetige Funktion $z(\cdot)$ mit $z(t) = \tilde{y} + \int_s^t V(u, z(u)) du$ stimmt in einer Umgebung von s mit dieser Lösung überein.

Beweis. 1. Wir beweisen zuerst die Eindeutigkeit: Seien $\mathbf{y}(\cdot)$ und $\mathbf{z}(\cdot)$ Lösungen der AWA mit $(t, \mathbf{y}(t))$ und $(t, \mathbf{z}(t))$ in dem Bereich der gleichmäßigen Lipschitz-Stetigkeit. $\Delta(\cdot)$ sei ihre Differenz. Es gilt dann

$$\begin{aligned}\|\Delta(t)\| &\leq \int_s^t \|\mathbf{V}(\mathbf{u}, \mathbf{y}(\mathbf{u})) - \mathbf{V}(\mathbf{u}, \mathbf{z}(\mathbf{u}))\| \, d\mathbf{u} \\ &\leq L \cdot \int_s^t \|\Delta(\mathbf{u})\| \, d\mathbf{u}\end{aligned}$$

Es kommt nicht darauf an, welche Norm auf dem \mathbb{R}^n wir wählen; für die Funktionen benutzen wir die Supremumsnorm $\|\cdot\|_\infty$. Wenn wir $|t - s|$ so kurz wählen, dass $L \cdot |t - s| < 1$, dann ergibt sich $\|\Delta(\cdot)\| = 0$ für diese t .

2. Wir wollen die Existenz nicht für den speziellen Anfangswert $\tilde{\mathbf{y}}$ beweisen, sondern für alle $\mathbf{y}(s)$ in der Nähe von $\tilde{\mathbf{y}}$. Wir suchen also eine lokale Lösung der 'Integralgleichung'

$$\mathbf{y}(t) = \mathbf{y}(s) + \int_s^t \mathbf{V}(\mathbf{u}, \mathbf{y}(\mathbf{u})) \, d\mathbf{u}.$$

Dazu verwenden wir den Banach'schen Fixpunktsatz. Wir beginnen mit der zeitlich konstanten Funktion $\mathbf{y}_0(\cdot) = \mathbf{y}(s)$ und konstruieren dann rekursiv

$$\mathbf{y}_{n+1}(t) = \mathbf{y}_n(s) + \int_s^t \mathbf{V}(\mathbf{u}, \mathbf{y}_n(\mathbf{u})) \, d\mathbf{u} = \Psi(\mathbf{y}_n)(t).$$

Wenn wir Konvergenz beweisen können für t in einer Umgebung von s für alle $\mathbf{y}(s)$ in einer Umgebung von $\tilde{\mathbf{y}}$, dann haben wir im Limes eine Schar von Kurven, parametrisiert durch ihren Wert zum Zeitpunkt s , welche allesamt die Gleichung erfüllen.

3. Die lokal gleichmäßige Lipschitz-Eigenschaft garantiert, dass die Voraussetzungen für den Fixpunktsatz erfüllt sind. Sind nämlich $\mathbf{y}(\cdot)$ und $\mathbf{z}(\cdot)$ Kurvenscharen mit $\mathbf{y}(s) = \mathbf{z}(s)$, so gilt

$$\|\Psi(\mathbf{y})(t) - \Psi(\mathbf{z})(t)\| \leq \int_s^t \|\mathbf{V}(\mathbf{u}, \mathbf{y}(\mathbf{u})) - \mathbf{V}(\mathbf{u}, \mathbf{z}(\mathbf{u}))\| \, d\mathbf{u} \leq L \cdot \int_s^t \|\mathbf{y}(\mathbf{u}) - \mathbf{z}(\mathbf{u})\| \, d\mathbf{u}$$

Wir wählen das Zeitintervall so klein, dass mit einem $\alpha < 1$ gilt

$$\sup_t \|\Psi(\mathbf{y})(t) - \Psi(\mathbf{z})(t)\| \leq \alpha \cdot \sup_t \|\mathbf{y}(t) - \mathbf{z}(t)\|.$$

Wir verweisen auf den Fixpunktsatz in seiner lokalen Form. Eventuell müssen wir das Zeitintervall und die r -Kugel um die Ausgangsposition noch weiter verkleinern, um auch noch zu sichern, dass der erste Schritt weg von der konstanten Kurvenschar \mathbf{y}_0 eine Norm kleiner als $(1 - \alpha) \cdot r$ liefert.

Glatte Tangentialvektorfelder.

Die lokal gleichmäßige Lipschitz-Eigenschaft des Tangentialvektorfelds spielt eine nützliche Rolle im Beweis des Satzes von Picard-Lindelöf. Für weitergehende Untersuchungen stützt man sich gerne auf eine schärfere Regularitäts-Annahme, die wir im Folgenden diskutieren wollen. Wir holen weiter aus und betrachten zunächst einmal zeitunabhängige Tangentialvektorfelder auf einer \mathcal{C}^2 -Mannigfaltigkeit (N, \mathcal{D}) .

Gegeben sei ein Tangentialvektorfeld $v = \{v(q) : q \in U\}$ auf einer Kartenumgebung U . Die Koordinatisierung auf U sei bezeichnet $\varphi(\cdot) = \{x^j(\cdot) : j \in J\}$, oder auch kurz $x(\cdot)$ mit Werten im Raum der J -Spalten. Der Anstieg der Koordinatenfunktion x^j in der Richtung des Feldes sei $v^j = \langle dx^j, v \rangle$. Wir fordern nun, dass diese Funktionen v^j einmal stetig differenzierbar ist; und wir drücken diese Forderung dadurch aus, dass wir v ein 1-glatte Vektorfeld (auf der 2-glatte Mannigfaltigkeit U) nennen. Wir schreiben $v = \sum v^j \cdot \frac{\partial}{\partial x^j}$.

Wir müssen zeigen, dass die so definierte 1-Glattheit wirklich eine Eigenschaft des Vektorfelds ist, also eine Eigenschaft, die nicht auf die spezielle Koordinatisierung Bezug nimmt. Sei $y(\cdot) = \{y^i(\cdot) : i \in I\}$ eine weitere Koordinatisierung. Die Einträge der Jacobi-Matrix $J^i_j = \frac{\partial y^i}{\partial x^j}$ sind dann einmal stetig differenzierbar, und ebenfalls stetig differenzierbar sind die Koeffizienten w^i des Vektorfelds.

$$w^i = \langle dy^i, v \rangle = \sum_j J^i_j \cdot \langle dx^j, v \rangle = \sum_j J^i_j \cdot v^j.$$

Die Rechnung zeigt in der Tat, dass das Vektorfeld v genau dann 1-glatte ist, wenn $\langle df, v \rangle$ stetig differenzierbar ist für jedes 2-glatte f . In der Tat ist $\langle \omega, v \rangle$ stetig differenzierbar für jede Pfaff'sche Form ω auf der \mathcal{C}^2 -Mannigfaltigkeit.

Definition 1.15. Ein Vektorfeld v auf einer \mathcal{C}^2 -Mannigfaltigkeit (N, \mathcal{D}) wird eine 1-glatte Vektorfeld genannt, wenn $\langle \omega, v \rangle$ stetig differenzierbar ist für jede Pfaff'sche Form.

Bemerke: Wenn $\{x^j(\cdot) : j \in J\}$ irgendeine lokale Koordinatisierung ist, dann hat jedes glatte Vektorfeld v lokal die Gestalt $v = \sum v^j \cdot \frac{\partial}{\partial x^j}$ mit stetig differenzierbaren v^j . Andererseits hat jede Pfaff'sche Form lokal die Gestalt $\omega = \sum_j g_j \cdot dx^j$ mit stetig differenzierbaren g_j . Und es gilt

$$\langle \omega, v \rangle = \left\langle \sum_j g_j \cdot dx^j, \sum_j v^j \cdot \frac{\partial}{\partial x^j} \right\rangle = \sum g_j \cdot v^j.$$

Ein 1-glatte Vektorfeld erfüllt (in jeder lokalen Koordinatisierung) die Voraussetzungen des Satzes von Picard-Lindelöf; denn eine stetig differenzierbare Funktion im \mathbb{R}^n ist auf jeder kompakten konvexen Teilmenge Lipschitz-stetig. Wenn $V^j(t, \cdot)$ für jedes t in einer Intervallumgebung von s stetig differenzierbar ist, und die Differentiale $dV^j(t, \cdot)$ in t stetig sind, dann sind in jedem (t, q) die Voraussetzungen des Satzes von Picard-Lindelöf erfüllt.

Sprechweise 1.6.2. Wenn das zeitabhängige Vektorfeld $v(t, q)$ diese Bedingungen erfüllt, dann sagen wir, dass es die verschärften P.L.-Bedingungen erfüllt.

Ein interessanter Aspekt der zeitunabhängigen Vektorfelder ergibt sich durch das sog. Lie-Produkt. Wir betrachten C^∞ -Vektorfelder, die wir mit Buchstaben wie X, Y, Z bezeichnen. Die Ableitung einer Funktion entlang des Vektorfelds bezeichnen wir mit Xf oder mit $X(f)$.

Satz 1.6.2. *Wenn X und Y C^∞ -Vektorfelder sind, dann auch der Kommutator $Z = [X, Y]$:*

$$f \mapsto [X, Y](f) = X(Yf) - Y(Xf)$$

Wenn h, k glatte Funktionen sind und $\hat{k} = Xk, \hat{h} = Yh$, dann gilt

$$[h \cdot X, k \cdot Y] = hk \cdot [X, Y] + h\hat{k} \cdot Y - k\hat{h} \cdot X.$$

Beweis.

Nach dem Satz von oben (3. Charakterisierung) müssen wir nur zeigen, dass die Zuordnung die Produktregel erfüllt.

$$[X, Y](f \cdot g) \stackrel{?}{=} f \cdot [X, Y](g) + g \cdot [X, Y](f).$$

Dies ist nun aber eine einfache Rechenaufgabe.

$$\begin{aligned} X(Y(f \cdot g)) &= X(f \cdot Y(g) + g \cdot Y(f)) = \\ &= X(f) \cdot Y(g) + X(g) \cdot Y(f) + f \cdot X(Yg) + g \cdot X(Yf) \\ Y(X(f \cdot g)) &= X(f) \cdot Y(g) + X(g) \cdot Y(f) + f \cdot Y(Xg) + g \cdot Y(Xf). \end{aligned}$$

Subtraktion ergibt die erste Behauptung. Die Formel für das Zusammenspiel mit der Modulstruktur ergibt sich ebenfalls durch eine einfache Rechnung.

Das Klammerprodukt ist eine antikommutative bilineare Verknüpfung: $([Y, X] = -[X, Y])$. Die Verknüpfung ist nicht assoziativ. An die Stelle des Assoziativgesetzes tritt die berühmte Jacobi-Identität, welche besagt:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0.$$

Das Nachrechnen überlassen wir dem Leser.— Wir wollen die algebraische Struktur hier nicht weiter verfolgen.

Die Sprache der dynamischen Systeme

Die Theorie der Anfangswertaufgaben hat ihre Wurzeln in der Newton'schen Mechanik. In Newtons Theorie ist der Zustand eines Systems durch Ort und Impuls der beteiligten Massenpunkte beschrieben. Wenn der Zustand zu einem bestimmten Zeitpunkt festgelegt ist, dann ist die zeitliche Entwicklung eindeutig festgelegt, übrigens nicht nur für die Zukunft, man kann auch in eindeutiger Weise auf die Vergangenheit schliessen. Die Menge aller möglichen Zustände heisst der Phasenraum.

Wir bemerken: In der Newton'schen Mechanik hat man neben dem Begriff des Phasenraums auch noch den Begriff des Konfigurationsraums. Man stellt sich vor: Unter dem Eindruck von „holonomen“ Zwangsbedingungen (die dem System m Freiheitsgrade belassen) wird der Konfigurationsraum zu einer m -dimensionalen Mannigfaltigkeit und der Phasenraum N zu einer $2m$ -dimensionalen Mannigfaltigkeit. Die den Gesetzen der Mechanik entsprechenden Zustandsänderungen können durch ein Vektorfeld beschrieben werden, welches auf die lokal wirkenden Kräfte zurückzuführen ist.

Wir bemerken: Ein Kraftfeld auf dem Konfigurationsraum (oder auf dem Tangentialbündel) ist zunächst einmal kein Vektorfeld, sondern ein Covektorfeld (wie z. B. das Differential eines Potentials.) Es ist eine der Mechanik eigentümliche Konstruktion, welche aus einem Cotangentialvektorfeld auf dem Tangentialbündel des Konfigurationsraum ein Tangentialvektorfeld auf dem Phasenraum macht. Die Konstruktion entspricht dem Übergang von den Lagrange-Mechanik zur Hamilton-Mechanik. —Der Sachverhalt wird sehr überzeugend aufgeklärt in dem mehrfach erwähnten Lehrbuch MMCM von V. I. Arnold. Wenn die herrschenden Kräfte im Laufe der Zeit nicht verändert werden, dann hat man ein „autonomes dynamisches System“; dem entspricht ein zeitunabhängiges Vektorfeld. Wenn ein zeitabhängiges Vektorfeld $\{v(t, \cdot) : t \in \mathbb{R}\}$ am Werke ist, dann spricht man von einem nichtautonomen System.

In der modernen Theorie der dynamischen Systeme wird diese Sprache auf allgemeinere Szenarien übertragen. Von einem Konfigurationsraum ist hier nicht mehr die Rede, und der 'Phasenraum' hat nicht notwendigerweise eine gerade Dimension. Mit einem Vektorfeld v auf einer Mannigfaltigkeit (N, \mathcal{D}) assoziiert man dennoch die Vorstellung von einem autonomen dynamischen System. Die Mannigfaltigkeit wird als der Phasenraum des dynamischen Systems bezeichnet, seine Punkte P heissen die möglichen Zustände des Systems. Mit einem zeitabhängigen Vektorfeld $v(\cdot, \cdot)$ assoziiert man die Vorstellung von einem nichtautonomen dynamischen System.

Es zeigt sich: Wenn das Vektorfeld eine gewisse Glattheit besitzt (wie z. B. die verschärften P.L.-Bedingungen), dann ist es die Vorstellung am Platze, dass $v(\cdot, \cdot)$ determiniert, wie sich die Zustände in der Zeit verändern. In jedem Fall definiert man

Definition 1.16.

Eine glatt parametrisierte Kurve $\gamma(t)$ auf N heißt ein Stück einer Lösungskurve für das zeitabhängige Vektorfeld $v(\cdot, \cdot)$, wenn gilt

$$\begin{aligned} \dot{\gamma}(t) &= v(t, \gamma(t)) && \text{für } t \in (t', t''), \text{ d.h.} \\ \langle df, \dot{\gamma}(t) \rangle &= \langle df, v(t, \gamma(t)) \rangle && \text{für alle glatten } f. \end{aligned}$$

Wir haben oben bewiesen

Satz 1.6.3 (Maximale Lösungskurven).

Ist $v(\cdot, \cdot)$ ein zeitabhängiges Vektorfeld, welches die verschärften P.L.-Bedingungen erfüllt, so existiert zu jedem $(s, P) \in \mathbb{R} \times M$ gibt es genau eine maximale Lösungskurve

$$\{\gamma(t) : t \in (t', t'')\} \quad \text{mit } \gamma(t_0) = P.$$

Maximal bedeutet, dass die Kurve nicht Teil einer in einem größeren Intervall definierten Lösungskurve ist. Der Wert $t'(s, P) < s$ heisst die Vergangenheit, $t''(s, P) > s$ heisst die Zukunft für (s, P) . Das Zeitintervall $(t'(s, P), t''(s, P))$ kann endlich sein, aber auch nach oben oder nach unten unendlich. Man kann leicht zeigen, dass bei festem s die Funktionen $t''(P)$ unterhalbstetig ist. Das bedeutet: Wenn $t''(\tilde{P})$ endlich ist, dann gilt in einer vollen Umgebung $t''(P) > t''(\tilde{P}) - \varepsilon$ für jedes $\varepsilon > 0$.

Man sagt von einer maximalen Lösungskurve, dass sie nach Unendlich strebt, wenn für jedes Kompaktum ein Zeitpunkt existiert, sodass die Lösungskurve das Kompaktum danach nicht mehr trifft. Lösungskurven mit endlicher Zukunft streben notwendigerweise nach Unendlich.

Beispiel 1.6.2 (Endliche und unendliche Zukunft).

Die skalare Differentialgleichung $\dot{y} = -y^2$ erfüllt überall in $\mathbb{R} \times \mathbb{R}$ die Bedingungen des Satzes von Picard-Lindelöf. Zu jedem Anfangswert (t_0, y_0) gibt es genau eine Lösung der AWA. Eine offensichtliche Lösungskurve ist die Konstante $y(\cdot) = 0$. Die übrigen Lösungskurven sind entweder strikt positiv oder strikt negativ. Sie sind allesamt monoton fallend. Die positiven haben eine unendliche Zukunft, aber eine beschränkte Vergangenheit. Die negativen haben eine beschränkte Zukunft. Man kann die Lösungskurven in der Tat explizit angeben. Zu jedem c gehören zwei Lösungskurven, nämlich

$$y_{c+}(t) = \frac{1}{t-c} \text{ für } t > c \quad \text{sowie} \quad y_{c-}(t) = \frac{1}{t-c} \text{ für } t < c.$$

Der erweiterte Phasenraum

Die Lösungsstücke einer Differentialgleichungen $\dot{q} = v(t, q)$ sind parametrisierte Kurven auf der Mannigfaltigkeit N , dem n -dimensionalen Phasenraum $\{q(t) : t \in (t', t'')\}$. Manchmal empfiehlt es sich, daneben auch die Kurven $\{(t, q(t)) : t \in (t', t'')\}$ im cartesianischen Produkt $\mathbb{R} \times N$ zu betrachten. Die nullte Koordinate steigt mit Einheitsgeschwindigkeit mit t . Das zeitabhängige Vektorfeld wird erweitert zu einem Vektorfeld $\hat{v}(\cdot, \cdot)$ auf dem 'erweiterten Phasenraum' $\mathbb{R} \times N$. Es ist da als ein zeitunabhängiges Vektorfeld zu verstehen. Entlang \hat{v} kann man nun auch zeitabhängige glatte Funktionen $f(t, q)$ ableiten. Zu der gewohnten $(v(t, \cdot))$ -Ableitung, welches für jedes feste t zu berechnen ist, kommt an der Stelle (t, q) noch (als Summand) die partielle Ableitung $\frac{\partial}{\partial t} f$ hinzu.

$$\langle df(\cdot, \cdot), \hat{v} \rangle|_{t,q} = \frac{\partial}{\partial t} f|_{t,q} + \langle df(t, \cdot), v(t, \cdot) \rangle|_q.$$

In einer lokalen Koordinatisierung haben wir also $\hat{v} = \frac{\partial}{\partial t} + \sum_j v^j \cdot \frac{\partial}{\partial x^j}$.

Der Übergang zum erweiterten Phasenraum ist bequem für die Formulierung der AWA's. Während in der ursprünglichen Auffassung maximale Lösungskurven durch eine Bedingung $\gamma(s) = P$ ausgewählt wurden, kann man diese Kurven jetzt einfach als die maximale Lösungskurve durch den Punkt (s, P) beschreiben. Offenbar gilt

Satz 1.6.4. *Wenn ein zeitabhängigen Vektorfeld $v(\cdot, \cdot)$ die verschärften P.L.-Bedingungen erfüllt, dann bilden die maximalen Lösungskurven eine Partition des erweiterten Phasenraums.*

Die folgenden Konstruktionen kann man sowohl auf den Phasenraum als auch auf den erweiterten Phasenraum beziehen. Der Ausgangspunkt ist in jedem Fall ein zeitabhängiges Vektorfeld $v(\cdot, \cdot)$, welches die verschärften P.L.-Bedingungen erfüllt.

Definition 1.17 (Globaler Fluß und Lie-Ableitung).

Man sagt von einem Vektorfeld $v(\cdot, \cdot)$, dass es einen globalen Fluss liefert, wenn alle Lösungskurven unendliche Zukunft und unendliche Vergangenheit haben.

Der Punkt auf der Mannigfaltigkeit, welchen die Lösungskurve, die zur Zeit s durch P geht, zur Zeit t erreicht, soll hier mit $\varphi_s^t(P)$ bezeichnet werden. Für die Schar der Abbildungen $\{\varphi_s^t(\cdot) : s, t \in \mathbb{R}\}$ folgt aus der eindeutigen Lösbarkeit der AWA's

- i) $\varphi_s^s(\cdot) = \text{Identität auf } M$, für alle s
- ii) $\varphi_t^u(\varphi_s^t(\cdot)) = \varphi_s^u(\cdot)$ für alle s, t, u .
- iii) $\frac{d}{dt}f(\varphi_s^t(P)) = \langle df, v(t, \varphi_s^t(P)) \rangle$ für alle f und $P \in M$.

Definition 1.18 (Lokaler Fluß und die Flussaxiome).

Wenn manche Lösungskurven nur eine endliche Zukunft (oder endliche Vergangenheit) besitzen, dann nennt man die entsprechen Abbildungen $\{\varphi_s^t : s, t \in \mathbb{R}\}$ den „lokalen Fluss“ zum Vektorfeld (\cdot, \cdot) .

Die Abbildungen $\varphi_s^t(\cdot)$ sind hier für t in Umgebungen von Punkten definiert, für welche die Lösungskurve zu (s, P) eine Zukunft $> t$ bzw. (im Falle $t < s$) eine Vergangenheit $< t$ hat. Das „Flussaxiom“ (ii) ist so zu lesen: Immer wenn für ein (s, t, u, P) die linke Seite definiert ist, dann ist auch die rechte Seite definiert und sie liefert denselben Punkt auf der Mannigfaltigkeit. Die lokalen Eigenschaften (iii) gelten unverändert auch für den lokalen Fluss.

Hinweis: Ein berühmter, aber nicht leicht zu beweisender Satz ist der folgende: Das Vektorfeld $v(\cdot, \cdot)$ erfülle die verschärften Picard-Lindelöf-Bedingungen. Wenn es einen globalen Fluss generiert, dann sind die φ_s^t allesamt C^1 -Diffeomorphismen; die stetige Differenzierbarkeit der $v(t, \cdot)$ garantiert also die stetige Differenzierbarkeit der Abbildungen φ_s^t . Im Falle eines lokalen Flusses ist die Aussage entsprechend zu modifizieren: Zu jedem (s, P) gibt es eine Intervallumgebung I von s und eine Umgebung U_s von P , sodass die Abbildungen φ_u^t stetig differenzierbare Abbildungen von $U_u = \varphi_s^u U_s$ auf $U_t = \varphi_s^t U_s$ liefert, wobei für $t, u \in I$ gilt $\varphi_t^u = (\varphi_u^t)^{-1}$ und $\varphi_t^u = \varphi_s^u \circ \varphi_s^t$.

Interpretation des Klammerprodukts

Auf der C^∞ -Mannigfaltigkeit seien ∞ -glatte Vektorfelder X und Y gegeben. $Z = [X, Y]$ sei das Klammerprodukt. $\{\varphi^h : h \in \mathbb{R}\}$ sei der lokale Fluss zu X , entsprechend seien ψ^h zu Y und χ^h zu $Z = [X, Y]$ definiert. Betrachten wir einen Punkt, der zuerst über eine kurze Zeit s gemäß X verschoben wird, und dann über eine kurze Zeit t gemäß Y . Die Bildpunkte bezeichnen wir mit $P_s = \varphi^s(P)$ bzw. $P_{st} = \psi^t \circ \varphi^s(P)$. Entsprechend seien Q_t

und Q_{ts} definiert. Für jede glatte Funktion f gilt dann

$$\begin{aligned} & f(P_{st}) - f(P_s) - f(Q_t) + f(P) \\ &= (f(\psi^t(P_s)) - f(P_s)) - (f(\psi^t(P)) - f(P)) \\ &\approx t \cdot Y(f)(P_s) - t \cdot Y(f)(P) \\ &\approx t \cdot s \cdot X(Yf)(P) \\ & f(Q_{ts}) - f(P_s) - f(Q_t) + f(P) \\ &\approx s \cdot t \cdot Y(Xf)(P) \end{aligned}$$

$$\lim_{s,t \rightarrow 0} \frac{1}{st} (f(P_{st}) - f(P_{ts})) = X(Yf)(P) - Y(Xf)(P) = \langle df, [X, Y] \rangle|_P.$$

Sprechweise 1.6.3 (Orbits und das Phasenporträt).

Wir denken an das autonome System zu einem stetig differenzierbaren Vektorfeld v auf einer \mathcal{C}^2 -Mannigfaltigkeit. Die Menge der Punkte, die auf derselben Lösungskurve wie P liegen, heißt die Bahn von P oder auch der Orbit von P .

$$O(P) := \{ \varphi^u(P) : u \in (t'(P), t''(P)) \} .$$

Die Orbits in einem dynamischen System sind spezielle Typen von Spuren von Kurven: Wenn ein Orbit nicht in dem Sinn entartet ist, dass er aus einem einzigen Punkt besteht („stationärer Punkt“), dann hat die Lösungskurve durch ihn überall einen nicht-verschwindenden Tangentialvektor; der Orbit ist dann in der Tat eine eindimensionale Mannigfaltigkeit.

Man kann zeigen: Die nichtentarteten Orbits können lokal als gemeinsame Lösungsmengen von $n - 1$ glatten Funktionen beschrieben werden können (Ohne Beweis! Das Stichwort zu dieser Aussage lautet ‘Erste Integrale’.)

Einem Orbit sieht man i. Allg. nicht an, ob er zu einer Lösungskurve mit unendlicher oder endlicher Zukunft (oder Vergangenheit) gehört. Es gibt nur drei Typen von Orbits: die einpunktigen, die „periodischen“ und die injektiven. Wenn nämlich eine Lösungskurve einen Doppelpunkt hat, $\varphi_{t_0}^{t_0+T}(P) = P$ für ein $T \neq 0$, dann impliziert das $\varphi_s^{s+T}(P) = P$ für alle s . Für autonome Systeme gilt eine Aussage, die analog ist zu dem obigen Satz über erweiterte Phasenräume. Die Menge aller Orbits ist eine einfache Überdeckung (‘Partition’) der Mannigfaltigkeit. Diese Überdeckung heißt das Phasenporträt zum stetig differenzierbaren Vektorfeld v .

Wir bemerken:

1. Man kann am Phasenporträt zu einem stetig differenzierbaren Vektorfeld nicht ablesen, welche Zeit nötig ist, von einem Punkt auf dem Orbit in einen anderen zu gelangen; man kann nur sehen, von wo nach wo man irgendwann hingelangen kann.
2. Das Phasenporträt hängt nicht wirklich von v ab; entscheidend ist das daraus abgeleitete „Richtungsfeld“; wenn k eine strikt positive Funktion ist, dann haben v und $k \cdot v$ dasselbe Richtungsfeld und damit dasselbe Phasenporträt. Ein solches k liefert eine Umparametrisierung der Lösungskurven.

Beispiel 1.6.3. (Das mathematische Pendel)

Betrachte einen schwerebenen Stab der Länge ℓ , der am einen Ende festgehalten wird und am anderen die Masse m trägt. ϑ sei die Auslenkung aus der Vertikalen. Es gilt

$$\frac{1}{2}m\ell^2\ddot{\vartheta} = -mgl \cdot \sin \vartheta .$$

Durch geeignete Wahl der Zeitskala erhalten wir $\ddot{x} = -\sin x$, und als System geschrieben

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -\sin x_1 \end{pmatrix}$$

Auf jedem Orbit ist die Gesamtenergie E konstant. Diese Energie ist die Summe von potentieller und kinetischer Energie

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

Wir haben stationäre Punkte $(k \cdot \pi, 0)$ auf der Abszissenachse. Die Orbits in der Nähe der $(2k\pi, 0)$ sehen aber ganz anders aus als diejenigen bei $((2k+1)\pi, 0)$.

Zu $E > 2$ gibt es zwei Lösungen. Der Winkel strebt nach $-\infty$ oder nach $+\infty$. Die potentielle Energie ist eine periodische Funktion des Winkel.

$$x_2 = \pm \sqrt{2E - 2(1 - \cos x_1)} .$$

Für $E < 2$ erhalten wir periodische Orbits. Für kleine E sehen sie aus wie kleine Kreise. Der Winkel schwankt um die Nulllage.

$$\frac{1}{2}x_2^2 + \frac{1}{2}x_1^2 = E .$$

Für $E = 2$ haben wir neben den stationären Punkten Orbits, die sich einem stationären Punkt nähern, ihn aber (natürlich) nicht erreichen.

1.7 Rückblick und Ausblick

Der Begriff der glatten Funktion setzt beim Definitionsbereich eine bestimmte Struktur voraus, die Struktur einer n -dimensionalen r -glatten Mannigfaltigkeit. Mit dem Studium der glatten Funktionen wird das Muster unseres axiomatischen Aufbaus der Analysis fortgesetzt. Wir haben früher gesehen: Der Begriff der (halb-)stetigen Funktion bezieht sich auf topologische Räume. Metrische Räume sind die Definitionsbereiche der lokal gleichmäßigstetigen Funktionen. Messbare Räume sind die Definitionsbereiche der messbaren Funktionen.

Zu jedem der genannten Raumtypen gehört ein bestimmter Typ von Morphismen. Bei den Mannigfaltigkeiten sind es die glatten Abbildungen; es wird gefordert, dass der Pullback die glatten Funktionen auf dem Zielraum in glatte Funktionen auf dem Urbildraum transformiert. Eine Neuigkeit gegenüber den einfacheren Raumtypen sind für uns die Koordinatisierungen. Hier kommen Begriffe der Linearen Algebra ins Spiel. Eine n -dimensionale Mannigfaltigkeit trägt in jedem ihrer Punkte P einen Tangentialraum T_P und den dazu dualen Cotangentialraum T_P^* . Eine glatte Abbildung der Mannigfaltigkeit φ liefert eine lineare Abbildung $\varphi_*|_P$ von T_P in $T_{\varphi(P)}$ ('Pushforward').

Jede lokale Koordinatisierung bei P $\{x^j : j \in J\}$ liefert in jedem Punkt bei P ein Paar zueinander dualer Basen von T_P^* und T_P , nämlich $\{dx^j : j \in J\}$ und $\{\frac{\partial}{\partial x^j} : j \in J\}$. Diese Basen sind geeignet, jedes Tangentialvektorfeld bei P mit Hilfe eines J -Tupels von Funktionen (als J -Spalte notiert) darzustellen: $\mathbf{v} = \sum \frac{\partial}{\partial x^j} \cdot v^j$. Tangentialvektorfelder haben wir als Herausforderung verstanden, Differentialgleichungen zu 'lösen'. Dabei geht es darum, einen Überblick zu gewinnen über die Menge der Integralkurven zum Vektorfeld. Das Vektorfeld liefert einen lokalen Fluss.

Eine ganz andere Rolle spielen die Cotangentialfelder, auch Pfaff'sche Formen oder 1-Formen genannt. Der Pullback einer glatten Abbildung bildet k -Formen auf der Zielmannigfaltigkeit in k -Formen auf der Urbildmannigfaltigkeit ab; und diese Abbildung ist verträglich mit der 'Corand-Abbildung' d . Insbesondere gilt für die glatten Funktionen f, g , die man als 0-Formen versteht $\varphi^*(g \cdot df) = \varphi^*g \cdot d(\varphi^*f)$.

Sowohl bei den Vektorfeldern (Differentialgleichungen) als auch bei den Differentialformen sind wir nun an die Grenzen der Kenntnisse gelangt, die man bei den Studierenden im zweiten Studienjahr voraussetzen kann. Wir müssen weiter ausholen, wenn wir tiefer in die Geometrie der Mannigfaltigkeiten eindringen wollen. Ein nächstes Ziel sollte dann sein, den berühmten Satz von Stokes zu verstehen und zu beweisen. Wir wollen hier noch an zwei Beispielen einen Eindruck vermitteln, um was es dabei geht.

Der Satz von Stokes kann in einer Formel von fast magischer Einfachheit ausgedrückt werden: $\int_{\partial \mathfrak{B}} \omega = \int_{\mathfrak{B}} d\omega$. In Worten: Das Integral einer Differentialform ω über

den Rand eines Bereichs ist gleich dem Integral der abgeleiteten Differentialform $d\omega$ über eben diesen Bereich.

Wir müssen eine Reihe von Begriffen klären: Was ist ein Bereich \mathfrak{B} und was ist sein Rand $\partial\mathfrak{B}$? Was ist eine k -Form ω und was ist ihr Corand $d\omega$? Und schliesslich: Was bedeutet in diesem Zusammenhang Integration? Zur letzten Frage haben wir bei anderer Gelegenheit eine erste Antwort gegeben. Wir haben Kurvenintegrale diskutiert: Wenn ω eine Pfaff'sche Form der einfachen Form $f dg$ ist und $\mathfrak{C} = \{\gamma(t) : t \in [a, b]\}$ eine glatt parametrisierte Kurve, dann liefern die Zerteilungen des Parameterintervalls $a = t_0 < t_1 < \dots < t_N = b$ Näherungswerte für das Kurvenintegral

$$\int_{\mathfrak{C}} f dg = \lim \sum_1^N f(\gamma(t_{n-1})) \cdot (g(\gamma(t_n)) - g(\gamma(t_{n-1}))) = \int_a^b f(\gamma(t)) \langle dg, \dot{\gamma}(t) \rangle dt.$$

Der Limes ist über Verfeinerungen zu erstrecken. $\langle dg, \dot{\gamma}(t) \rangle$ ist wie üblich der infinitesimale Zuwachs der Funktion g beim Durchlaufen der Kurve. Es sollte klar sein, dass eine glatte Umparametrisierung der Kurve das Kurvenintegral der Form nicht verändert. Das Kurvenintegral ist auch gleich $\int_a^b F(t) dG(t)$, wo F und G die durch die Kurve zurückgenommenen Funktionen sind: $F = \gamma^*f$, $G = \gamma^*g$.

Beispiel 1.7.1. Es sei $f(x)$ eine stetige nichtnegative Funktion über dem Intervall $[a, b]$ und \mathfrak{B} der 'Bereich unter der Kurve': $\mathfrak{B} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : a \leq x \leq b; y \leq f(x) \right\}$. Der Rand $\partial\mathfrak{B}$ setzt sich aus vier Kurvenstücken zusammen: Wir gehen zuerst entlang der Abszissenachse von $(a, 0)$ nach $(b, 0)$, dann 'senkrecht' weiter bis $(b, f(b))$, sodann entlang des Funktionsgraphen nach $(a, f(a))$ und schliesslich senkrecht zurück zum Ausgangspunkt.

Das Kurvenintegral der Pfaff'schen Form $\omega = -y dx$ über den Rand $\partial\mathfrak{B}$ liefert nur auf dem dritten Teilstück des Randes einen Beitrag, und zwar die Fläche unter der Kurve $\int_{\mathfrak{B}} dx \wedge dy$. Und das ist es, was der Satz von Stokes sagt, denn für $\omega = -y dx$ haben wir $d\omega = dx \wedge dy$.

Beispiel 1.7.2. Gegeben sei eine Pfaff'sche Form $\omega = a(x, y) dx + b(x, y) dy$ mit a, b stetig differenzierbar in einer Umgebung eines achsenparalleles Rechtecks $\mathfrak{B} = [x^l, x^r] \times [y^l, y^r]$. Der Rand setzt sich aus vier achsenparallelen Kurven zusammen $\partial\mathfrak{B} = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$, die wir im positiven Sinn durchlaufen. Wir fassen die einander gegenüberliegenden in entgegengesetzten Richtungen zu durchlaufenden Kurven zusammenfassen.

$$\int_{\gamma_1 + \gamma_3} \omega = \int_{\gamma_1 + \gamma_3} a dx = \int_{x^l}^{x^r} (a(x, y^l) - a(x, y^r)) dx = \iint \left(-\frac{\partial a}{\partial y} \right) (x, y) dx \wedge dy$$

$$\int_{\gamma_2 + \gamma_4} \omega = \int_{\gamma_2 + \gamma_4} b dy = \int_{x^l}^{x^r} (b(x^r, y) - b(x^l, y)) dy = \iint \left(+\frac{\partial b}{\partial x} \right) (x, y) dx \wedge dy.$$

$$\int_{\partial\mathfrak{B}} \omega = \int_{\mathfrak{B}} d\omega \quad \text{mit} \quad d\omega = a_y dy \wedge dx + b_x dx \wedge dy = \left(-\frac{\partial a}{\partial y} + \frac{\partial b}{\partial x} \right) dx \wedge dy.$$

Didaktische Anmerkung

Die Differentialgleichungen sind ein klassisches Terrain für Rechenaufgaben. Die Freude am Errechnen ‘expliziter’ Lösungen haben wir in die Übungsstunden verschoben. In den Vorlesungen haben wir die Begriffsklärung in den Vordergrund gestellt.

2 Einschlägige Themen der Linearen Mathematik

Bei oberflächlicher Betrachtung scheint der nächste Unterabschnitt anzuschließen an die Thematik der dynamischen Systeme. Wir werden damit aber nicht tiefer eindringen in die Begriffswelt der Mannigfaltigkeiten. Zuerst deuten wir Fragen um lineare Differentialgleichungen an, die sich nicht auch Anfangswertaufgaben reduzieren. Im weiteren behandeln wir linearisierte dynamische Systeme mit dem Matrizenkalkül.

2.1 Lineare Differentialgleichungen

Die Gleichung des freien harmonischen Oszillators ist eine skalare homogene lineare Differentialgleichung zweiter Ordnung mit konstanten Koeffizienten

$$m\ddot{x} + r \cdot \dot{x} + kx = 0 \quad \text{für } t \in \mathbb{R} \text{ .}$$

Homogene lineare Differentialgleichungen zweiter Ordnung spielen in vielen Bereichen der mathematischen Physik eine zentrale Rolle, und zwar nicht nur die allbekannten partiellen Differentialgleichungen (PDE), sondern auch eine ganze Reihe skalarer gewöhnlicher Differentialgleichungen (ODE). Diese haben allerdings in der Regel nichtkonstante Koeffizienten. Es geht also um Gleichungen der Form

$$a_2(x) y'' + a_1(x) y' + a_0(x) y = 0.$$

Als Beispiele erwähnen wir die eindimensionale Schrödingergleichung, die Besselsche Differentialgleichung, die Legendresche und die hypergeometrische Differentialgleichung

$$\begin{aligned} y''(x) + (E - U(x)) y(x) &= 0, \\ x^2 y'' + x y' + (x^2 - n^2)y &= 0, \\ (1 - x^2) y'' - 2x y' + n(n + 1)y &= 0, \\ x(1 - x) y'' + [\gamma - (\alpha + \beta + 1)x] y' - \alpha\beta y &= 0. \end{aligned}$$

Diesen Gleichungen ist gemeinsam, dass der Lösungsraum ein zweidimensionaler Vektorraum von Funktionen ist. Eine 'Lösung' ist (jedenfalls lokal) eindeutig bestimmt, wenn man den Wert und die Ableitung in einem Punkt kennt; das ergibt sich aus dem Satz von Picard-Lindelöf (oder einer passenden Variante dieses Satzes).

Die Definitionsbereiche der Funktionen verdienen eingehende Beachtung. Grundsätzlich ist voranzusetzen, dass $a_2(x)$ im Bereich nicht verschwindet. Im Fall der eindimensionalen Schrödingergleichung sucht man nach Lösungen, die im Unendlichen verschwinden. Die Bessel'sche Gleichung untersucht man auf der positiven Halbachse \mathbb{R}_+ ; (man nennt die Variable gerne auch r .) Die Legendresche wird auf dem Intervall $(-1, +1)$ untersucht. Wir bemerken aber, dass sich diese Gleichung durch die Variablensubstitution $x = 1 - 2z$

als ein Spezialfall der hypergeometrischen Gleichung erweist; ($\gamma = 1$, $\alpha = n+1$, $\beta = -n$). Und bei der Theorie dieser Gleichung versteht man z als eine komplexe Variable.

Es ist bemerkenswert, dass sich Lösungen der hypergeometrischen Gleichung durch einen Potenzreihenansatz gewinnen lassen; man gewinnt in der Tat (für $\gamma \notin -\mathbb{Z}$) eine partikuläre Lösung durch eine Potenzreihe, die im Einheitskreis konvergiert. Wir erwähnen das, um an die ursprüngliche Idee von Newton zu erinnern, Differentialgleichungen durch Koeffizientenvergleich in einem Potenzreihenansatz zu 'lösen', sozusagen als eine Anfangswertaufgabe. Newtons Idee trägt aber nicht weit. Für die heutige Theorie der angegebenen Gleichungen ist die Potenzreihendarstellung einer partikulären Lösung nicht der zentrale Punkt. Die moderne Theorie geht ganz andere Wege, man interessiert sich z. B. für das Randverhalten passender Lösungen bzw. ihr globales Verhalten. Diese Fragen und Wege sollen hier aber nicht unser Thema sein.

Wir orientieren uns bei den weiteren Erörterungen dieses Abschnitts an der Vorstellungswelt des Oszillators. In dieser Vorstellungswelt ist es (ebenso wie bei den dynamischen Systemen) angemessen, die unabhängige Variable als die Zeit zu interpretieren, die von $-\infty$ nach $+\infty$ läuft.

Bei den linearen Differentialgleichungen, die wir hier näher studieren wollen, betrachtet man neben den 'homogenen' Gleichungen, deren Lösungsraum ein n -dimensionaler Vektorraum ist, auch sog. inhomogene Gleichungen; sie entstehen aus der homogenen Gleichung dadurch, dass man die 0 auf der rechten Seite durch eine lokal integrierbare Funktion $f(t)$ ersetzt. Der Lösungsraum ist da ein affiner Raum. Die allgemeine Lösung erhält man, indem man zu einer speziellen ('partikulären') Lösung der inhomogenen Gleichung die allgemeine Lösung der homogenen Gleichung dazuaddiert.

Komplexifizierung Beim Studium der linearen Differentialgleichungen empfiehlt es sich u. U., ins Komplexe zu gehen. Wir erläutern das zunächst an einfachen Fällen:

Sprechweise 2.1.1. Eine skalare lineare Differentialgleichung mit konstanten Koeffizienten von der Ordnung p ist eine Gleichung der Form

$$a_p \cdot x^{(p)}(t) + a_{p-1} \cdot x^{(p-1)}(t) + \dots + a_1 \cdot \dot{x}(t) + a_0 \cdot x(t) = b(t) \quad .$$

Die a_j sind hier (reelle oder komplexe) Zahlen mit $a_p \neq 0$. Die homogene Gleichung heisst die Gleichung zum charakteristischen Polynom $a_p \cdot z^p + a_{p-1}z^{p-1} + \dots + a_1 \cdot z + a_0$. Die rechte Seite ist eine vorgegebene lokal integrierbare Funktion.

Eine partikuläre Lösung ist eine komplexwertige Funktion $\{z(t) : t \in \mathbb{R}\}$ mit der angegebenen Beziehung zwischen den Ableitungen, (wobei die $(p-1)$ -te Ableitung totalstetig ist).

Beispiel 2.1.1 (Der harmonische Oszillator).

Wir behandeln die Gleichung des harmonischen Oszillators als ein konkretes Beispiel. Für diesen speziellen Fall werden wir eine Basis des Lösungsraums der homogenen Gleichung explizit angeben. Die Diskussion der inhomogenen Gleichung stellen wir vorerst noch zurück.

Mit den Parametern $\gamma = r/m$; $\omega_0^2 = k/m$ wird aus der homogenen Gleichung

$$\ddot{x} + \gamma \cdot \dot{x} + \omega_0^2 \cdot x = 0,$$

und da unterscheiden wir nun zwei Fälle, den ‘schwach gedämpften’ Fall zu $\omega_0 > \frac{1}{2}\gamma$ und den ‘stark gedämpften’ Fall zu $\omega_0 \leq \frac{1}{2}\gamma$. Zwar führt der Ansatz $x(t) = \tilde{x} \cdot e^{i\omega t}$ in beiden Fällen zum Ziel; die Lösungen haben aber ein recht verschiedenes Aussehen.

Den stark gedämpften Fall diskutieren wir später.

Im schwach gedämpften Fall besitzt das charakteristische Polynom zwei zu einander komplex konjugierte Nullstellen. Und daraus ergeben sich zwei linear unabhängige Lösungen der homogenen Differentialgleichung, nämlich mit $\tilde{\omega} = \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}$

$$x_+(t) = \exp(-\frac{1}{2}\gamma \cdot t) \cdot \exp(i\tilde{\omega}t), \quad x_-(t) = \exp(-\frac{1}{2}\gamma \cdot t) \cdot \exp(-i\tilde{\omega}t).$$

Wenn $x(t)$ eine komplexwertige Lösung ist, dann sind Real- und Imaginärteil reellwertige Lösungen. Beliebige Präsentationsweisen für die reellwertigen Lösungen sind

$$x(t) = \exp(-\frac{1}{2}\gamma \cdot t) \cdot \Re(A \cdot \exp(i\tilde{\omega}t)) = \exp(-\frac{1}{2}\gamma \cdot t) \cdot |A| \cdot \cos(\tilde{\omega}(t - t_0))$$

Satz 2.1.1.

Gegeben sei eine homogene lineare Differentialgleichung mit konstanten Koeffizienten.

$$a_p \cdot z^{(p)}(t) + a_{p-1} \cdot z^{(p-1)}(t) + \dots + a_1 \cdot \dot{z}(t) + a_0 \cdot z(t) = 0$$

Ist λ eine Nullstelle des charakteristischen Polynoms ist, so ist die Funktion $z(t) = e^{\lambda t}$ eine Lösung. Ist λ eine doppelte Nullstelle, so ist auch $t \cdot e^{\lambda t}$ eine Lösung. Ist λ eine ℓ -fache Nullstelle ist, so sind die Funktionen

$$e^{\lambda t}, \quad t \cdot e^{\lambda t}, \quad t^2 \cdot e^{\lambda t}, \quad \dots, \quad t^{\ell-1} \cdot e^{\lambda t}$$

linear unabhängige Lösungen der homogenen Gleichung.

Sind λ_k die Nullstellen des charakteristischen Polynoms, ℓ_k die Vielfachheit von λ_k , dann sind die Lösungen der homogenen Gleichung die Funktionen der Gestalt

$$x(t) = \sum_k p_k(t) \cdot e^{\lambda_k \cdot t} \quad \text{mit Polynomen } p_k(t) \text{ vom Grad } \leq \ell_k - 1.$$

Beweis. Die erste Behauptung ergibt sich aus einer trivialen Rechnung.

Die folgende Überlegung von Euler macht die zweite Aussage plausibel: Wenn λ_1 und λ_2 ($\lambda_1 \neq \lambda_2$) nahe aneinander liegende Nullstellen sind, dann sind

$$e^{\lambda_1 \cdot t} \text{ und } \frac{1}{\lambda_2 - \lambda_1} [e^{\lambda_1 \cdot t} - e^{\lambda_2 \cdot t}]$$

linear unabhängige Lösungen. Im Grenzfall $\lambda_1 = \lambda_2$ wird aus der zweiten Lösung $t \cdot e^{\lambda t}$. Man kann aber auch nachrechnen

$$\begin{aligned} (t \cdot e^{\lambda t})' &= e^{\lambda t} + t \cdot \lambda e^{\lambda t}; & (te^{\lambda t})'' &= [2\lambda + t \cdot \lambda^2] e^{\lambda t}, \dots \\ & \mathbf{a}_p(t \cdot e^{\lambda t})^{(p)} + \dots + \mathbf{a}_1 \cdot (te^{\lambda t})' + \mathbf{a}_0 \cdot e^{\lambda t} &= \\ &= e^{\lambda t} [\mathbf{a}_p \cdot p \cdot \lambda^{p-1} + \dots + \mathbf{a}_1 \cdot \lambda] + te^{\lambda t} [\mathbf{a}_p \cdot \lambda^p + \dots + \mathbf{a}_1 \lambda + \mathbf{a}_0] \quad . \end{aligned}$$

Die erste Klammer ist die Ableitung des charakteristischen Polynoms im Punkt λ ; diese verschwindet, wenn λ eine doppelte Nullstelle ist.

Auf dieselbe Weise ergibt sich der Beweis im Falle einer ℓ -fachen Nullstelle.

Die letzte Behauptung ergibt sich folgendermaßen. Die angegebenen Funktionen erfüllen die homogene Gleichung. Sie bilden einen p -dimensionalen Vektorraum wegen $\sum_k \ell_k = p$. Wir werden uns später (aufgrund sehr viel allgemeinerer Überlegungen) überzeugen, dass die Lösungsmenge in dem hier vorliegenden komplexen(!) Fall ein p -dimensionaler komplexer Vektorraum ist.

Wir haben gesehen: Für die Charakterisierung des Lösungsraums ist der Fundamentalsatz der Algebra ein bequemes Hilfsmittel. Wenn nun aber die Koeffizienten der Gleichung reell sind, dann interessiert man sich natürlich primär für den reell- p -dimensionalen Vektorraum der reellwertigen Lösungen. Wenn $z(t)$ eine komplexwertige Lösung ist, dann auch die komplex konjugierte Funktion $\bar{z}(t)$. Der Raum der reellwertigen Lösungen besteht aus den Realteilen der komplexwertigen Lösungen. Wir bemerken dazu: Wenn die Koeffizienten des Polynoms reell sind, dann ist mit λ auch $\bar{\lambda}$ eine Wurzel. Zu einem Paar $\lambda \neq \bar{\lambda}$ sind

$$\frac{1}{2} (e^{\lambda t} + e^{\bar{\lambda} t}) \quad \text{und} \quad \frac{1}{2i} (e^{\lambda t} - e^{\bar{\lambda} t})$$

linear unabhängige reelle Lösungen. Wenn $\lambda = \alpha + i\omega$, $\bar{\lambda} = \alpha - i\omega$, dann sind auch

$$e^{\alpha t} \cdot \cos(\omega t) \quad \text{und} \quad e^{\alpha t} \cdot \sin(\omega t)$$

linear unabhängige Lösungen. Allgemein: Wenn λ eine nichtreelle ℓ -fache Nullstelle eines reellen charakteristischen Polynoms ist, dann sind

$$\begin{aligned} e^{\alpha t} \cdot \cos \omega t, & \quad t \cdot e^{\alpha t} \cdot \cos \omega t, \dots, & t^{\ell-1} \cdot e^{\alpha t} \cdot \cos \omega t; \\ e^{\alpha t} \cdot \sin \omega t, & \quad t \cdot e^{\alpha t} \cdot \sin \omega t, \dots, & t^{\ell-1} \cdot e^{\alpha t} \cdot \sin \omega t \end{aligned}$$

linear unabhängige reelle Lösungen der homogenen Gleichung.

Auf diese Weise gewinnt man alle reellen Lösungen zu einem charakteristischem Polynom mit reellen Koeffizienten.

Gegeben sei eine homogene lineare Differentialgleichung mit konstanten Koeffizienten

$$x^{(p)} + \mathbf{a}_{p-1} \cdot x^{(p-1)} + \dots + \mathbf{a}_1 \cdot \dot{x} + \mathbf{a}_0 \cdot x = 0.$$

Das p -Tupel $\mathbf{y}(t)$ löst das Gleichungssystem

$$\dot{\mathbf{y}} = \mathbf{A} \cdot \mathbf{y} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \\ -\mathbf{a}_0 & -\mathbf{a}_1 & -\mathbf{a}_2 & \dots & -\mathbf{a}_{p-1} \end{pmatrix}$$

genau dann, wenn der erste Eintrag $\mathbf{y}(t) = \mathbf{x}(t)$ die gegebene skalare Gleichung löst.

Wir beschränken uns daher von nun an auf lineare Systeme erster Ordnung.

Systeme erster Ordnung

Im Falle der konstanten Koeffizienten ist eine homogene lineare Differentialgleichung durch eine einzige Matrix \mathbf{A} gegeben; und jede quadratische Matrix \mathbf{A} kommt in Betracht.

$$\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x}.$$

Im Matrizenkalkül kann man die Lösungen sehr einfach beschreiben.

Satz 2.1.2.

Sei \mathbf{A} eine beliebige $n \times n$ -Matrix, $\mathbf{x}_0 \in \mathbb{C}_{\text{Sp}}^n$. Die (eindeutig bestimmte!) Lösung der Anfangswertaufgabe $\dot{\mathbf{x}} = \mathbf{A} \cdot \mathbf{x}$, $\mathbf{x}(0) = \mathbf{x}_0$ ist die zeitabhängige Spalte

$$\mathbf{x}(t) = \exp(t\mathbf{A}) \cdot \mathbf{x}_0 \quad \text{für alle } t \in \mathbb{R}$$

Beweis. Nach dem Satz von Picard-Lindelöf gibt es genau eine Lösung der AWA. Unser $\mathbf{x}(t)$ ist eine Lösung; denn

$$\lim_{h \rightarrow 0} \frac{1}{h} [\mathbf{x}(t+h) - \mathbf{x}(t)] = \lim_{h \rightarrow 0} \frac{1}{h} [\exp(t+h)\mathbf{A} - \exp t\mathbf{A}] \cdot \mathbf{x}_0 = \mathbf{A} \cdot \exp t\mathbf{A} \cdot \mathbf{x}_0 = \mathbf{A} \cdot \mathbf{x}(t) \quad .$$

Es handelt sich sogar um eine globale Lösung. Durch jeden Punkt (t_0, \mathbf{x}_0) im „erweiterten Phasenraum“ $\mathbb{R} \times \mathbb{R}^n$ gibt es genau eine Kurve, welche die Differentialgleichung löst.

Variation der Konstanten Wir haben gefunden: Jede konstante Spalte \mathbf{c} liefert eine Lösung der homogenen Gleichung $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, nämlich $\mathbf{x}(t) = \exp(t \cdot \mathbf{A}) \cdot \mathbf{c}$.

Wir betrachten die inhomogene Gleichung zu einer Spalte $\mathbf{b}(t)$, die in lokal Lebesgue-integrierbarer Weise von der Zeit abhängt.

$$\dot{\mathbf{x}}(t) = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{b}(t) \quad .$$

Diese inhomogene Gleichung „löst“ man mittels eines Tricks, der auf Lagrange zurückgeht. Der Trick heißt die Variation der Konstanten. Man macht den Ansatz

$$\mathbf{x}(t) = \exp(t \cdot \mathbf{A}) \cdot \mathbf{c}(t) \quad .$$

Die Produktregel liefert eine Differentialgleichung für $c(\cdot)$

$$\dot{x}(t) = \exp(t \cdot A) \cdot \dot{c}(t) + A \cdot \exp(t \cdot A) \cdot c(t) = \exp(t \cdot A) \cdot \dot{c}(t) + A \cdot x(t) \quad .$$

Die inhomogene Gleichung ist erfüllt, wenn

$$c(t) = \text{const} + \int^t \exp(-s \cdot A) \cdot b(s) ds.$$

Somit ergibt sich also $c(\cdot)$ durch eine „Quadratur“, wenn man die Schar der Matrizen $\exp(-t \cdot A)$ bereits hat. $c(\cdot)$ ist absolutstetig.

Beispiel 2.1.2. Für reelle Zahlen $\alpha \neq \lambda$ betrachten wir die inhomogene Gleichung

$$\dot{x} = \alpha \cdot x + e^{\lambda t}.$$

Die Lösungen der homogenen Gleichungen sind $x(t) = e^{\alpha t} \cdot \text{const.}$

Die Lösungen der inhomogenen Gleichungen sind $x(t) = e^{\alpha t} \cdot \text{const} + \frac{1}{\lambda - \alpha} e^{\lambda t}.$

(Den Fall $\lambda = \alpha$ überlassen wir dem Leser.)

Koordinatenwechsel

Satz 2.1.3.

Sei A irgendeine $n \times n$ -Matrix und B eine invertierbare $n \times n$ -Matrix.

Die zeitabhängige Spalte $x(t) = B \cdot y(t)$ löst die homogene lineare Gleichung $\dot{x} = A \cdot x$ genau dann, wenn $y(\cdot)$ die transformierte Gleichung löst $\dot{y} = B^{-1} \cdot A \cdot B \cdot y.$

Beweis. $\dot{y} = B^{-1} \cdot \dot{x} = B^{-1} \cdot A \cdot x = B^{-1} \cdot A \cdot B y.$

Satz 2.1.4. Sei λ ein Eigenwert der Matrix A und \tilde{x}_λ ein Eigenvektor dazu

$$A \tilde{x}_\lambda = \lambda \cdot \tilde{x}_\lambda \quad , \quad \tilde{x}_\lambda \neq 0 \quad .$$

Dann ist $x(t) = e^{\lambda t} \cdot \tilde{x}_\lambda$ eine Lösung der Gleichung $\dot{x}(t) = A \cdot x(t).$

Satz 2.1.5.

Sei A eine $n \times n$ -Matrix, deren Eigenvektoren $\tilde{x}_1, \dots, \tilde{x}_n$ (zu den Eigenwerten $\lambda_1, \dots, \lambda_n$) den vollen n -dimensionalen Spaltenraum aufspannen. Dann hat jede Lösung der Gleichung $\dot{x} = Ax$ die Form

$$x(t) = \sum \alpha_j \cdot e^{\lambda_j t} \tilde{x}_j \quad , \quad \alpha_j \in \mathbb{C} \quad \text{beliebig.}$$

Beweis.

Alle diese Linearkombinationen sind Lösungen. Nach dem Satz von Picard-Lindelöf gibt es zu jedem Anfangswert genau eine Lösung. Jeden Anfangswert kann man durch eine Linearkombination erhalten.

Jordanblöcke Es gibt nun $n \times n$ -Matrizen, deren Eigenvektoren nicht den gesamten Spaltenraum aufspannen. Beispiele sind die sog. Jordanblöcke.

Der Jordan-Block von der Länge k zum Eigenwert λ ist die obere Dreiecksmatrix, die in der Diagonalen den Eintrag λ und in der ersten Nebendiagonalen den Eintrag 1 hat (und sonst lauter Nullen). Für $k = 4$

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} = \lambda \cdot E + N .$$

Die Matrix N hat die Eigenschaft $N^4 = 0$, während $N \neq 0$, $N^2 \neq 0$, $N^3 \neq 0$. N^k hat nämlich in der k -ten Nebendiagonale den Eintrag = 1 und sonst Nullen.

Satz

Die Spalte $x(t)$ löst die Gleichung $\dot{x} = (\lambda E + N)x$ genau dann, wenn

$$x(t) = e^{\lambda t} \cdot y(t) \quad \text{mit einem } y(t) \text{ sodass } \dot{y} = Ny.$$

Die $y(t)$ haben die Gestalt

$$y(t) = e^{tN} \cdot y_0 = \left(I + tN + \frac{1}{2!}t^2N^2 + \dots + \frac{1}{n!}t^nN^n \right) \cdot y_0 .$$

Der Beweis ist eine leichte Rechnung.

Satz 2.1.6. „Jordan'sche Normalform“

Zu jeder $n \times n$ -Matrix A existiert eine ähnliche Matrix

$$\tilde{A} = B^{-1}AB ,$$

welche aus Jordanblöcken besteht.

Die Längen der Jordanblöcke zum Eigenwert λ summieren sich zur Vielfachheit dieses Eigenwerts. Es handelt sich um einen berühmten Satz der linearen Algebra, den wir hier nicht beweisen.

Beispiel 2.1.3 (Der harmonische Oszillator mit starker Reibung).

Wenn wir passend skalieren, dann lautet die homogene Schwingungsgleichung zum gedämpften Oszillator $\ddot{x} + r\dot{x} + x = 0$ oder in Matrixform

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - rx_2 \end{aligned} \quad \frac{d}{dt} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -r \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} .$$

Die Summe der Eigenwerte ist $= -r$, das Produkt ist $= 1$. Die charakteristische Gleichung

$$0 = \det(A - \lambda I) = \lambda^2 + r\lambda + 1 = 0$$

hat für $r > 2$ zwei verschiedene (negative!) Wurzeln. Die Matrix A lässt sich diagonalisieren. Im Falle $r = 2$ ist $\lambda = -1$ doppelte Nullstelle und die Kurven der Auslenkung haben die Gestalt $\mathbf{x}(t) = (\alpha + \beta t)e^{-t}$; sie gehen höchstens einmal durch die Nulllage.

Lineare Differentialgleichungen mit stetigen Koeffizienten

Einige der Ideen, die wir eben für lineare Differentialgleichungen mit konstanten Koeffizienten entwickelt haben, lassen sich verallgemeinern.

Satz 2.1.7 (Unendliche Zukunft).

Sei $\{A(t) : t \in \mathbb{R}\}$ eine stetige Schar von $n \times n$ -Matrizen und $\{b(t) : t \in \mathbb{R}\}$ eine stetige Schar von n -Spalten. Zu jedem $(t_0, \mathbf{x}_0) \in \mathbb{R} \times \mathbb{R}^n$ existiert genau eine Lösung der Anfangswertaufgabe

$$\dot{\mathbf{y}}(t) = A(t) \cdot \mathbf{y}(t) + \mathbf{b}(t) \quad \text{für alle } t \in \mathbb{R} .$$

Beweis. Der Satz von Picard-Lindelöf garantiert lokal die Existenz und die Eindeutigkeit. Die folgende Abschätzung zeigt, dass die Lösungskurven nicht in endlicher Zeit explodieren können. Sei $r(t)$ der euklidische Abstand einer Lösungskurve vom Koordinatenursprung

$$r(t) = \sqrt{\mathbf{y}^T(t) \cdot \mathbf{y}(t)} .$$

Wir zeigen, dass in jedem endlichen Intervall gilt:

$$\dot{r}(t) \leq \text{Const} \cdot r(t) + \text{Const} .$$

Daraus ergibt sich, dass $\|\mathbf{x}(t)\|$ nicht in endlicher Zeit nach ∞ streben kann. In der Tat gilt

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} r^2(t) \right) &= \frac{d}{dt} \left(\frac{1}{2} \mathbf{y}^T(t) \cdot \mathbf{y}(t) \right) = \mathbf{y}^T(t) \cdot \dot{\mathbf{y}}(t) = \\ &= \mathbf{y}^T(t) \cdot (A(t) \cdot \mathbf{y}(t) + \mathbf{b}(t)) \\ \dot{r}(t) &= \frac{\mathbf{y}^T(t)}{r(t)} \cdot (A(t) \cdot \mathbf{y}(t) + \mathbf{b}(t)) . \end{aligned}$$

Nach der Schwarz'schen Ungleichung

$$\|\dot{r}(t)\| \leq \|A(t)\mathbf{y}(t) + \mathbf{b}(t)\| \leq \text{Const} \cdot r(t) + \text{Const} .$$

Die Lösungsmatrizen

Der Raum aller Lösungen der homogenen Gleichung ist ein Vektorraum.

Die Auswertung in einem beliebigen s ist eine Basis.

Der Basiswechsel von $\mathbf{y}(s)$ nach $\mathbf{y}(t)$ ist durch eine Matrix Y_s^t gegeben. $Y_s^t \mathbf{y}_0$ ist der Punkt, in welchem sich diejenige Lösungskurve zur Zeit t befindet, welche zur Zeit s im Punkt $\mathbf{y}(s) = \mathbf{y}_0$ startet.

$$\mathbf{y}(t) = \mathbf{y}(s) + \int_s^t A(u) \cdot \mathbf{y}(u) \, du = Y_s^t \cdot \mathbf{y}_0.$$

Satz 2.1.8.

Für die Schar der Matrizen $\{Y_s^t : s, t \in \mathbb{R}\}$ gilt

$$(i) \quad Y_t^t = I \quad (\text{Einheitsmatrix}) \quad \text{für alle } t$$

$$(ii) \quad Y_s^t = Y_u^t \cdot Y_s^u \quad \text{für alle } t_0, t_1, t_2$$

$$(iii) \quad \frac{d}{dt} Y_s^t = \lim_{h \rightarrow 0} \frac{1}{h} (Y_t^{t+h} - I) \cdot Y_s^t = A(t) \cdot Y_s^t$$

Beweis. von (iii): Die j -te Spalte von Y_t^{t+h} erfüllt die Gleichung

$$\begin{aligned} \frac{1}{h} [y_j(t+h) - y_j(t)] &= \frac{1}{h} \int_0^h A(t+s) \cdot y_j(t+s) ds \\ \frac{d}{dt} Y_s^t &= \lim_{h \rightarrow 0} \frac{1}{h} (Y_s^{t+h} - Y_s^t) = \lim_{h \rightarrow 0} \frac{1}{h} [Y_t^{t+h} - I] Y_s^t = A(t) \cdot Y_s^t . \end{aligned}$$

Einparametrische Gruppen von Bijektionen

Eine Schar von Bijektionen einer Grundmenge $\{\varphi^t : t \in \mathbb{R}\}$ nennt man eine einparametrische Gruppe von Bijektionen, wenn gilt

$$\varphi^0 = \text{Identität} , \quad \varphi^t \circ \varphi^s(\cdot) = \varphi^{t+s}(\cdot) \quad \text{für alle } s, t \in \mathbb{R} .$$

Eine Schar von Bijektionen $\{\varphi_t^s : s, t \in \mathbb{R}\}$ nennt man eine verallgemeinerte einparametrische Gruppe von Bijektionen, wenn gilt

$$\begin{aligned} \varphi_t^t &= \text{Id} \quad \text{für alle } t ; \\ \varphi_t^u \circ \varphi_s^t &= \varphi_s^u \quad \text{für alle } s, t, u . \end{aligned}$$

Unsere Schar $\{Y_s^t : s, t \in \mathbb{R}\}$ ist also eine verallgemeinerte einparametrische Gruppe von $n \times n$ -Matrizen, (aufgefasst als lineare Abbildungen des Spaltenraums.) Die Schar ist stetig differenzierbar. Die stetige Schar $\{A(t) : t \in \mathbb{R}\}$ wird der infinitesimale Generator dieser verallgemeinerten einparametrischen Gruppe genannt. Eine einparametrische Gruppe erhalten wir, wenn $A(\cdot) \equiv A$ konstant ist. In diesem Falle gilt

$$Y^s = Y_0^s = \exp(s \cdot A) \quad \text{für alle } s \in \mathbb{R} .$$

Ein nicht ganz so spezieller Fall eines infinitesimalen Generators ist der, wo die $A(t)$ paarweise kommutieren

$$A(s) \cdot A(t) = A(t) \cdot A(s) \quad \text{für alle } s, t .$$

In diesem Falle kommutieren auch die Lösungsmatrizen und es gilt

$$Y_t^s = \exp\left(\int_s^t A(u) du\right) .$$

In der Tat: Wenn die $A(t)$ paarweise kommutieren, dann kommutieren auch die Integrale und leisten in der Tat $\frac{d}{dt}Y_s^t = A(t) \cdot Y_s^t$; denn

$$\exp\left(\int_s^{t+h} A(u) du\right) = \exp\left(\int_t^{t+h} A(u) du\right) \cdot \exp\left(\int_s^t A(u) du\right)$$

$$\frac{d}{dt} \exp\left(\int_s^t A(u) du\right) = A(t) \cdot \exp\left(\int_s^t A(u) du\right) = A(t) \cdot Y_s^t .$$

Der eindimensionale Fall Die Lösungen $y(\cdot)$ der homogenen skalaren Gleichung

$$\dot{y} = a(t) \cdot y$$

ergeben sich durch eine Quadratur $\tilde{y}(t) = \exp\left(\int_s^t a(u) du\right) \cdot \tilde{y}(s)$ für alle s, t .

Durch eine weitere Quadratur finden wir eine interessante Lösung der inhomogenen Gleichung

$$\dot{y} = a(t) \cdot y + b(t) .$$

Wenn wir nämlich wieder die Variation der Konstanten ansetzen, dann finden wir für die partikuläre Lösung, die zum Zeitpunkt s verschwindet

$$y(t) = \int_s^t g(t, v) b(v) dv \quad \text{mit} \quad g(t, v) = \exp\left(\int_v^t a(u) du\right) .$$

Betrachten wir den Fall, wo die ‘Störung’ $b(\cdot)$ für $v \geq s$ verschwindet. Unsere Partikuläre Lösung diejenige Lösung, die für kleine s verschwindet. Die Lösung ‘erhebt’ sich aus der Nulllösung unter dem Einfluss der einsetzenden Störung. Manchmal studiert man auch den allgemeineren Fall

$$y(t) = \int_{-\infty}^t g(t, v) dB(v) \quad \text{mit} \quad B(\cdot) \text{ von beschränkter Variation.}$$

Satz 2.1.9 (Variation der Konstanten im n -dimensionalen nichtautonomen Fall).

Es sei $\{A(t) : t \in \mathbb{R}\}$ sei eine stetige Schar von $n \times n$ -Matrizen; und $\{b(t) : t \in \mathbb{R}\}$ eine lokal integrable Schar von n -Spalten.

Es existiert dann eine stetig differenzierbare Schar von Matrizen $\{Y_s^t : s, t \in \mathbb{R}\}$ sodass

a) Für jedes (s, \tilde{y}) ist

$$\tilde{y}(t) = Y_s^t \cdot \tilde{y} \quad \text{die Lösung der AWA} \quad \dot{y}(t) = A(t) \cdot y(t), \quad y(s) = \tilde{y} .$$

b) Die Lösung der inhomogenen Gleichung für Zeitpunkte $t \geq s$

$$\dot{y}(t) = A(t) \cdot y(t) + b(t); \quad y(s) = \tilde{y}$$

ergibt sich durch die Quadratur

$$y(t) = Y_s^t \cdot \tilde{y} + \int_s^t Y_s^v \cdot b(v) dv .$$

Beweis von b) („Variation der Konstanten“)

Wir machen den Ansatz

$$\mathbf{y}(t) = \mathbf{Y}_s^t \cdot \mathbf{c}(t) \quad .$$

Die Produktregel liefert eine Differentialgleichung für $\mathbf{c}(\cdot)$:

$$\dot{\mathbf{y}} = \left(\frac{d}{dt} \mathbf{Y}_s^t \right) \cdot \mathbf{c}(t) + \mathbf{Y}_s^t \cdot \dot{\mathbf{c}}(t) = \mathbf{A}(t) \mathbf{Y}_s^t \cdot \mathbf{c}(t) + \mathbf{Y}_s^t \cdot \dot{\mathbf{c}}(t)$$

Die Gleichung ist gelöst, wenn wir $\mathbf{c}(\cdot)$ so finden, dass

$$\mathbf{Y}_s^t \cdot \dot{\mathbf{c}}(t) = \mathbf{b}(t) \quad : \quad \dot{\mathbf{c}}(t) = \mathbf{Y}_t^s \cdot \mathbf{b}(t)$$

$$\mathbf{c}(t) = \tilde{\mathbf{y}} + \int_s^t \mathbf{Y}_v^s \mathbf{b}(v) dv$$

$$\mathbf{y}(t) = \mathbf{Y}_s^t \cdot \mathbf{c}(t) = \mathbf{Y}_s^t \left(\tilde{\mathbf{y}} + \int_s^t \mathbf{Y}_v^s \mathbf{b}(v) dv \right) = \mathbf{Y}_s^t \cdot \tilde{\mathbf{y}} + \int_s^t \mathbf{Y}_v^t \mathbf{b}(v) dv \quad \text{q.e.d.}$$

Die Influenzfunktion.

Die Lösung der inhomogenen Gleichung für $t > s$ (fest) betrachten wir als das Resultat einer Störung, die zu den Zeiten v ($s < v < t$) von $\mathbf{b}(s)$ hervorgerufen wurde. Es handelt sich um eine Überlagerung von Störungen

$$\mathbf{y}(t) = \mathbf{Y}_s^t \cdot \tilde{\mathbf{y}} + \int_s^t \mathbf{Y}_v^t \cdot \mathbf{b}(v) dv = \mathbf{Y}_s^t \cdot \tilde{\mathbf{y}} + \int_s^t \mathbf{G}(t, v) \mathbf{b}(v) dv \quad \text{für } t \geq s \quad .$$

Wie oben haben wir eine Interpretation der (hier matrixwertigen) Funktionen $\mathbf{G}(\cdot, \cdot)$.

$$\mathbf{G}(t, v) = \begin{cases} \mathbf{Y}_v^t & \text{falls } t > v \\ 0 & \text{für } t \leq v \end{cases}$$

Die Funktion $\mathbf{G}(\cdot, \cdot)$ heißt die Influenzfunktion; denn die Störung zum Zeitpunkt wirkt in der Zukunft gemäß $\mathbf{G}(\cdot, \cdot)$

Ein anderer geläufiger Name ist Green's Funktion (oder genauer verzögerte Green's Funktion, *retarded Green's function*) zur Schar $\{\mathbf{A}(t) : t \in \mathbb{R}\}$. Die Physiker sagen, $\mathbf{G}(\cdot, v)\mathbf{b}$ sei die Lösung der Gleichung

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \delta_v \cdot \mathbf{b} \quad \text{mit} \quad \mathbf{y}(t) = 0 \quad \text{für } t < v \quad .$$

Über ein sehr kurzes Zeitintervall bei v ist auf das (bis dahin ruhende) System eine Anregung \mathbf{b} ausgeübt worden. Die Lösung der inhomogenen Gleichung

$$\mathbf{y}(t) = \mathbf{Y}_s^t \cdot \tilde{\mathbf{y}} + \int_s^t \mathbf{G}(t, v) \mathbf{b}(v) dv$$

entsteht aus dem ungestörten $\tilde{\mathbf{y}}(t) = \mathbf{Y}_s^t \tilde{\mathbf{y}}$ durch eine Überlagerung der Lösungen mit δ -Impulsen zu den verschiedenen Zeiten v .

Von δ -Impulsen wollen wir nicht sprechen. Die Überlegung zur verzögerten Green'schen Funktion lässt sich aber jedenfalls problemlos auf den Fall ausdehnen, wo $A(\cdot)$ und $b(\cdot)$ lokal integrierbar sind. Von Interesse ist möglicherweise der Fall, wo $b(\cdot)$ eine Indikatorfunktion oder eine elementare Treppenfunktion ist, ist. Man schreibt dann die lineare Differentialgleichung als Integralgleichung. Gesucht sind nun die totalstetigen $y(\cdot)$ mit

$$y(t'') - y(t') = \int_{t'}^{t''} [A(s) \cdot y(s) + b(s)] ds \quad \text{für alle } (t', t'') .$$

Beispiel 2.1.4 (Das Einschalten einer konstanten Kraft).

$$\dot{y}(t) = y(t) + \begin{cases} 0 & \text{für } t < 0 \\ 1 & \text{für } t > 0 \end{cases}$$

$$\text{Partikuläre Lösung : } y(t) = \begin{cases} 0 & \text{für } t \leq 0 \\ e^t - 1 & \text{für } t \geq 0 . \end{cases}$$

Ergänzung: Nochmals nichtlineare Flüsse

Wir gehen aus von einem autonomen dynamischen System zu einem stetig differenzierbaren Vektorfeld; wir studieren dieses in der Nähe eines stationären Punkts \tilde{P} , ($V(\tilde{P}) = 0$):

$$\dot{\gamma}(t) = V(\gamma(t)) \quad \text{mit} \quad \gamma(0) = P \sim \tilde{P}.$$

Zu diesem Zweck assoziiert man eine lineare Differentialgleichung, das sog. linearisierte System. Zuerst wählen wir eine Koordinatisierung $\{x^j : j \in J\}$ bei \tilde{P} . Es sei nun

$$V(\cdot) = \sum_j v^j(\cdot) \cdot \frac{\partial}{\partial x^j}, \quad dv^j = \sum_k v^j_k \cdot dx^k.$$

Die $J \times J$ -Matrix A mit den Einträgen $v^j_k(\tilde{P})$ verstehen wir als den infinitesimalen Generator eines lineares dynamisches System auf dem Tangentialraum $T_{\tilde{P}}$. Man überlegt sich leicht, dass dieses im Punkt \tilde{P} linearisierte System wohldefiniert ist; wenn man ein anderes Koordinatensystem $\{y^i : i \in I\}$ wählt, dann gelangt man zum gleichen linearen System, (wo jetzt die Tangentialvektoren durch I-Spalten beschrieben sind.)

Wenn alle Eigenwerte der Matrix A negativen Realteil haben, dann nennt man den Punkt einen stabilen stationären Punkt.

Beispiel 2.1.5 (Kleine Oszillationen).

Man spricht von einem natürlichen Lagrange'schen System, wenn die Lagrange-Funktion die Differenz von kinetischer und potentieller Energie ist, wobei die kinetische Energie eine positivdefinite quadratische Form in den Geschwindigkeiten ist:

$$L(q, \dot{q}) = T - U = \frac{1}{2} \sum_{jk} a_{jk} \dot{q}^j \dot{q}^k - U(q).$$

Die Punkte \tilde{q} mit $dU|_{\tilde{q}} = 0$ liefern offenbar stationäre Punkte des Systems; wenn die Geschwindigkeiten zur Zeit 0 allesamt Null sind, dann gibt es keine zeitliche Veränderung.

Die Hamilton-Funktion ist in unserem Falle $H = T + U$, wobei man hier aber davon ausgeht, dass die kinetische durch die Impulse ausgedrückt ist. Das ist aber kein Problem; denn auch in dieser Darstellung erscheint die kinetische Energie als eine positivdefinite quadratische Form.

$$p_j = \frac{\partial}{\partial \dot{q}_j} T = \frac{1}{2} \frac{\partial}{\partial \dot{q}_j} (\dot{q}^T \cdot A \cdot \dot{q}) = (A\dot{q})_j = \sum_k a_{jk} \cdot \dot{q}^k.$$

Wenn man die p_j als Zeile notiert, dann hat man also $T = \frac{1}{2} p \cdot A^{-1} \cdot p^T$. Die Hamiltongleichungen werden zu einem System von $2n$ linearen Differentialgleichungen, wenn man in der Nähe des Gleichgewichtspunkts Punkts \tilde{q} (o. B. d. A mit den Koordinaten $\tilde{q}^j = 0$) die potentielle Energie durch das quadratische Taylorpolynom ersetzt. ('Linearisierung')

$$U(q) = \frac{1}{2} \sum_{jk} b_{jk} q^j q^k + o(\|q\|^2) = \frac{1}{2} q^T \cdot B \cdot q + o(\|q\|^2).$$

In der Linearen Algebra lernt man: Wenn A und B symmetrische Matrizen sind mit A positiv definit, dann kann man die beiden Matrizen simultan auf Diagonalgestalt bringen.

$$U^{-1} \cdot A \cdot U = I \quad (\text{Einheitsmatrix}); \quad U^{-1} \cdot B \cdot U = \tilde{B} \quad (\text{Diagonalmatrix}).$$

Bei der entsprechenden Wahl des Koordinatensystems entsprechend zerfällt das System der Gleichungen, gemäß den Eigenvektoren zur Matrix \tilde{B} .

Wenn man gar nicht zu den Hamilton-Gleichungen übergeht, und gleich bei Lagrange-Gleichungen bleibt, dann bewirkt der Koordinatenwechsel, dass die Bewegungsgleichungen in n unabhängige Gleichungen zweiter Ordnung zerfällt.

$$\ddot{Q}_i = -\lambda_i \cdot Q_i,$$

wobei die reellen Zahlen λ_i die Nullstellen des Polynoms $\det |B - \lambda A|$ sind. Die Einzelschwingungen heissen die 'normal modes' zu den Kreisfrequenzen λ_i . Die Schwingungen des Systems sind Überlagerungen von ungedämpften harmonischen Schwingungen.

Plausibilitätsbetrachtung: Linearisierung zur Behandlung von Störungen.

Wir gehen aus von der Lösung $x(\cdot)$ einer AWA zu einer nichtlinearen Differentialgleichung

$$\dot{x} = V(t, x) \quad x(s) = \tilde{x} .$$

Im gewählten Koordinatensystem ist das Vektorfeld durch eine zeitabhängige Spalte gegeben, von der wir annehmen, dass sie für jedes t stetig differenzierbar ist mit $W(t, \cdot) = \frac{\partial V}{\partial x}(t, \cdot)$ stetig in t . (Wir fordern also die verschärften Picard-Lindelöf-Bedingungen) Wir interessieren uns für 'kleine Störungen' der 'Anfangsposition' zur Zeit s .

Die folgende Plausibilitätsüberlegung soll nun zeigen, dass die kleine Störung einer linearen Differentialgleichung folgt: Die ungestörte Lösungskurve durch (s, \tilde{x}) befindet sich nach der kurzen Zeit ε im Punkt $\tilde{x} + \varepsilon \cdot v(s, \tilde{x})$; die Lösungskurve durch den benachbarten Punkt $(s, \tilde{x} + h \cdot y)$ ($|h|$ klein) befindet sich nach der kurzen Zeit ε im Punkt

$$\begin{aligned} (\tilde{x} + h \cdot y) + \varepsilon \cdot V(s, \tilde{x} + h \cdot y) &\sim \tilde{x} + \varepsilon \cdot V(s, \tilde{x}) + h \cdot (y + \varepsilon \cdot W(s, \tilde{x}) \cdot y) \\ &= \tilde{x} + \varepsilon \cdot V(s, \tilde{x}) + h \cdot (I + \varepsilon \cdot W(s, \tilde{x})) \cdot y. \end{aligned}$$

In Worten: Die Störung des Startpunkts in Richtung des Tangentialvektors y wird in infinitesimaler Zeit ε zur Störung des Endpunkts in Richtung $(I + \varepsilon \cdot W) \cdot y$.

$$\dot{y}(s) = W(s, x(s)) \cdot y(s) = A(s) \cdot y(s).$$

Statt $W(s, x(s))$ schreibt man auch $v_*(s, x(s))$. Es handelt sich um eine lineare Abbildung des Tangentialraums in sich.

Wir bemerken: Wenn man ein n -Tupel linear unabhängiger Anfangsbedingungen als die Spalten einer Matrix z schreibt, dann ergibt das eine Matrixgleichung infinitesimaler Störungen entlang der Lösungskurven $\{(s, x(s)) : s\}$.

Die Plausibilitätsüberlegung dient der Motivation für das Konzept der Gleichung der Variationen zu einem nichtautonomen System.

Wir schicken voraus: Zu jeder stetig differenzierbaren Abbildung φ gehört in jedem Punkt P eine lineare Abbildung des Tangentialraums T_P in den Tangentialraum $T_{\varphi(P)}$, der Pushforward $\varphi_*|_P$. In derselben Weise assoziiert man mit der Differentialgleichung $\dot{x} = v(t, x)$ (für die x in der Karte U) ein System von Differentialgleichungen

$$\begin{cases} \dot{x} = v(t, x) & \text{für } x \in U \\ \dot{y} = v_*(t, x(t)) \cdot y & \text{für } y \in \bigcup_{x \in U} T_x \end{cases}$$

Es zeigt sich, dass man über das Studium dieser ‘Gleichung der Variationen’ feine Aussagen über den lokalen Fluss $\{\varphi_s^t : s, t\}$ gewinnen kann.— Wir verweisen auf das wunderbare Lehrbuch

V. I. Arnol'd: Ordinary Differential Equations: (Chapter 4: Proofs of the main Theorems) Springer-Verlag 1992.

Lineare Gleichungen mit periodischen Koeffizienten

Wir betrachten Lösungskurven einer homogenen linearen Gleichung

$$\dot{y} = A(t) \cdot y \quad y(s) = \tilde{y} \quad ,$$

wo $A(\cdot)$ die Periode T hat, $A(t + T) = A(t)$.

Wir betrachten die Lösung zu den Zeitpunkten

$$s, s + T, s + 2T, \dots, \quad \text{also } y(s), y(s + T), y(s + 2T), \dots$$

Mit der Abkürzung $B = B_s = Y_s^{s+T}$ haben wir

$$y(s) = \tilde{y}, \quad y(s + T) = B\tilde{y}, \dots, \quad y(s + nT) = B^n\tilde{y}, \dots \quad .$$

Die Matrix B_s heißt der Monodromie-Operator (zur Startzeit s). Man bemerkt, dass die Monodromie-Operatoren zu verschiedenen Startzeiten zueinander konjugiert (‘ähnlich’) sind.

$$B_u = M^{-1}B_sM \quad \text{mit} \quad M = Y_u^s \quad .$$

Interessant sind nun die Eigenvektoren des Monodromie-Operators, besonders diejenigen zum Eigenwert $= 1$. Wir machen plausibel, dass diese zu periodischen Lösungen der Gleichung $\dot{y} = Ay$ führen.

Gegeben sei ein zeitunabhängiges Vektorfeld auf einer $(d + 1)$ -dimensionalen Mannigfaltigkeit, welches eine periodische Lösungskurve $\tilde{\gamma}(t)$ besitzt.

Sei $\{y^0, y^1, \dots, y^d\}$ ein Koordinatensystem mit

$$y^j(\tilde{\gamma}(t)) = 0 \quad \text{für } j = 1, 2, \dots, d, \text{ alle } t \quad .$$

Wir nehmen an, dass es eine derartige Koordinatisierung gibt. Wir nehmen auch an, dass das Vektorfeld in der Nähe der Kurve $\tilde{\gamma}(\cdot)$ nicht verschwindet; dann gilt also $\dot{\mathbf{y}}^0(t) \neq 0$ entlang von $\tilde{\gamma}(\cdot)$, während die übrigen $\mathbf{y}^j(t)$ auf der Kurve verschwinden. Wir interessieren uns für das Verhalten der Lösungskurven in der Nähe der periodischen Kurve $\tilde{\gamma}(\cdot)$.

Im Falle $d = 1$ gibt es im Wesentlichen nur zwei Möglichkeiten. Sie könnten sich „spiralenförmig“ von $\tilde{\gamma}(\cdot)$ weg oder auf $\tilde{\gamma}(\cdot)$ zu bewegen. (Denkbar ist natürlich auch der Fall, dass sie selbst geschlossene Kurven sind.)

Einen Eindruck vom Verhalten der Lösungskurven (auch im Fall $d > 1$) erhält man, wenn man die Differentialgleichung „linearisiert“. Die gegebene Gleichung kann man offenbar in die Form bringen.

$$\frac{d\mathbf{y}^j}{d\mathbf{y}^0} = W^j(\mathbf{y}^0, \mathbf{y}^1, \dots, \mathbf{y}^d) \quad \text{mit} \quad W^j(\mathbf{y}^0, 0, \dots, 0) = 0 \quad .$$

Die Taylor-Approximation (für jedes feste \mathbf{y}^0) liefert

$$W^j(\mathbf{y}^0, \mathbf{y}) = A(\mathbf{y}^0) \cdot \mathbf{y} + o(\|\mathbf{y}\|) \quad .$$

Wir führen die gewohnten Notationen ein, $\mathbf{t} = \mathbf{y}^0$, $A(\mathbf{t}), \dots$

Die linearisierte Gleichung lautet $\dot{\mathbf{y}} = A(\mathbf{t}) \cdot \mathbf{y}$, wo $A(\cdot)$ periodisch ist.

Beispiel 2.1.6.

Wir studieren die Kurven in der Ebene in der Nähe des Einheitskreises $r = 1$, welche die folgende Differentialgleichung erfüllen

$$\dot{r} = (r^2 - 1) \cdot \frac{1}{2}(2r \cos \varphi - 1); \quad \dot{\varphi} = 1 \quad .$$

Mit $\mathbf{y} = r - 1$, $\varphi = \mathbf{t}$ haben wir die gewohnten Bezeichnungen

$$\dot{\mathbf{y}} = \mathbf{y} \cdot \frac{1}{2}(\mathbf{y} + 2)(2(\mathbf{y} + 1) \cdot \cos \mathbf{t} - 1) \quad .$$

Die linearisierte Gleichung lautet $\dot{\mathbf{y}} = \mathbf{y} \cdot (2 \cos \mathbf{t} - 1)$. Die Koeffizienten haben die Periode 2π . Die Lösungen sind

$$\tilde{\mathbf{y}}(\mathbf{t}) = \exp\left(\int_0^{\mathbf{t}} (2 \cos u - 1) du\right) \cdot \tilde{\mathbf{y}} = \exp(-\mathbf{t}) \cdot \exp(2 \sin \mathbf{t}) \cdot \tilde{\mathbf{y}}$$

Der Monodromie-Operator ist die Multiplikation mit $\exp(-2\pi)$. Die Lösungskurven nähern sich spiralenförmig dem Einheitskreis.

2.2 Die Idee des Randoperators, Orientierung

Ein affines Simplex der Dimension m ist die konvexe Hülle von $m + 1$ Punkten in allgemeiner Lage. Je $l + 1$ Extrempunkte bestimmen eine l -dimensionale Seite. Das Simplex selbst ist die einzige m -dimensionale Seite. Es gibt genau $m + 1$ Randsimplizes der Dimension $m - 1$. Die Extrempunkte sind die 0-dimensionalen Seiten.

Durch eine Aufzählung seiner Extrempunkte P_0, P_1, \dots, P_m erhält ein m -dimensionales Simplex eine Orientierung. Zwei Aufzählungen, die durch eine gerade Permutation auseinander hervorgehen, definieren dieselbe Orientierung. Die entgegengesetzte Orientierung erhält man, wenn man das Tupel der Extrempunkte einer ungeraden Permutation unterwirft. Eine Orientierung eines m -dimensionalen Simplex ($m \geq 1$) induziert eine Orientierung auf jedem $(m - 1)$ -dimensionalen Randsimplex. Wie das funktioniert, erläutern wir zuerst für die Dimension $m = 2$.

$$[P_0, P_1, P_2] \quad \rightsquigarrow \quad \left\{ \begin{array}{l} [P_1, P_2] \\ [P_2, P_0] \\ [P_0, P_1] \end{array} \right.$$

Man notiert

$$\begin{aligned} \partial[P_0, P_1, P_2] &= [P_1, P_2] + [P_2, P_0] + [P_0, P_1] = [P_1, P_2] - [P_0, P_1] + [P_0, P_1]; \\ \partial[P_0, P_1, \dots, P_m] &= \sum_0^m (-1)^k \cdot [P_0, \dots, \overline{P_k}, \dots, P_m] \end{aligned}$$

Die Überstreichung bedeutet, dass der überstrichene Punkt wegzulassen ist. Das Minuszeichen zeigt an, dass die nachfolgende Punktfolge einer ungeraden Permutation unterzogen ist. Man überzeugt sich, dass der Randoperator ∂ für ein orientiertes Simplex S wohldefiniert ist; d. h. wenn man die Folge $[P_0, P_1, \dots, P_m]$ einer ungeraden Permutation unterzieht, dann ändern alle Summanden das Vorzeichen. Einem positivorientierten Simplex ordnen wir das algebraische Symbol $|S\rangle$ zu, einem negativorientierten Simplex das algebraische Symbol $-|S\rangle$. Offenbar gilt

$$\partial(-|S\rangle) = -\partial(|S\rangle).$$

Es sei nun $\{S_\alpha : \alpha \in I\}$ eine Familie von orientierten Simplizes mit der Eigenschaft, dass auch alle Randsimplizes der S_α (in der einen oder anderen Orientierung) zur Familie gehören. Der Randoperator wird auf formale Summen ausgedehnt. Die formalen ganzzahligen Summen bezeichnen wir $\sum c_\alpha \cdot |S_\alpha\rangle$, wo $\{c_\alpha : \alpha \in I\}$ eine Familie ganzer Zahlen ist mit nur endlich vielen $\neq 0$. Diese algebraischen Objekte heißen die ganzzahligen Ketten. Man definiert den Randoperator auf dieser abelschen Gruppe \mathfrak{C}

$$\partial\left(\sum c_\alpha \cdot |S_\alpha\rangle\right) = \sum c_\alpha \cdot \partial(S_\alpha).$$

Der Rand eines 0-dimensionalen Simplex ist $= 0$. Mit den formalen Summen der Simplizes S_α wird wie üblich gerechnet: $\sum c_\alpha \cdot |S_\alpha\rangle + \sum d_\alpha \cdot |S_\alpha\rangle = \sum (c_\alpha + d_\alpha) \cdot |S_\alpha\rangle$.

Wenn in einer Kette nur Simplizes mit der Dimension $\leq q$ nichtverschwindende Koeffizienten haben, dann sagen wir, es handle sich um eine Kette von der Dimension $\leq q$. Die Gruppe dieser Ketten heisst die Kettengruppe zum q -dimensionalen Gerüst und wird mit $\mathfrak{C}_{\leq q}$ bezeichnet. Eine Kette, in welcher nur Simplizes der Dimension q mit einem Koeffizienten $\neq 0$ vorkommen, heisst eine Kette der Dimension q . (Man beachte, dass die Null eine Kette von jeder Dimension ist.) Die Gruppe der Ketten von der Dimension q wird mit \mathfrak{C}_q bezeichnet. \mathfrak{C} ist die direkte Summe dieser Untergruppen:

$$\mathfrak{C} = \mathfrak{C}_0 \oplus \mathfrak{C}_1 \oplus \mathfrak{C}_2 \oplus \cdots = \bigoplus_{q=0} \mathfrak{C}_q.$$

Der Randoperator bildet \mathfrak{C}_q in \mathfrak{C}_{q-1} ab; ∂_q bezeichnet seine Einschränkung auf \mathfrak{C}_q .

$$0 \leftarrow \mathfrak{C}_0 \xleftarrow{\partial_1} \mathfrak{C}_1 \xleftarrow{\partial_2} \mathfrak{C}_2 \xleftarrow{\partial_3} \cdots$$

Satz 2.2.1. *Der Rand eines Rands verschwindet; $\partial \circ \partial = 0$*

Beweis. *Es genügt, das für Simplizes der Dimension $m \geq 2$ nachzuweisen. Sei $S = [P_0, P_1, \dots, P_m]$ ein Simplex mit aufgezählten Ecken. In ∂S tauchen nur solche Tupel auf, die einen Extrempunkt weniger haben; in $\partial \circ \partial(S)$ kommen diejenigen mit zwei Extrempunkten weniger vor, jedoch mit dem Koeffizienten $= 0$, weil die Reihenfolge des Weglassens verschiedene Vorzeichen ergibt.*

Die Ketten der Dimension q , deren Rand verschwindet, heissen die q -dimensionalen Zyklen. \mathfrak{Z}_q bezeichnet die Gruppe der q -dimensionalen Zyklen: $\mathfrak{Z}_q = \{c : \partial_q c = 0\}$. Die Untergruppe der q -dimensionalen Ränder wird mit \mathfrak{B}_q bezeichnet: $\mathfrak{B}_q = \partial_{q+1} \mathfrak{C}_{q+1}$. Die Faktorgruppe

$$\mathfrak{H}_q = \mathfrak{Z}_q / \mathfrak{B}_q = \ker \partial_q / \text{im } \partial_{q+1}.$$

heisst die q -te ganzzahlige Homologiegruppe der Familie $\{S_\alpha : \alpha \in I\}$.

Gedächtnishilfe: Die Buchstaben \mathfrak{C} , \mathfrak{Z} , \mathfrak{B} erklären sich durch die englischen Wörter:

Kette = chain, Rand = boundary, Jeder Rand ist ein Zyklus.

Orientierungen im Sinne der Geometrie:

Ein reellaffiner Raum wird zu einem orientierten affinen Raum, indem man für jede aufgezählte Basis seines Tangentialraums festlegt, ob sie als positiv oder negativ gelten soll; dabei ist Verträglichkeit gefordert: das Vorzeichen ändert sich, wenn die Matrix des Basiswechsels eine negative Determinante hat. — Offensichtlich besitzt ein affiner Raum genau zwei Orientierungen.

Konvention: Wenn P_0, P_1, \dots, P_m ein affines Simplex mit aufgezählten Ecken ist, dessen Orientierung als positiv festgelegt wurde, dann induziert das eine Orientierung des aufgespannten affinen Raums: Wenn $P_1 = P_0 + e_1$, $P_2 = P_0 + e_1 + e_2, \dots$ dann soll die so aufgezählte Basis des Tangentialraums als positiv gelten.

Eine Mannigfaltigkeit wird zu einer orientierten Mannigfaltigkeit, indem man für jede Aufzählung der Elemente einer lokalen Basis festlegt, ob sie als positiv und als negativ gelten soll. Dabei ist Verträglichkeit gefordert: Wenn zwei Kartengebiete einen nichtleeren Durchschnitt haben, dann haben die Aufzählungen der lokalen Koordinatensysteme genau dann dasselbe Vorzeichen, wenn die Jacobimatrix überall im Durchschnitt positives Vorzeichen hat. Man überlegt sich leicht: Wenn eine zusammenhängende Mannigfaltigkeit überhaupt eine Orientierung besitzt, dann besitzt sie genau zwei Orientierungen, Die einfachste Mannigfaltigkeit, die keine Orientierung besitzt ist das Möbiusband.

Affine Kettenkomplexe

Der Rand wurde oben rein kombinatorisch konstruiert. Wir denken nun aber auch kurz (zur Unterstützung der Anschaulichkeit) an Simplizes (oder Zellen), die als konvexe Mengen in einem hochdimensionalen affinen Raum liegen.

Definition. Es sei \mathcal{S} eine Menge von affinen Simplizes (in irgendeinem affinen Raum) mit den Eigenschaften

- (i) Gehört das Simplex S zu \mathcal{S} , so gehört auch jede Seite zu \mathcal{S}
- (ii) Für je zwei Simplizes in \mathcal{S} ist der Durchschnitt entweder leer oder eine gemeinsame Seite.

Die Vereinigung dieser Simplizes heisst dann ein simplizialer Komplex oder auch ein simplizial zerlegter affiner Komplex. Die Topologen definieren: Ist ein topologischer Raum X zu einem simplizialen Komplex homöomorph, so heisst er triangulierbar oder auch ein topologisches Polyeder. Eine Triangulation oder simpliziale Zerlegung des topologischen Raums X liegt vor, wenn eine topologische Abbildung eines simplizialen Komplexes auf X fixiert ist.

Wichtig ist die Verallgemeinerung, die man einen affinen Zellenkomplex nennt. Statt der Simplizes hat man kompakte konvexe Polyeder mit den obigen Eigenschaften. Jede m -dimensionale orientierte Zelle K induziert eine Orientierung für jede der $(m - 1)$ -dimensionalen Randzellen gemäß der folgenden Verabredung: Es sei $[P_1, \dots, P_m]$ ein Simplex mit aufgezählten Ecken in einer Randfläche und $[P_0, P_1, \dots, P_m]$ ein Simplex in K ; der Randoperator überträgt dann die Orientierung der Eckenanzählung.

Die weiteren Begrifflichkeiten um die ‘topologischen Polyeder’ müssen wir hier noch nicht entwickeln; wir interessieren uns vorerst nur für die kombinatorischen Gegebenheiten. Wir diskutieren zunächst einmal das eindimensionale Gerüst eines simplizialen Komplexes; denn hier können wir an die Vorstellungsweisen der Graphentheorie anknüpfen.

Der Kettenkomplex zu einem Graphen

Gegeben sei ein ungerichteter Graph ohne Schleifen und Doppelkanten über der Scheitelmenge S^0 . Die Kantenmenge sei mit S^1 bezeichnet. Wir spezifizieren eine Orientierung der Kanten, d. h. wir spezifizieren für jede Kante $k \in S^1$, welchen ihrer Randpunkte wir als den Anfangspunkt $\alpha(k)$ ansehen wollen (und welchen als den Endpunkt $\beta(k)$). Die Kante k mit $\alpha(k) = P$, $\beta(k) = Q$ notieren wir auch $k = [P, Q]$. Eine formale Summe $\sum c_k \cdot k = \sum c_k \cdot [\alpha(k), \beta(k)]$ mit ganzzahligen ‘Gewichtungen’ c_k nennen wir eine

1-Kette. Die Gruppe der 1-Ketten bezeichnen wir mit \mathfrak{C}_1 . Die Menge der ganzzahligen Gewichtungen der Scheitel des Graphen bezeichnen wir mit \mathfrak{C}_0 . Der Randoperator ∂ bildet \mathfrak{C}_1 in additiver Weise in \mathfrak{C}_0 ab mit $\partial k = \beta(k) - \alpha(k)$. Bemerke: Wenn k die Kante von P nach Q ist, dann entspricht $-k$ einer Kante von Q nach P . Wir notieren auch $-k = [Q, P]$.

Was man sich natürlicherweise als einen *Weg von $\tilde{P} = P_0$ nach $\tilde{Q} = P_n$* vorstellen möchte (ohne scharfe mathematische Definition!) liefert eine 1-Kette $c = [P_0, P_1] + [P_1, P_2] + \cdots + [P_{n-1}, P_n]$ mit $\partial c = P_n - P_0 = \tilde{Q} - \tilde{P}$. Man spricht bekanntlich von einem doppelpunktfreien Weg, wenn die P_k paarweise verschieden sind. Einen doppelpunktfreien geschlossenen Weg nennt man eine Masche. Zu den geschlossenen Wegen und speziell zu den Maschen gewinnt man 1-dimensionale Zyklen. Die Zyklen sind diejenigen formalen Linearkombinationen $c = \sum_n [P_n, Q_n]$, in welchen jeder Scheitel (wenn überhaupt) genau so oft an erster wie an zweiter Stelle vorkommt.

Mit dem Begriff des aufspannenden Baums kann man sich ein sehr einfaches Bild von der Gruppe der 1-Zyklen \mathfrak{Z}_1 machen. Man beweist nämlich leicht den

Satz. *In einem ungerichteten Graphen ohne Schleifen und Doppelkanten existiert eine Menge von Maschen, sodass jeder Zyklus in eindeutiger Weise als ganzzahlige Linearkombination der Maschen darstellbar ist.*

Der Satz impliziert offenbar, dass \mathfrak{Z}_1 isomorph ist zur Gruppe \mathbb{Z}^n für ein gewisses n .

Wir bemerken: Das zweidimensionale Gerüst eines kombinatorischen Komplexes entsteht aus dem eindimensionalen dadurch, dass man erklärt, welche Tripel von Ecken die Ecken eines Dreiecks im Komplex sind und welche nicht. Besonders übersichtlich ist die Lage bei den sog. planaren Graphen. Das sind diejenigen Graphen, deren Scheitel man so in die Ebene legen kann, dass sich die durch Kurvenstücke realisierten Kanten des Graphen nur in Scheiteln treffen.

Warnung

Wir haben das alles recht ausführlich geschildert, um Verständnisschwierigkeiten vorzubeugen, die beim Anfänger manchmal entstehen, wenn später den Gruppen $\mathfrak{C}_q, \mathfrak{Z}_q, \mathfrak{B}_q$ die reellen Vektorräume der Coketten, Cozyklen und Coränder $\mathfrak{C}^q, \mathfrak{Z}^q, \mathfrak{B}^q$ gegenübergestellt werden. Die Coketten sind ganz andere Objekte als die Ketten.

Man darf sich vorstellen, dass der Vektorraum der q -Coketten \mathfrak{C}^q der Dualraum des Raums der q -Ketten \mathfrak{C}_q ist. Streng genommen gilt das aber nur, wenn man zuerst von den ganzzahligen Gewichtungen auf der Menge der q -Simplizes zu den reellen Gewichtungen auf der Menge der q -Simplizes übergegangen ist. — Das werden wir auch tun, wenn wir im Beispiel unten elektrische Netzwerke diskutieren.

Die 0-Coketten kann man in jedem Fall als die reellwertige Funktionen auf der Scheitelmengemenge S^0 verstehen; die 1-Coketten sind die ‘Funktionen’ auf der Kantenmenge; später in der Theorie der Mannigfaltigkeiten werden sie uns als die 1-Formen erscheinen. Der Corand-Operator d ist der zum Randoperator ∂ duale Operator.

$$0 \rightarrow \mathfrak{C}^0 \xrightarrow{d} \mathfrak{C}^1 \xrightarrow{d} \mathfrak{C}^2 \xrightarrow{d} \cdots$$

In der Dimension 0 haben wir eine lineare Abbildung $\mathfrak{C}^0 \rightarrow \mathfrak{C}^1$. Sie macht aus der Funktion f auf der Scheitelmenge die 1-Cokette df , welche der Kante den Anstieg der Funktion entlang dieser Kante zuordnet: Wir notieren

$$\langle df, k \rangle = f(\beta(k)) - f(\alpha(k)) = \langle f, \partial k \rangle \quad \text{für } f \in \mathfrak{C}^0, k \in \mathfrak{C}_1.$$

In der Dimension 1 haben wir eine lineare Abbildung der Kantenfunktionen $\omega^1 \mapsto \omega^2 = d\omega^1$. Eine 1-Form $\omega^1 \in \mathfrak{C}^1$ liefert für jedes 2-Simplex $[P_0, P_1, P_2]$ eine Zahl, nämlich

$$\langle d\omega^1, [P_0, P_1, P_2] \rangle = \langle \omega^1, [P_1, P_2] \rangle + \langle \omega^1, [P_2, P_0] \rangle + \langle \omega^1, [P_0, P_1] \rangle.$$

Diese Zahl nennt man manchmal die Rotation von ω^1 über dem Simplex.

Es muss empfohlen werden, dass man sich die Elemente aus \mathfrak{C}_q immer als 'Gewichtungen' vorstellt (es ist üblich, nur ganzzahlige Gewichtungen zuzulassen) und die Elemente \mathfrak{C}^q (mit hochgestelltem q) als 'Funktionen'. Die Operatoren ∂ ('Rand') und d ('Corand') sind zueinander dual. Der Rand erniedrigt die Dimension, der Corand erhöht sie.

Beispiel 2.2.1 (Elektrische Netzwerke). S^0 sei eine Menge von Lötstellen und S^1 eine Menge von leitenden Verbindungen. Wenn wir die Lötstellen auf ein Potential f bringen, dann bringt das auf die leitenden Verbindungen einen Potentialabfall ('Spannung') df . Wenn wir Ströme in den Leitungen haben, dann beschreiben wir diese durch 1-Ketten. Ein beliebiger Strom kann durch Abflüsse nach und Zuflüsse von draussen in Gang gehalten werden; im Knoten \tilde{P} muss der Überschuss der von den Nachbarknoten zu- und abfließenden Stromstärken ab- oder zugeleitet werden. Der Rand ∂c beschreibt den von den Lötstellen nach aussen fließenden Strom. Für $c = [P_0, P_1] + [P_1, P_2] + \dots + [P_{n-1}, P_n]$ haben wir $\partial c = [P_n] - [P_0]$. Beispiel: Wenn für einen Strom c gilt $\partial c = g \cdot [\tilde{Q}] - g \cdot [P]$, dann interpretieren wir das so, dass ein Strom von g Ampère fließt von \tilde{Q} nach P fließt. Die Zyklen sind die stationären Ströme ohne Ein- und Abfluss. Zweidimensionale Simplexe gibt es hier nicht.

Hinweis: Die Theorie der elektrischen Schaltkreise ist eine Approximation an die Theorie des Elektromagnetismus. In dieser Approximation wird angenommen, dass die Phänomene dadurch beschrieben werden können, dass man sagt, was in endlich vielen Schaltelementen vorgeht, dass also der umgebende Raum mit seinen Feldern keine Rolle spielt. (Siehe z. B. Feynman's Lecture Notes, Bd. II, Kap. 22 ff.)

Ein Schaltkreis setzt sich aus passiven und aktiven Komponenten k zusammen, welche irgendwie in Knoten P verbunden sind. Es geht um den Spannungsabfall V^k und den Strom I_k in allen Komponenten k , wenn in gewissen 'aktiven' Komponenten elektromotorische Kräfte \mathcal{E}^k bzw. Ströme eingespeist werden. Die elementaren Schaltelemente sind Induktoren ('Spulen'), Kondensatoren und Ohm'sche Widerstände. (In Wirklichkeit kann natürlich die Trennung des Gesamtsystems in elementare Komponenten nicht vollkommen sein.) Der einfachste Schaltkreis, in welchem alle drei Typen sowie ein Generator vorkommen, ist durch den harmonischen Oszillator gegeben; dort sind alle Elemente 'in Reihe' geschaltet. Die Schaltelemente kann man auch 'parallel' schalten. Die Theorie der Schaltkreise befasst sich mit komplizierteren 'Verlötungen'.

Das Beispiel wird uns noch weiter beschäftigen. Bei schwachen Strömen kann man mit linearen Zusammenhängen zwischen den Strömen und den Spannungen rechnen. Die elektrische Energie auf den Kondensatoren und die in der Zeiteinheit dissipierte Energie in den Widerständen sind durch quadratische Formen beschrieben.

Eine für Mathematiker gut lesbare Beschreibung der Theorie findet sich in dem hervorragenden Lehrbuch von Paul Bamberg und Shlomo Sternberg: *A course of mathematics for students of physics in the second year, Chap. 12 -14, p. 407-526.*

Auch in den Szenarien der nächsten ‘Beispiele’ werden wir keine substantiellen Sätze beweisen. Es kann nur darum gehen, erste Ideen vom algebraischen Umgang mit dem Randoperator ∂ zu vermitteln.

Beispiel 2.2.2 (Nullhomologe Wege in Gebieten $G \subseteq \mathbb{C}$).

Bei der Diskussion holomorpher Differentiale $f(z) dz$ über Gebieten G in der komplexen Ebene haben wir polygonale Kurvenzüge homotop deformiert. Ein zentrales Ergebnis war der Cauchy’sche Integralsatz in der Homotopieversion: Das Integral einer holomorphen Form über eine geschlossene Kurve ergibt jedenfalls dann den Wert 0, wenn man die Kurve in G auf einen Punkt zusammenziehen kann. (Die Kurven dieser Art heissen die nullhomotopen Kurven.) Eine Verallgemeinerung dieses Satzes ist die Homologieversion des Cauchy’schen Integralsatzes; sie besagt, dass auch das Integral über nullhomologe Kurven verschwindet. Die nullhomogenen Polygonzüge sind die Ränder der zweidimensionalen affinen Komplexe im Gebiet G . In den Lehrbüchern der Funktionentheorie findet man einfache Beispiele von geschlossenen Kurven in der zweifach punktierten komplexen Ebene $G = \mathbb{C} \setminus \{+1, -1\}$, die nullhomolog sind, aber nicht auf einen Punkt zusammengezogen werden können. Man kann zeigen, dass eine geschlossene Kurve im Gebiet G genau dann nullhomolog ist, wenn für jeden Punkt \tilde{z} ausserhalb G die Umlaufszahl der Kurve in Bezug auf \tilde{z} verschwindet, wenn also, salopp gesprochen, jedes $\tilde{z} \notin G$ von der Kurve genauso oft im positiven wie im negativen Sinn umlaufen wird.

Beispiel 2.2.3 (Singuläre Simplexe und ihre Ränder).

Wir haben früher definiert: Ein parametrisiertes glattes Kurvenstück auf einer Mannigfaltigkeit ist eine stetig differenzierbare Abbildung eines Intervalls in die Mannigfaltigkeit $\{\gamma(t) : t' \leq t \leq t''\}$. Solche parametrisierten Kurvenstücke liefern dieselbe Kurve, wenn sie durch eine stetig differenzierbare Umparametrisierung aus einander hervorgehen.

Verabredung: Stetige Differenzierbarkeit einer Abbildung φ der Menge S soll im Folgenden immer heissen, dass φ die Einschränkung einer auf einer Umgebung von S stetig differenzierbaren Abbildung ist.

Definition (Glatte singuläre Simplexe in einer Mannigfaltigkeit).

Ein parametrisiertes glattes orientiertes m -Simplex auf einer Mannigfaltigkeit ist eine stetig differenzierbare Abbildung eines affinen m -Simplexes mit aufgezählten Ecken. Zwei solche Abbildungen $\{\gamma'(t) : t \in [P_0, P_1, \dots, P_m]\}$ und $\{\gamma''(s) : s \in [Q_0, Q_1, \dots, Q_m]\}$ beschreiben (definitionsgemäß!) dasselbe orientierte glatte parametrisierbare Simplex $\gamma(\cdot)$ auf der Mannigfaltigkeit, wenn stetig differenzierbare Umparametrisierungen $T(\cdot), S(\cdot)$

existieren, wenn also $\gamma'(T(s)) = \gamma''(s)$, $\gamma''(S(t)) = \gamma(t)$ mit Diffeomorphismen $S(\cdot), T(\cdot)$

$$T(S(t)) = t \quad \text{für } t \in [P_0, \dots, P_m], \quad \text{und } S(T(s)) = s \quad \text{für } s \in [Q_0, \dots, Q_m]$$

sowie $T(Q_k) = P_{\pi(k)}$ mit einer geraden Permutation π . Wenn eine Umparametrisierung mit einer ungeraden Permutation $\pi(\cdot)$ existiert, dann sagt man, die singulären Simplexes seien bis auf die Orientierung gleich; und bezeichnet dieses singuläre Simplex mit $-\gamma(\cdot)$.

Man beachte: Es ist nicht gefordert, dass $\gamma(\cdot)$ eine Immersion ist. Es ist eine sehr spezielle Situation, wenn das Bild des offenen affinen Simplexes eine Untermannigfaltigkeit ist; in diesem Fall nennt man das Bild eine glatt parametrisierbare Untermannigfaltigkeit.

Die Bezeichnung 'singulär' hat keine gute Begründung; es dient traditionell der Unterscheidung von anderen Typen von Simplexes.

Die formalen Summen von orientierten Simplexes heißen die ganzzahligen singulären Ketten. Der Rand eines m -dimensionalen orientierten singulären Simplexes $\gamma(\cdot)$ ist eine $(m-1)$ -dimensionale singuläre Kette, die in offensichtlicher Weise als Summe von $m+1$ orientierten Simplexes geschrieben werden kann. Man betrachtet einfach die Einschränkungen von $\gamma(\cdot)$ auf die Randsimplexes (mit ihren Orientierungen): $\gamma_k(\cdot) : k = 0, 1, \dots, m$ und summiert

$$\partial\gamma = \sum_{k=0}^m (-1)^k \cdot \langle \gamma_k \rangle.$$

Der Rand eines Rands ist 0. Eine singuläre Kette mit verschwindendem Rand heißt ein singulärer Zyklus. Die Menge der q -dimensionalen Zyklen ist eine kommutative Gruppe, die Menge der q -dimensionalen Ränder ist eine Untergruppe; die Faktorgruppe heißt die q -te singuläre Homologiegruppe der Mannigfaltigkeit.— Die singulären Homologiegruppen sind ein reichlich kompliziertes Thema der algebraischen Topologie. Es zeigt sich, dass die Theorie der Cohomologie-Gruppen weniger kompliziert ist. Es gibt also gute Gründe, sich mit den Coketten, Cozyklen und Corändern anzufreunden, auch wenn diese auf den ersten Blick weniger anschaulich sein mögen. In diesem Sinne werden wir uns gründlich mit den k -Formen befassen. Dazu brauchen wir aber aus der Linearen Algebra den Begriff der alternierenden Multilinearform. Wir holen weiter aus.

2.3 Tensoren und ihre Indizes

Es seien

$\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_k$ Vektorräume, $\mathbf{U}_1^*, \mathbf{U}_2^*, \dots, \mathbf{U}_k^*$ ihre Dualräume,
und $\langle \cdot, \cdot \rangle_1, \langle \cdot, \cdot \rangle_2, \dots, \langle \cdot, \cdot \rangle_k$ die dazugehörigen Dualitäten

Für $\mathbf{u} \in \mathbf{U}_1$ und $\mathbf{f} \in \mathbf{U}_1^*$ bezeichnet also $\langle \mathbf{f}, \mathbf{u} \rangle_1$ den Wert der Linearform \mathbf{f} im ‘Punkt’ \mathbf{u} , oder anders gesagt, den ‘Wert der Linearform’ \mathbf{u} im ‘Punkt’ \mathbf{f} .

So wie wir den Vektor \mathbf{u}_1 oder $\langle \cdot, \mathbf{u}_1 \rangle_1$ als Linearform auf \mathbf{U}_1^* verstehen können, so können wir das Produkt

$\langle \cdot, \mathbf{u}_1 \rangle \cdots \langle \cdot, \mathbf{u}_k \rangle_k$ als eine Multilinearform auf $\mathbf{U}_1^* \times \cdots \times \mathbf{U}_k^*$

verstehen. Man nennt eine solche sehr spezielle Multilinearform einen faktorisierten k -stufigen Tensor und notiert sie

$$\mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_k,$$

Den Vektorraum der Linearkombinationen solcher faktorisierbarer Tensoren, aufgefasst als Multilinearformen auf $\mathbf{U}_1^* \times \cdots \times \mathbf{U}_k^*$, nennt man das Tensorprodukt der Vektorräume \mathbf{U}_1 und bezeichnet diesen Vektorraum mit $\mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_k$. Wir werden sehen, dass man die Elemente dieses Vektorraums als die Linearformen auf dem Tensorprodukt $\mathbf{U}_1^* \otimes \mathbf{U}_2^* \otimes \cdots \otimes \mathbf{U}_k^*$ auffassen kann.

Satz 2.3.1 (Spezielle Basisdarstellungen).

Jede Multilinearform auf $\mathbf{U}_1^* \times \mathbf{U}_2^* \times \cdots \times \mathbf{U}_k^*$ lässt sich als Linearkombination von faktorisierten Tensoren schreiben. Wenn $\{\mathbf{u}_i : i \in I_1\}$ eine Basis von \mathbf{U}_1 ist, $\{\mathbf{u}_i : i \in I_2\}$ eine Basis von \mathbf{U}_2 usw., dann gewinnen wir im System aller Tensoren von der Gestalt

$$\mathbf{u}_{i_1} \otimes \mathbf{u}_{i_2} \otimes \cdots \otimes \mathbf{u}_{i_k} \quad \text{mit} \quad i_1 \in I_1, i_2 \in I_2, \dots, i_k \in I_k$$

eine Basis des Tensorprodukts $\mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_k$. Jeder Tensor besitzt eine Darstellung

$$\mathcal{T} = \sum_{I_1 \times \cdots \times I_k} \mathbf{a}^{i_1 i_2 \dots i_k} \cdot \mathbf{u}_{i_1} \otimes \mathbf{u}_{i_2} \otimes \cdots \otimes \mathbf{u}_{i_k}$$

mit eindeutig bestimmten Koeffizienten.

Im Falle der Bilinearformen, d. h. im Spezialfall $k = 2$ haben wir den Beweis ausführlich dargestellt. Für k -stufige Tensoren geht das alles genauso. Fazit: Für jemanden, der sich in jedem der ‘Faktoren’ des Tensorprodukts auf eine Basis festgelegt hat, ist ein Tensor ein System von k -fach indizierten Einträgen $(\mathbf{a}^{i_1 i_2 \dots i_k})_{I_1 \times \cdots \times I_k}$.

(Mit der Hochstellung der Indizes deutet man an, dass man die Tensoren als ‘ k -fach kontrahredient’ (oder ‘ k -fach kontravariant’) ansehen will, Wir werden unten auch noch sog. cogrediente und gemischte Tensoren kennenlernen; da gibt es dann noch weitere Konventionen, was das Hoch- und Tiefstellen der Indizes betrifft. Wir werden dort diskutieren, was es mit den Wörtern co- und kontrahredient oder -variant auf sich hat.)

Sprechweise 2.3.1 (Rang eines Tensors). Man sagt von einem Tensor, er habe den Rang $\leq r$, wenn er sich als Linearkombination von r faktorisierbaren Tensoren präsentieren lässt.

Multilineare Umformungen

Die Operation \otimes ist in jeder Position additiv und damit linear. (Skalare Faktoren kann man in jeden am Produkt beteiligten Vektoren 'hineinziehen'). Sie ist multilinear.

$$\begin{aligned} (\mathbf{u}'_1 + \mathbf{u}''_1) \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_k &= \mathbf{u}'_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_k + \mathbf{u}''_1 \otimes \mathbf{u}_2 \otimes \cdots \otimes \mathbf{u}_k \\ \mathbf{u}_1 \otimes (\mathbf{u}'_2 + \mathbf{u}''_2) \otimes \cdots \otimes \mathbf{u}_k &= \mathbf{u}_1 \otimes \mathbf{u}'_2 \otimes \cdots \otimes \mathbf{u}_k + \mathbf{u}_1 \otimes \mathbf{u}''_2 \otimes \cdots \otimes \mathbf{u}_k \\ &\dots \end{aligned}$$

Man beachte: Wenn uns zwei Summen gegeben sind

$$\sum_{\alpha} \mathbf{u}_1^{\alpha} \otimes \mathbf{u}_2^{\alpha} \otimes \cdots \otimes \mathbf{u}_k^{\alpha} \quad \text{und} \quad \sum_{\beta} \mathbf{u}_1^{\beta} \otimes \mathbf{u}_2^{\beta} \otimes \cdots \otimes \mathbf{u}_k^{\beta},$$

dann gibt es viele Möglichkeiten, eine evtl. vorliegende Gleichheit festzustellen. Man kann beide Summen in eine (willkürlich gegebene) Basis umrechnen. Dies läuft darauf hinaus, dass man für eine genügend große Familie von k -Tupeln $\{\omega_{\gamma} = \mathbf{f}_{\gamma}^1 \times \cdots \times \mathbf{f}_{\gamma}^k : \gamma \in \Gamma\}$ die Summen $\langle \omega_{\gamma}, \sum_{\alpha} \dots \rangle$ und $\langle \omega_{\gamma}, \sum_{\beta} \dots \rangle$ auswertet. Manchmal kann man aber auch mit einigen geschickten multilinearen Rechenschritten von der einen Summe zur anderen gelangen. Diese Bemerkung macht man manchmal zur Grundlage einer Definition des Tensorprodukts, welche die Wirkung auf Linearformen nicht ins Spiel bringt. Man definiert:

Definition 2.1. Das Tensorprodukt ergibt sich aus dem System aller formalen Summen von faktorisierbaren Tensoren, wenn man Summen $\sum_{\alpha} \mathbf{u}_1^{\alpha} \otimes \mathbf{u}_2^{\alpha} \otimes \cdots \otimes \mathbf{u}_k^{\alpha}$, die sich durch multilineare Operationen in einander umrechnen lassen, als gleich betrachtet.

Wenn man in einen Tensor aus $\mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_k$ an einer Stelle, sagen wir an der k -ten, eine feste Linearform $\mathbf{f}^k \in \mathbf{U}_k^*$ einsetzt, dann erhält man einen $k-1$ -stufigen Tensor aus $\mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_{k-1}$. Und diese Zuordnung ist linear. Wenn man an zwei Stellen, sagen wir an den beiden letzten, ein Paar von Linearformen einsetzen, dann erhält man einen $k-2$ -stufigen Tensor, und diese Zuordnung ist bilinear.

$$\begin{aligned} \varphi_{(k)}: \quad \mathbf{U}_k^* &\longrightarrow \mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_{k-1}, \\ \varphi_{(k-1,k)}: \quad \mathbf{U}_{k-1}^* \times \mathbf{U}_k^* &\longrightarrow \mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_{k-2}. \end{aligned}$$

Gemischte Tensoren

Die Vektoren aus \mathbf{U} und die Linearformen auf \mathbf{U} (d. h. also die Elemente von \mathbf{U}^*) bezeichnet man auch als Tensoren erster Stufe. Die Vektoren $\mathbf{v} \in \mathbf{U}$ heißen kontragrediente Tensoren erster Stufe, die Kovektoren $\mathbf{f} \in \mathbf{U}^*$ heißen cogrediente Tensoren erster Stufe.

Bei den Tensoren zweiter Stufe hat man vier Typen: Neben $\mathbf{U}_1 \otimes \mathbf{U}_2$ werden nämlich noch drei weitere Tensorprodukte ins Auge gefasst, nämlich $\mathbf{U}_1 \otimes \mathbf{U}_2^*$, $\mathbf{U}_1^* \otimes \mathbf{U}_2$ und $\mathbf{U}_1^* \otimes \mathbf{U}_2^*$. Die Elemente des ursprünglichen Tensorprodukts nennt man die zweifach kontragrediente Tensoren, die Elemente der weiteren werden kontra-cogredient, co-kontragredient und zweifach cogredient genannt. Die zweifach cogredienten Tensoren sind (in natürlicher Weise) mit den Linearformen auf dem Raum $\mathbf{U}_1 \otimes \mathbf{U}_2$ zu identifizieren. Im gleichen Sinn sind die Räume $\mathbf{U}_1 \otimes \mathbf{U}_2^*$, $\mathbf{U}_1^* \otimes \mathbf{U}_2$ zueinander dual.

Für allgemeines κ haben wir neben $\mathbf{U}_1 \otimes \mathbf{U}_2 \otimes \cdots \otimes \mathbf{U}_\kappa$ die weiteren $2^\kappa - 1$ Tensorprodukte zu betrachten, die wir erhalten, indem wir an irgendwelchen Positionen den Vektorraum \mathbf{U}_κ durch den Vektorraum \mathbf{U}_κ^* ersetzen. Die Elemente dieser Tensorprodukte nennt man gemischt kontra-cogrediente Tensoren (kontragredient in den Positionen \dots , und cogredient in den Positionen \dots). Nehmen wir an, wir hätten in jedem \mathbf{U}_κ eine Basis $\{\mathbf{u}_i : i \in I_\kappa\}$ gewählt und $\{\ell^i : i \in I_\kappa\}$ sei die duale Basis. Wir erhalten dann für jedes gemischte Tensorprodukt eine spezielle Basis bestehend aus Produkten von Vektoren und Covektoren, wo zu jedem κ ein \mathbf{u}_i bzw. ein ℓ^i zu wählen ist. Im Falle des dreistufigen Tensorprodukts $\mathbf{U}_1 \otimes \mathbf{U}_2^* \otimes \mathbf{U}_3$ beispielsweise hat ein solches Basiselement die Gestalt $\mathbf{u}_{i_1} \otimes \ell^{i_2} \otimes \mathbf{u}_{i_3}$. Die Tensoren besitzen eine eindeutige Darstellung

$$\mathcal{T} = \sum_{i_1 i_2 i_3} a^{i_1 i_2 i_3} \cdot \mathbf{u}_{i_1} \otimes \ell^{i_2} \otimes \mathbf{u}_{i_3}.$$

Dieses Beispiel dürfte die Konventionen des Hoch- und Tiefstellens des Indizes hinlänglich illustrieren. Die Konvention wird manchmal durch die Regeln erläutert, die beim Übergang zu neuen Basen zu beachten sind. Solche Regeln sind nur dann nötig, wenn man die Ausdrücke $\mathbf{u}_{i_1} \otimes \ell^{i_2} \otimes \mathbf{u}_{i_3}$ unterdrückt, wenn man also nur das System der Koeffizienten betrachtet (was zweifellos manchmal die Übersichtlichkeit verbessern kann, siehe das Zitat von H. Weyl im Ende des Abschnitts). Da wir die Produkte der Basiselemente nicht unterdrücken, ist alles geklärt durch den

Satz 2.3.2 (Basis- und Koordinatenwechsel).

Es sei \mathbf{G} die $I \times J$ -Matrix des Basiswechsels von $\{\mathbf{u}_i : i \in I\}$ nach $\{\mathbf{v}_j : j \in J\}$

$$(\mathbf{v}_1, \dots, \mathbf{v}_m) = (\mathbf{u}_1, \dots, \mathbf{u}_m) \cdot \mathbf{G} \quad \mathbf{v}_j = \sum \mathbf{u}_i \cdot g^i_j.$$

und $\mathbf{H} = \mathbf{G}^{-1}$ die inverse Matrix vom Format $J \times I$. $\mathbf{u}_i = \sum_{j \in J} \mathbf{v}_j \cdot h^j_i$

Für die Transformation der dualen Basen benötigt man dieselben Matrizen

$$\mathbf{m}^j = \sum g^j_i \cdot \ell^i; \quad \ell^i = \sum h^i_j \cdot \mathbf{m}^j.$$

Allerdings: Während man sich die Basen als Zeilen aufgelistet vorstellen sollte, sollte man sich die Koordinaten, d. h. die dualen Basen als Spalten aufgelistet denken.

Wenn man Covektoren (bzgl. einer Basis) durch Tupel von Skalaren ausdrückt, dann organisiert man diese üblicherweise als Zeilen; die Spalten transformieren sich wie eine

Basis, also ‘cogredient’; die Spalten, mit welchen man die Vektoren beschreibt, werden ‘contragredient’ transformiert.

Multilineare Abbildungen

Bei der Identifikation des Typs eines gemischten Tensors ist es oft hilfreich, zu fragen, welche Objekte der Tensor ‘verarbeiten’ kann, und welche Objekte dabei herauskommen. Ein Vektor in der k -ten Position kann einen Kovektor in der k -ten Position verarbeiten; und ein Kovektor in der k -ten Position kann einen Vektor in der k -ten Position verarbeiten. Ein co-contragredienter Tensor verarbeitet einen kontra-cogredienten Tensor in diesen Positionen, u. s. w.. Das Ergebnis der ‘Verarbeitung’ eines eingesetzten Objekts ist ein Tensor geringerer Stufe. So können die Tensoren k -ter Stufe auf vielfältige Weise als multilineare Abbildungen eines Tensorprodukts der Stufe m in ein Tensorprodukt der Stufe $k - m$ verstanden werden. Das Umgekehrte ist ebenso richtig: eine multilineare Abbildung eines (gemischten) Tensorprodukts der $(k - m)$ -ten Stufe in ein (gemischtes) Tensorprodukt der m -ten Stufe kann als ein (gemischter) Tensor der Stufe k verstanden werden. Konkret (und etwas speziell) gesprochen: Es sei

$$\mathcal{T} \in \mathbf{U}_1 \otimes \cdots \otimes \mathbf{U}_m \otimes \mathbf{U}_{m+1}^* \otimes \cdots \otimes \mathbf{U}_k^*.$$

Wenn man in diesen Tensor in den Positionen $m + 1, \dots, k$ Vektoren $(\mathbf{u}_{m+1}, \dots, \mathbf{u}_k)$ einsetzt, dann erhält man eine Multilinearform auf $\mathbf{U}_1^* \otimes \cdots \otimes \mathbf{U}_m^*$, oder anders gesagt, einen Tensor $\mathcal{T}_{(\mathbf{u}_{m+1}, \dots, \mathbf{u}_k)} \in \mathbf{U}_1 \otimes \cdots \otimes \mathbf{U}_m$. Die Zuordnung

$$\mathbf{U}_{m+1} \times \cdots \times \mathbf{U}_k \ni (\mathbf{u}_{m+1}, \dots, \mathbf{u}_k) \longmapsto \mathcal{T}_{(\mathbf{u}_{m+1}, \dots, \mathbf{u}_k)} \in \mathbf{U}_1 \otimes \cdots \otimes \mathbf{U}_m$$

ist multilinear. Sie ist also eine lineare Abbildung des Tensorprodukts $\mathbf{U}_{m+1}^* \otimes \cdots \otimes \mathbf{U}_k^*$ in das Tensorprodukt $\mathbf{U}_1 \otimes \cdots \otimes \mathbf{U}_m$.

Beispiel 2.3.1. Wir erläutern den Standpunktwechsel am Beispiel einer $I \times J$ -Matrix $\mathbf{A} = \mathbf{a}^i_j$, die zunächst, wie in der Linearen Algebra üblich, als eine lineare Abbildung der J -Spalten in den Raum der I -Spalten verstanden werden soll

$$\varphi_{\mathbf{A}}: \quad \mathbb{R}_{\text{Sp}}^J \ni \mathbf{x} \mapsto \mathbf{y} = \mathbf{A} \cdot \mathbf{x} \in \mathbb{R}_{\text{Sp}}^I; \quad y^i = \sum_j \mathbf{a}^i_j x^j,$$

Wenn man zur Auffassung der Tensorrechnung übergehen will, dann fragt man erst einmal, welche Art von Vektoren der Tensor verarbeiten kann: er kann aus $(\boldsymbol{\xi}, \mathbf{x})$ einen Skalar machen, wenn $\boldsymbol{\xi}$ eine I -Zeile und \mathbf{x} eine J -Spalte ist.

$$\mathcal{T}_{\mathbf{A}}: \quad \mathbb{R}_Z^I \otimes \mathbb{R}_{\text{Sp}}^J \ni (\boldsymbol{\xi}, \mathbf{x}) \mapsto \boldsymbol{\xi} \mathbf{A} \mathbf{x} = \sum_{ij} \xi_i \mathbf{a}^i_j x^j \in \mathbb{R}.$$

$\mathcal{T}_{\mathbf{A}}$ ist also ein Element des Tensorprodukts $\mathbb{R}_{\text{Sp}}^I \otimes \mathbb{R}_Z^J$, dem Dualraum von $\mathbb{R}_Z^I \otimes \mathbb{R}_{\text{Sp}}^J$. Wenn \mathbf{A} den Rang 1 hat, wenn also $\mathbf{a}^i_j = z^i \zeta_j$, dann gilt $\mathcal{T}_{\mathbf{A}} = \mathbf{z} \otimes \boldsymbol{\zeta}$ das Tensorprodukt der I -Spalte \mathbf{z} mit der J -Zeile $\boldsymbol{\zeta}$. Die $I \times J$ -Matrizen sind die Linearkombinationen dieser Matrizen vom Rang 1.

Zur Übung formulieren wir hier in der Sprache der Tensoren den wichtigen Satz der Linearen Algebra, welcher besagt, dass Zeilenrang und Spaltenrang einer Matrix gleich sind. Wenn man die Matrix als einen zweistufigen Tensor auffasst, dann ist r der Rang dieses Tensors.

Satz 2.3.3. *Wenn eine $I \times J$ -Matrix A den Rang r hat dann kann man sie als Summe von r Matrizen vom Rang 1 darstellen.*

Beweis. *Wir gehen von dem bekannten Satz aus, dass jede Matrix vom Rang r eine Faktorisierung der folgenden Art besitzt*

$$A = B \cdot \tilde{A} \cdot K \quad \text{mit} \quad \tilde{A} \text{ vom Format } r \times r$$

Den Tensor zur Matrix A können wir nicht nur mit Hilfe der Familie der Einheits-I-Spalten $\{e_i : i \in I\}$ und der Einheits-J-Zeilen $\{e^j : j \in J\}$ als Summe von $|J|$ Tensoren vom Rang 1 schreiben: $\mathcal{T}_A = \sum_{ij} e_i \otimes e^j a^i_j = \sum_j w_j \otimes e^j$ mit $w_j = \sum_i e_i a^i_j$, sondern auch als Summe der Länge r

$$\mathcal{T}_A = \sum_r u_r \otimes \ell^r \quad \text{mit} \quad u_r = \sum_i e_i \left(\sum b^i_m \tilde{a}^m_r \right), \quad \ell^r = \sum_j k^r_j e^j$$

Beispiel 2.3.2. Ein interessanter gemischter Tensor der Stufe 3 ist die Matrizen-Multiplikation. Normalerweise wird man sie als eine bilineare Abbildung $V \times V \longrightarrow V$ verstehen.

Die Interpretation als Tensor erscheint aber bei manchen Fragestellungen angemessener, z. B. bei der Frage nach dem Rang des Tensors. Hier gibt es ein berühmtes Resultat: der Rang nicht so groß, wie man das auf den ersten Blick erwartet.

Betrachten wir den vierdimensionalen Vektorraum der 2×2 -Matrizen. V . Strassen hat bemerkt: Um die vier Einträge in der Produktmatrix $C = A \cdot B$ zu berechnen, braucht man keine acht Multiplikationen durchzuführen, $a_{11}b_{11}$, $a_{12}b_{21}$, $a_{11}b_{12}$, $a_{21}b_{22}$, $a_{21}b_{11}$, Es genügen 7 Multiplikationen. Wenn man E_1, \dots, E_7 wie unten angegeben berechnet, dann ergeben sich die Einträge der Produktmatrix durch einfache Summationen

$$\begin{aligned} c_{11} &= E_1 + E_2 + E_3 + E_4, & c_{22} &= E_1 + E_5 + E_6 - E_7, \\ c_{21} &= E_3 + E_6, & c_{12} &= E_5 - E_4, \quad \text{wobei} \\ E_3 &= a_{22} \cdot (b_{11} + b_{21}), & E_4 &= (a_{11} - a_{12}) \cdot b_{22}, \\ E_5 &= a_{11} \cdot (b_{12} + b_{22}), & E_6 &= (a_{21} - a_{22}) \cdot b_{11}, \\ E_1 &= (a_{11} - a_{22}) \cdot (b_{11} - b_{22}), & E_2 &= (a_{12} - a_{22}) \cdot (b_{21} + b_{22}), \\ & & E_7 &= (a_{11} + a_{21}) \cdot (b_{11} + b_{12}). \end{aligned}$$

Zum Matrizenkalkül

Der Matrixkalkül hat nichts mit geometrischer Anschauung zu tun. Bei den Matrizen liefert die Geometrie keinen universellen Grund für das Hoch- und Tiefstellen der Indizes.

In den Situationen, denen der Anfänger üblicherweise zuerst begegnet, kann man die Spalten als kontragrediente und die Zeilen als cogrediente Vektoren verstehen. Demgemäß tragen bei unseren $I \times J$ -Matrizen die Einträge in j -ten Spalte häufig die Indizes unten, während die Einträge in i -ten Zeile obere Indizes tragen: \mathbf{a}^i_j . Diese Konvention haben wir bereits oben bei den Matrizen zu einem Basiswechsel eingehalten. Auch bei den Jacobi-Matrizen sind wir so verfahren: $\frac{\partial y^i}{\partial x^j} = J^i_j$. Der elementarste Fall ist der, wo die $I \times J$ -Matrizen dazu benützt werden, lineare Abbildungen des Spaltenraums \mathbb{K}^J in den Spaltenraum \mathbb{K}^I -Spalten zu beschreiben.

Ein anderer wichtiger elementarer Verwendungszusammenhang sind die symmetrischen Bilinear-Formen und (im komplexen Fall) die hermiteschen Formen. Der Anfänger begegnet ihnen zuerst bei der Hesse-Matrix.

Wenn man eine symmetrische(!) Matrix zur Darstellung einer quadratischen Funktion auf einem Vektorraum benützt, $q(\mathbf{x}) = \sum_{ij} \mathbf{a}_{ij} \cdot x^i x^j$, dann sollten beide Indizes unten stehen. Die dahinterstehende Vorstellung ist die, dass zur quadratischen Funktion (und zur dazugehörigen symmetrischen Bilinearform $\mathbf{b}(\mathbf{x}, \mathbf{y}) = \sum_{ij} \mathbf{a}_{ij} \cdot x^i y^j$) ein zweifach cogredienter Tensor gehört, der einem Paar von Vektoren eine Zahl zuordnet. Die x^j werden hier als die Werte der 'Koordinaten' eines Vektors $\mathbf{v} = \sum_i x^i \cdot \mathbf{e}_i$ verstanden, $x^j = \langle \ell^j, \mathbf{v} \rangle$. Man könnte daher schreiben

$$\mathbf{b}(\cdot, \cdot) = \sum_{ij} \mathbf{a}_{ij} \ell^i \otimes \ell^j.$$

Einige algebraische Operationen

1) Die Verjüngung: Im Tensorprodukt $\bigotimes_{l \in L} V_l$ sei V_{l_1} der Dualraum von V_{l_2} . Man definiert dann eine bilineare Abbildung $j_{l_1, l_2} : \bigotimes_{l \in L} V_l \longrightarrow \bigotimes_{l \in L \setminus \{l_1, l_2\}} V_l$, indem man für faktorisierte Tensoren festlegt

$$\begin{aligned} j_{l_1, l_2}(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_{l_1} \otimes \cdots \otimes \mathbf{v}_{l_2} \otimes \cdots \otimes \mathbf{v}_k) &= \\ &= \langle \mathbf{v}_{l_1}, \mathbf{v}_{l_2} \rangle \cdot (\mathbf{v}_1 \otimes \cdots \otimes \overline{\mathbf{v}_{l_1}} \otimes \cdots \otimes \overline{\mathbf{v}_{l_2}} \otimes \cdots \otimes \mathbf{v}_k) \end{aligned}$$

Diese Abbildung heisst die Verjüngung um die Positionen l_1, l_2 .

Wenn beispielsweise A eine $I \times I$ -Matrix ist, welche einen Endomorphismus des Spaltenraums \mathbb{R}^I darstellt, $\mathcal{T}_A = \sum_{ik} \mathbf{a}^i_k \mathbf{e}_i \otimes \mathbf{e}^k \in \mathbb{R}_{Sp}^I \otimes \mathbb{R}_Z^I$, dann ist die Verjüngung die Spurbildung $j_{1,2}(\mathcal{T}_A) = \text{trace}(A) = \sum_i \mathbf{a}^i_i$.

2) Faktorisierung in Blöcken:

Sind $W^I = \bigotimes_{l \in L^I} V_l$ und $W^{II} = \bigotimes_{l \in L^{II}} V_l$ Vektorprodukte mit disjunkten Indexmengen L^I, L^{II} , dann definiert man $W^I \otimes W^{II} = \bigotimes_{l \in L^I} V_l \otimes \bigotimes_{l \in L^{II}} V_l = \bigotimes_{l \in L^I + L^{II}} V_l$. Die faktorisierten Tensoren ergeben sich hier also als die Paare faktorisierter Tensoren.

Man kann diese Operation des Zusammenbauens natürlich auch als eine Operation des Zerlegens eines Tensorprodukts in Blöcke verstehen:

Wenn $\bigotimes_{l \in L} V_l$ ein Tensorprodukt zur Indexmenge L ist, und $L = L_1 + \cdots + L_m$ eine

Partition der Indexmenge, so haben wir

$$\bigotimes_{l \in L} V_l = \bigotimes_{l \in L_1} V_l \otimes \bigotimes_{l \in L_2} V_l \otimes \cdots \otimes \bigotimes_{l \in L_m} V_l$$

Formal sieht das ähnlich aus wie bei der Konstruktion cartesischer Produkte oder bei der Konstruktion direkter Summen $\bigoplus_{l \in L}$. Man sollte aber auch die Unterschiede im Auge behalten: bei den cartesischen Produkten multiplizieren sich die Mächtigkeiten, bei den direkten Summen addieren sich die Dimensionen und bei den Tensorprodukten multiplizieren sich die Dimensionen der beteiligten Vektorräume.

3) Identische Faktoren; symmetrische Tensoren

Ein wichtiger Spezialfall eines Tensorprodukts liegt vor, wenn alle Faktoren derselbe Vektorraum V sind: $W = \bigotimes_{l \in L} V_l$ mit $V_l = V$ für alle l . Die Elemente nennt man die Tensoren der k -Stufe, wenn $|L| = k$. Man kann sie als Multilinearformen auf dem Dualraum V^* verstehen. Der Dualraum des k -fachen Tensorprodukts ist das k -fache Tensorprodukt des Dualraums V^* . Seine Elemente heissen die Multilinearformen. Man nennt sie auch cogrediente Tensoren der Stufe k (Der Name k -Form auf V ist vergeben für die alternierenden Multilinearformen vom Grad k , die wir im nächsten Abschnitt diskutieren werden.) Der Begriff des cogredienten Tensors der Stufe k taucht in der Analysis-Vorlesung zuerst bei der sog Taylor-Formel auf. Da haben wir die zusätzliche Eigenschaft der Symmetrie, die wir unten genauer studieren werden. Wir wiederholen kurz die Konstruktion

Satz (Taylorformel):

In der Nähe des Nullpunkts des \mathbb{R}^n sei $F(x^1, x^2, \dots, x^n)$ r -mal stetig differenzierbar. Es existieren dann Koeffizientensysteme $a_i, a_{i_1 i_2}, \dots, a_{i_1 i_2 \dots i_r}$, sodass

$$F(x) - F(0) = \sum_i a_i x^i + \frac{1}{2!} \sum_{i_1 i_2} a_{i_1 i_2} x^{i_1} \cdot x^{i_2} + \cdots + \frac{1}{r!} \sum_{i_1 i_2 \dots i_r} a_{i_1 i_2 \dots i_r} x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_r} + o(\|x\|^r).$$

Der Beweis ergibt sich aus der bekannten Taylor-Approximation einer Funktion $f(t)$ von einer Variablen, nämlich $f(t) = F(tx^1, tx^2, \dots, tx^n)$ (für festes x). Bei den Auswertungen der Ableitungen kommen gemäß der Kettenregel die partiellen Ableitungen von F ins Spiel. Die Summationsindizes laufen allesamt von 1 bis n ; es wird also über $I^k = I \times I \times I \cdots \times I$ summiert. Die Koeffizienten sind die partiellen Ableitungen. Da es bei den partiellen Ableitungen nicht darauf ankommt, in welcher Reihenfolge man sie vornimmt, notiert man bei den partiellen Ableitungen nur die Häufigkeiten α_l , mit welcher nach den verschiedenen Variablen x^l abgeleitet wird ($l = 1, 2, \dots, n$)

$$a_\alpha = a_{i_1 i_2 \dots i_r} = D_\alpha F(0) = \frac{\partial^r}{\partial x^\alpha} F(0) \quad F(x) = \sum_{|\alpha| \leq r} \frac{1}{\alpha!} D_\alpha F(0) \cdot x^\alpha.$$

Hier bedeutet x^α das Monom $x^{i_1} \cdot x^{i_2} \cdot \dots \cdot x^{i_r}$, in welchem der Faktoren x^l α_l -mal vorkommt. Es wird über alle Multiindizes $\alpha = (\alpha_1, \dots, \alpha_n)$ vom Gewicht $|\alpha| = \alpha_1 + \dots + \alpha_n \leq r$

summiert. Der Nenner, welcher den Häufigkeiten Rechnung trägt, ist $\alpha! = \alpha_1! \cdot \dots \cdot \alpha_n!$.

$$\frac{k!}{\alpha!} = \binom{k}{\alpha_1 \dots \alpha_n} \quad (\text{'Multinomialkoeffizient'}).$$

Die Notation bewährt sich ganz allgemein beim Begriff des symmetrischen Tensors.

Definition (Symmetrische Tensoren).

Als eine Abkürzung für das Tensorprodukt $\bigotimes_{l=1}^k V_l$ mit $V_l = V$ für alle l notieren wir $\bigotimes^k V$. Ein Tensor $\mathcal{T} \in \bigotimes^k V$ soll als Multilinearform auf $V^* \times \dots \times V^*$ aufgefasst werden. Ein solcher Tensor heisst ein symmetrischer Tensor, wenn für jedes k -Tupel von Linearformen und jede Permutation π gilt

$$\langle (\ell^1, \dots, \ell^k), \mathcal{T} \rangle = \langle (\ell^{\pi(1)}, \dots, \ell^{\pi(k)}), \mathcal{T} \rangle.$$

Den Vektorraum der symmetrischen Tensoren k -ter Stufe bezeichnen wir mit $\bigotimes_{\text{symm}}^k V$. Wir definieren eine Projektion Π_{symm} von $\bigotimes^k V$ auf diesen Teilvektorraum, welche wir die Symmetrisierung nennen; dabei beschreiben wir die Tensoren wieder durch ihre Auswertungen in den k -Tupeln von Linearformen

$$\langle (\ell^1, \dots, \ell^k), \Pi_{\text{symm}}(\mathcal{T}) \rangle = \frac{1}{k!} \sum_{\pi} \langle (\ell^{\pi(1)}, \dots, \ell^{\pi(k)}), \mathcal{T} \rangle.$$

Für einen faktorisierten Tensor können wir den symmetrisierten Tensor kürzer beschreiben

$$\Pi_{\text{symm}}(\mathbf{v}_1 \otimes \dots \otimes \mathbf{v}_k) = \frac{1}{k!} \sum_{\pi} (\mathbf{v}_{\pi(1)} \otimes \dots \otimes \mathbf{v}_{\pi(k)}).$$

Es sei $\{\mathbf{e}_i : i \in I\}$ eine Basis des n -dimensionalen Vektorraums V . Für den Multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$ vom Gewicht k definieren wir

$$\mathbf{e}_{\alpha} = \Pi_{\text{symm}}(\mathbf{e}_1 \otimes \dots \otimes \mathbf{e}_1 \otimes \mathbf{e}_2 \otimes \dots \otimes \mathbf{e}_n) = \Pi_{\text{symm}} \tilde{\mathbf{e}}_{\alpha},$$

wo im faktorisierten Tensor $\tilde{\mathbf{e}}_{\alpha}$ auf der rechten Seite der Basisvektor \mathbf{e}_l α_l -mal auftritt, ($l = 1, \dots, n$). Nennen wir $I^{(k)}$ die Menge der Multiindizes α vom Gewicht k , so können wir sage: $\{\mathbf{e}_{\alpha} : \alpha \in I^{(k)}\}$ ist eine Basis des Raums $\bigotimes_{\text{symm}}^k V$ der symmetrischen Tensoren der Stufe k .

Wenn wir dieselben Konstruktionen für den Dualraum V^* durchführen, dann erhalten wir eine eine Basis für den Dualraum $(\bigotimes_{\text{symm}}^k V)^* = \bigotimes_{\text{symm}}^k V^*$. Die \mathbf{e}^{α} sind bis auf Faktoren die Elemente der dualen Basis. Es gilt

$$\langle \mathbf{e}^{\alpha}, \mathbf{e}_{\beta} \rangle = \begin{cases} 0 & \text{für } \beta \neq \alpha \\ \frac{\alpha!}{k!} & \text{für } \beta = \alpha \end{cases}$$

Betrachten wir nochmals die Polynome in n Unbestimmten x^i . Es spricht einiges dafür, ein solches Polynom als eine symmetrische Multilinearform zu verstehen, ähnlich, wie man ein quadratisches Polynom gerne mit einer symmetrischen Bilinearform identifiziert. Ein homogenes Polynom vom Grad 2 schreibt man klassisch $p(x) = \frac{1}{2!} \sum_{ij} a_{ij} x^i x^j$. Dabei kann man annehmen, dass die ‘Koeffizientenmatrix’ symmetrisch ist; andernfalls ersetzt man die a_{ij} durch $\frac{1}{2}(a_{ij} + a_{ji})$. Die symmetrische Koeffizientenmatrix beschreibt eine symmetrische Bilinearform. Entsprechend haben die Polynome, die homogen vom Grad k sind, die Form

$$p(x) = \frac{1}{k!} \sum_{i_1 \dots i_k} a_{i_1 \dots i_k} (x^{i_1} \cdot \dots \cdot x^{i_k}) \quad \text{mit einem symmetrischen Koeffizientenschema.}$$

Die Symmetrie des Koeffizientenschemas bedeutet hier, dass die Koeffizienten nur vom Multiindex α abhängen, dessen Einträge α_l angeben, wie oft die Variable x^l im Monom $x^{i_1} \cdot \dots \cdot x^{i_k} = x^\alpha$ vorkommt. ($l = 1, \dots, n$). $a_{i_1 \dots i_k} = a_\alpha$. Zum Multiindex α gibt es $\binom{k}{\alpha_1 \dots \alpha_n} = \frac{k!}{\alpha!}$ k -Tupel $(i_1 \dots i_k)$. Die Aufschlüsselung der Summe nach den Multiindizes ergibt daher

$$p(x) = \frac{1}{k!} \sum_{i_1 \dots i_k \in I^k} a_{i_1 \dots i_k} (x^{i_1} \cdot \dots \cdot x^{i_k}) = \sum_{\alpha \in I^{(k)}} \frac{1}{\alpha!} a_\alpha \cdot x^\alpha.$$

Die rechte Seite kann (und sollte man gelegentlich) als eine symmetrische Multilinearform verstehen, indem man x^α durch den symmetrisierten Tensor e^α ersetzt. Die Koeffizienten sind eindeutig bestimmt; denn die e^α bilden eine Basis des Raums $\bigotimes_{\text{symm}}^k V$ der symmetrischen Multilinearformen. Die Koeffizienten ergeben sich durch die Auswertung in den Basiselementen e_β .

$$\left\langle \sum_{\alpha \in I^{(k)}} \frac{1}{\alpha!} a_\alpha \cdot e^\alpha, e_\beta \right\rangle = \frac{1}{k!} a_\beta$$

‘Anschauung’. Klassische Umstände

Die Tensoren zweiter Stufe können offenbar in allerlei Varianten als anschauliche mathematische Objekte gelten. Schwieriger ist es, intuitive anschauliche Vorstellungen von den Tensoren höherer Stufe zu gewinnen. Die abstrakte Algebra bietet kaum Ansatzpunkte; man braucht Geometrie oder etwas Physik als Hintergrund.

Die Physiker kannten bis vor wenigen Jahrzehnten vor allem die symmetrischen Tensoren; insbesondere in der Kontinuumsmechanik spielen sie seit langem eine Rolle. Bei den Physikern ist die Idee allerdings traditionell an den dreidimensionalen Anschauungsraum geknüpft. Diese traditionelle Limitierung der Anschauung ist in der heutigen Mathematikerausbildung ein didaktisches Problem. Die Anschaulichkeit in der Mathematik darf heute nicht mehr auf die traditionelle geometrische Anschaulichkeit reduziert werden. Die Studierenden müssen darauf hingeführt werden, dass es nicht nur *einen* anschaulichen Raum gibt, und dass die für alle Mathematik wichtige Anschaulichkeit nicht überall und

nicht von vorneherein für jeden dasselbe bedeutet. 'Anschauung' muß dringlich auf immer weitere Bereiche der Mathematik ausgedehnt werden.

Man kann Beispiele nennen: In gewissem Sinne anschaulich können dem einen oder anderen Studierenden die affinen und die projektiven Räume über beliebigen Körpern werden, die heute in der Anfängervorlesung zur Linearen Algebra in den Vordergrund gestellt werden. Neben der Anschauung zur dreidimensionalen euklidischen Geometrie sollte aber in der heutigen Linearen Algebra noch dringlicher ein Platz offengehalten werden für den vierdimensionalen 'Anschauungsraum' der speziellen Relativitätstheorie, das Raum-Zeit-Kontinuum; die Elektrodynamik hat bekanntlich diesem 'anschaulichen' Raum viel zu verdanken. Hier in der Analysis III müssen wir nun noch einen etwas anderen Schritt gehen. Wir müssen Anschauung entwickeln für die Mannigfaltigkeiten und schliesslich auch für die 'gekrümmten' Räume. Übrigens: Auch eine Theorie wie die Hamilton'sche Mechanik erfordert eine erweiterte Form der Anschauung; für einen Einstieg in die Vorstellungswelt der symplektischen Mannigfaltigkeiten werden wir allerdings in dieser Veranstaltung keine Zeit mehr finden.

Hier noch einige Andeutungen zur Vorgeschichte der allgemeinen Tensorrechnung. Von Tensoren im Sinne der Infinitesimalrechnung reden die Physiker schon sehr lange. Das Wort kommt aus der Elastizitätstheorie. (Es ist abgeleitet aus dem lateinischen Wort 'tendere' = spannen, oder 'tensus' = gespannt.) Wenn auf ein elastisches Medium Kräfte wirken, dann entstehen einerseits Spannungen und andererseits Verformungen. Die Verformung ist als ein (ortsabhängiger) zweifach kontragrredienter symmetrischer Tensor zu deuten; entsprechend den elastischen Eigenschaften des Materials steht sie in Korrespondenz zur Spannung im betreffenden Punkt, einem zweifach cogredienten symmetrischen Tensor. Verantwortlich für diese Korrespondenz ist ein Tensor vierter Stufe, der (bei kleinen Kräften in approximativ linearer Weise) der Spannung die Verformung zuordnet; dieser Tensor ist eine Eigenschaft des elastischen Materials. Die beiden aus den zueinander dualen Vektorräumen stammenden Tensoren ergeben zusammen die 'infinitesimal' geleistete Verformungsarbeit.— Es handelt sich um eine Weiterentwicklung der Formel *Arbeit ist Kraft mal Weg*. So wie Arbeit geleistet wird, wenn man (auf einer Kurve) eine Verschiebung gegen ein Kraftfeld bewerkstelligt, so wird Arbeit geleistet, wenn man eine Verformung gegen eine Spannung aufbaut.

Ähnlich ist die Lage in der sog. Elektrostatik für ein dielektrisches Material. Man geht davon aus, dass die elektrische Spannung eine Polarisierung hervorruft, und zwar gemäß dem Dielektrizitätstensor ϵ . Die Spannung \mathcal{E} ist vom Typ eines Kovektors (aus der Potentialdifferenz abgeleitet), die dielektrische Verschiebung \mathcal{D} ist vom Typ eines Vektors, einer Art Verformung. Das Produkt liefert die elektrische Energiedichte. Nur im 'isotropen' Fall kann man (in irgendeinem Sinne) sagen, dass die Verformung in dieselbe Richtung weist wie die Spannung; nur in einem isotropen Medium ist es mit einer Dielektrizitätskonstanten getan. In allgemeineren Medien ist der Dielektrizitätstensor verantwortlich für den Zusammenhang. Mehr über klassischphysikalische Aspekte des Tensorbegriffs (sowohl in der nichtrelativistischen als auch in der relativistischen Welt) findet man in den berühmten Feynman-Lectures, Band II, Kapitel 31, 'Tensoren'.

Kommen wir zur Geometrie in einem traditionelleren Sinn. Im Minkowski-Raum hat man wie im euklidischen Raum ein symmetrisches inneres Produkt zwischen Vektoren, und das bedeutet im Sinne der Tensorrechnung einen Isomorphismus zwischen dem Raum der Vektoren und dem Raum der Covektoren. Dies ist der Ansatz für die (semi-)Riemann'sche Geometrie. Ausschlaggebend ist da ein (ortsabhängiger!) (semi-)metrischer Tensor.

Hier ist ein Platz für eine kleine Polemik: Durch die in der Schule gelehrt Anschauung zum euklidischen Raum ist manchen Studierenden der metrische Tensor des euklidischen Raums so eingeebnet, dass sie ihn gar nicht mehr wegdenken können; unausgesprochen (und das ist schlecht für einen Mathematiker!) nimmt man eine Identifizierung der Vektoren mit den Covektoren vor. Das drückt sich z. B. darin aus, dass man den Anstieg einer affinen Funktion durch einen Vektor(!) beschreibt, den sog. Gradienten, der senkrecht (!) auf den Niveaulinien steht. Manchem Studierenden fällt es nach einer Einweisung in diesem Sinn schwer, intuitiv einzusehen, dass die Auswertung einer Linearform in einem Vektor (logisch gesehen) etwas anderes ist als das euklidische innere Produkt zweier Vektoren. Wir schliessen daraus: Bei der Behandlung der Tensoren fällt der mathematischen Anfängervorlesung auch die Aufgabe zu, einer kurzschlüssigen von der euklidischen Geometrie dominierten Anschauung entgegenarbeiten, mit dem Ziel, eine differenziertere Anschauungswelt aufzubauen. Als allererstes gilt es u. E. , den Unterschied zwischen den Vektoren und den Covektoren herauszuarbeiten. Der Schlüsselbegriff ist u. E. die Vektorraumdualität.; die Tensorrechnung muss darauf aufbauen. Die Auffassung ist allerdings kontrovers, wie wir unten sehen werden.

Dies ist keine Absage an die Vorstellungsweisen der klassischen Geometrie; es geht um eine Bereinigung. Die weiterführenden Vorstellungen von den Tangentialvektoren und den Covektoren auf einer Mannigfaltigkeit schliessen sehr wohl an an die Vorstellungen der klassischen euklidischen Geometrie. Das zeigt sich in dem, was wir schon über Mannigfaltigkeiten gelernt haben. Die Tangentialvektoren entsprechen den 'echten' Vektoren, die (irgendwie) mit Verschiebungen zu tun haben. Ihnen gegenüber stehen die Covektoren, die in jedem Punkt den Anstieg einer Funktion beschreiben. Wir haben bereits gesehen: Von den Tangentialfeldern gelangt man zu Kurven; die Cotangentialfelder werden entlang von Kurven integriert. Es wird weiter gehen: k -Formen werden über k -dimensionale Bereiche integriert. Im Folgenden werden wir weitere Ansatzpunkte für eine flexible Raumschauung finden; wir werden die Anschauung in differenzierter Weise von den Besonderheiten der euklidischen Geometrie lösen.

In einem kurzen Anhang wollen wir nun noch eine (der unseren entgegenstehenden) Auffassung von den Tensoren skizzieren, mit welcher es möglicherweise mancher Leser an anderer Stelle (unter dem Stichwort Ricci-Kalkül) zu tun bekommt.

Exkurs: Das Herauf- und Herunterschaffen der Indizes bzgl. eines metrischen Tensors

Der Tensorkalkül wurde in der ersten Hälfte des 20. Jahrhunderts ganz auf Koordinatendarstellungen gegründet. Eine wichtige Rolle spielt das Hoch- und Tiefstellen der Indizes. Die maßgeblichen Impulse kamen aus den berühmten Vorlesungen über allgemeine Relativitätstheorie, die Hermann Weyl zum ersten Mal 1917 an der ETH Zürich gehalten

hat. (H. Weyl: Raum Zeit Materie, in sechster Auflage 1970 im Springer-Verlag). Der von Weyl und Einstein entwickelte Formalismus beherrscht bis heute die Lehrbücher, und soll deshalb hier kurz vorgestellt werden, obwohl wir ihn nicht befürworten. —

Der Anfänger sollte den Exkurs überspringen.

Der Sachverhalt, auf den sich der Formalismus bezieht, entspricht, wenn wir hier die Physik ausblenden, dem folgenden Szenario: Gegeben ist ein reeller Vektorraum V mit einem inneren Produkt $(\cdot | \cdot)$ auf $V \times V$, welches man als den metrischen Tensor auf V verstehen sollte. Das innere Produkt ist symmetrisch, im euklidischen Fall positiv definit, in der speziellen Relativitätstheorie vierdimensional mit der Signatur $(1, 3)$.

Wenn eine Basis $\{\mathbf{u}_i : i \in I\}$ gewählt wurde, dann nennt man das Zahlenschema $g_{ik} = (\mathbf{u}_i | \mathbf{u}_k)$ die cogredienten Koeffizienten des metrischen Tensors. Für einen Vektor $\mathbf{v} = \sum_i \mathbf{a}^i \mathbf{u}_i$ nennt man den Skalar \mathbf{a}^i die i -te kontragrediente Komponente des Vektors (Der Index muss hochgestellt werden!).

Es sei nun $\{\mathbf{u}^i : i \in I\}$ die duale Basis (sodass also $\mathbf{u}^i(\mathbf{v}) = \mathbf{a}^i$). Die Hochstellung des Index ist essentiell. Wenn man die Linearform $\ell(\cdot) = (\mathbf{v} | \cdot)$ in dieser dualen Basis darstellen will, dann braucht man dazu die sog. cogredienten Komponenten des Vektors \mathbf{v} , die man jetzt mit tiefgestellten Indizes auflistet.

$$\left(\sum_i \mathbf{a}^i \cdot \mathbf{u}_i | \cdot \right) = \sum_k \mathbf{a}_k \cdot \mathbf{u}^k(\cdot) \quad \text{mit} \quad \mathbf{a}_k = \sum_i g_{ik} \mathbf{a}^i.$$

In der Tat liefern diese \mathbf{a}_k für jeden Vektor $\mathbf{w} = \sum_k \mathbf{b}^k \cdot \mathbf{u}_k$ das richtige Ergebnis

$$(\mathbf{v} | \mathbf{w}) = \left(\sum_i \mathbf{a}^i \mathbf{u}_i | \sum_k \mathbf{b}^k \mathbf{u}_k \right) = \sum_{ik} \mathbf{a}^i \mathbf{b}^k g_{ik} = \sum_k \mathbf{a}_k \mathbf{b}^k = \sum_k \mathbf{a}_k \cdot \mathbf{u}^k(\mathbf{w}).$$

Im Tensorkalkül, den man in der Relativitätstheorie benützt, bleiben die gewählten Basen im Hintergrund, sie tauchen in der Notation nicht auf. Der Vektor \mathbf{v} wird alternativ mit \mathbf{a}^i oder mit \mathbf{a}_k dargestellt, wobei i und k als Summationsindizes die vorgegebene Basis (oder die Dualbasis) indizieren, und die Zahlenwerte, die cogredienten und kontragredienten Komponenten des Vektors \mathbf{v} in der oben angegebenen Weise ineinander umzurechnen sind. Die Umrechnungsprozedur nennt man das Herauf- und Herunterschaffen des Index (*raising and lowering the index of a tensor*) Die Notation unterdrückt nicht nur die \mathbf{u}_i und die \mathbf{u}^i , sondern auch das Summenzeichen, und man schreibt

$$\mathbf{a}_i = \mathbf{a}^k g_{ik}, \quad g^{ki} \mathbf{a}_i = \mathbf{a}^k.$$

Das Summenzeichen wird auch in allgemeineren Situationen mit mehrfach indizierten Größen unterdrückt, gemäß der sog. Einstein-Konvention, welche sagt: Wenn irgendwo in einem Aggregat von indizierten Größen ein Index sowohl oben als auch unten auftritt, dann wird über ihn summiert.

Bei einem dreistufigen Tensor sieht das Herunterziehen so aus: $\mathbf{a}^{ijk} g_{jl} = \mathbf{a}^i{}_l{}^k$. Der Buchstabe \mathbf{a} bleibt unverändert, weil auf beiden Seite der Gleichung derselbe Tensor gemeint ist. Übrigens: Wenn wir beim metrischen Tensor selbst einen Index hinaufziehen,

dann erhalten wir das Kroneckersymbol $g_{ij}g^{ik} = g_j^k = \delta_j^k$ ('Kroneckersymbol'). Anders gesagt: die co-kontragredienten Komponenten des metrischen Tensors sind das Kroneckersymbol.

Auffallend ist, dass in diesem Tensorkalkül der Unterschied zwischen den kontragredienten und den cogredienten Vektoren (und Tensoren) mittels des metrischen Tensors zum Verschwinden gebracht wird. An die Stelle dieser (in anderen Zusammenhängen nach unserer Meinung sehr wichtigen) Unterscheidung tritt die Unterscheidung zwischen den kontra- und cogredienten Komponenten desselben Vektors (bzw. Tensors).

H.Weyl schreibt auf Seite 37 des erwähnten Buchs:

Es ist vielfach versucht worden, in unserm Gebiet eine solche invariante, mit den Tensoren selbst und nicht mit ihren Komponenten arbeitende Bezeichnungsweise auszubilden, wie sie in der Vektorrechnung besteht. Was aber dort am Platze ist, erweist sich für den viel weiter gespannten Rahmen des Tensorkalküls als äußerst unzweckmäßig. Es werden eine solche Fülle von Namen und ein solcher Apparat von Rechenregeln nötig (wenn man nicht doch immer wieder auf die Komponenten zurückgreifen will), dass damit ein Gewinn von sehr erheblichem negativem Betrag erreicht wird. Man muß gegen diese Orgien des Formalismus, mit dem man heute sogar die Techniker zu belästigen beginnt, nachdrücklich protestieren.

Die Zeit geht weiter. Der traditionelle Tensorkalkül hat sich zweifellos in wichtigen Bereichen bewährt. In der reinmathematischen Tensorrechnung wird man jedoch auf einer klaren Unterscheidung von Vektoren und Covektoren bestehen müssen, obwohl es von vorneherein keine Gründe gibt, den einen oder den anderen Partner in einem Dualitätspaar heraufzuheben. Eine 'Identifizierung' der Covektoren mit den Vektoren kommt u. E. nicht in Betracht, wenn man kein ausgezeichnetes inneres Produkt auf $V \times V$ besitzt.

Kontroverse didaktische Positionen

Da wir nun schon einmal bei didaktischen Fragen sind, möchte ich mein Plädoyer für die Vektorraumdualität in den Zusammenhang der Anfängervorlesung stellen: Nach meiner Auffassung sollte man mit den Anfängern die elementaren Beispiele für Dualitätspaare sauber auseinanderhalten

- Die Matrizenrechnung sieht vor, eine J-Zeile \mathbf{t} mit einer J-Spalte \mathbf{s} zu 'multiplizieren', sodass das Resultat ein Skalar ist: $\langle \mathbf{t}, \mathbf{s} \rangle = \mathbf{t} \cdot \mathbf{s}$. Die Matrizenrechnung sieht nicht vor, zwei Spalten (oder zwei Zeilen) miteinander zu multiplizieren. Es ist sehr wohl plausibel, $\mathbf{s} \cdot \mathbf{t}$ als eine $J \times J$ -Matrix (vom Rang 1) zu betrachten. Allgemeiner versteht man das matrizentheoretische Produkt einer I-Spalte \mathbf{y} mit einer J-Zeile ξ , als eine $I \times J$ -Matrix vom Rang 1.

- Die Theorie des euklidischen Raums (wie auch die Theorie des Minkowski-Raums) kennt das Skalarprodukt zweier Vektoren. Es handelt sich um eine symmetrische bilineare Operation $(\cdot | \cdot)$ auf $V \times V$.
- Für die Theorie des Hilbert-Raums sind die Bezeichnungen von P. Dirac zu empfehlen. Das ‘innere’ Produkt zweier Vektoren im Hilbertraum wird als das Produkt eines ‘bra-Vektors’ $\langle w |$ mit einem ‘ket-Vektor’ $|v \rangle$ verstanden. (‘bracket’ für Klammer) Es handelt sich um eine sesquilineare Operation $\langle \cdot | \cdot \rangle$ auf $V \times V$. In der Matrizenrechnung entsteht der bra-Vektor zu einer komplexen Spalte, einem ket-Vektor, als die Zeile mit den konjugierten Einträgen, entsprechend der hermiteschen Konjugation der komplexen Matrizen.

Alle drei Strukturen gehören u. E. in die Anfängerausbildung. Die (mehr oder weniger delikaten) Gemeinsamkeiten dieser Formalismen und Notationen sollten aber kein Thema der ersten Anfängervorlesungen sein. Euklidische Geometrie, Hilbertraum-Geometrie und das Rechnen mit Zeilen und Spalten sollten zunächst einmal als getrennte Bereiche mit Blick auf spezifische Anwendungen entwickelt werden. Auf dem Niveau der Anfängervorlesungen sind nicht die Gemeinsamkeiten wichtig, sondern die verschiedenartigen Ausrichtungen auf Anwendungen.

Die Vermischung mit der dreidimensionalen euklidischen Geometrie in der traditionellen ‘Vektoranalysis’ ist u. E. für die Mathematikerausbildung abzulehnen, weil sie die Idee der Vektorraumdualität konterkariert..

2.4 Die Algebra der schiefsymmetrischen Formen

Die Determinanten als alternierende Multilinearformen vom Grad n auf einem n -dimensionalen \mathbb{K} -Vektorraum werden in der Grundvorlesung Lineare Algebra ausführlich studiert. Die schiefsymmetrischen (oder ‘alternierenden’) k -Formen für $k \leq n$ werden jedoch selten mit der Gründlichkeit behandelt, die wir für die geometrische Betrachtung der Mannigfaltigkeiten benötigen. Sie soll deshalb hier kurz vorgestellt werden. Dabei verlangen wir vom Grundkörper \mathbb{K} nur, dass die Charakteristik $\neq 2$ ist ($1 + 1 \neq 0$).

Definition 2.2 (k -Vektoren über V).

Es sei V ein n -dimensionaler \mathbb{K} -Vektorraum. Für $k = 1, 2, \dots, n$ definieren wir einen Vektorraum $\bigwedge^k V$ (wobei $\bigwedge^1 V = V$), den Vektorraum der k -Vektoren über V .

Die Elemente gewinnt ausgehend von den formale Linearkombinationen von k -fachen „Dachprodukten“ $w_1 \wedge \dots \wedge w_k$, die zunächst einfach als neuartige Symbole zu betrachten sind. Solche formalen Linearkombinationen

$$\sum_{\alpha} a_{\alpha} \cdot (w_{\alpha_1} \wedge \dots \wedge w_{\alpha_k}) \quad \text{und} \quad \sum_{\beta} a_{\beta} \cdot (w_{\beta_1} \wedge \dots \wedge w_{\beta_k})$$

sind als gleich zu betrachten, wenn sie mit Hilfe der folgenden Regeln ineinander umgerechnet werden können

(i) Für jedes k -Tupel von Vektoren und jede Permutation $(\sigma_1, \dots, \sigma_k)$ gilt

$$1 \cdot (w_{\sigma_1} \wedge \dots \wedge w_{\sigma_k}) = (\text{sign} \sigma) \cdot (w_1 \wedge \dots \wedge w_k)$$

(ii) für alle $\lambda, \mu \in \mathbb{K}$ gilt

$$\lambda \cdot (\mu w_1 \wedge w_2 \wedge \dots \wedge w_k) = 1 \cdot (\lambda \mu w_1) \wedge w_2 \wedge \dots \wedge w_k$$

(Statt $1 \cdot (w_1 \wedge \dots \wedge w_k)$ schreiben wir auch einfach $w_1 \wedge \dots \wedge w_k$)

(iii) $(w'_1 + w''_1) \wedge w_2 \wedge \dots \wedge w_k = w'_1 \wedge w_2 \wedge \dots \wedge w_k + w''_1 \wedge w_2 \wedge \dots \wedge w_k$.

Die Äquivalenzklassen nennt man die k -Vektoren über V . Diejenigen Äquivalenzklassen, in welchen es einem Repräsentanten gibt, welcher ein Dachprodukt ist, nennt man die faktorisierten k -Vektoren. Wir werden eine Besonderheit des Grades $k = n - 1$ kennenlernen: man kann jeden $n - 1$ -Vektor faktorisieren, Für Grade k mit $2 \leq k \leq n - 2$ ist das nicht der Fall.

Didaktische Bemerkung zum Muster der Definition:

Wir definieren hier die Objekte nach dem bewährten Muster, dass wir zunächst auf eine mehr oder weniger triviale Weise eine recht uninteressante Menge gewinnen und dann durch eine durchaus nichttriviale Äquivalenzrelation zu den interessanten Objekten gelangen. Die Menge dieser Objekte hat hier die Struktur eines \mathbb{K} -Vektorraums.

Das Muster haben wir beispielsweise auch angewandt, als wir von der unüberschaubaren, recht uninteressanten Menge der Cauchy-Folgen (in einem metrischen Raum) zur Menge der Äquivalenzklassen übergingen, die wir dann als die Vervollständigung des Raums deuten konnten. Die konstruierte Menge der Äquivalenzklassen hat die Struktur eines metrischen Raums, der sich als vollständig erweist.

Das Muster haben wir auch angewandt, als wir von der unüberschaubaren, recht uninteressanten Menge aller geschlossenen Kurven (in einem topologischen Raum) zur Menge der Homotopieklassen übergingen. Die konstruierte Menge hat hier die Struktur einer Gruppe.

Wer den Verdacht hegt, dass man möglicherweise jede formale Linearkombination von Dachprodukten in das Nullelement umrechnen kann, wird durch die folgenden Rechnungen eines Besseren belehrt. Wir werden sehen, dass $\bigwedge^k V$ ein $\binom{n}{k}$ -dimensionaler Vektorraum ist. Man kann also $\binom{n}{k}$ linear unabhängige Elemente finden und mit diesen kann man dann jedes Element von $\bigwedge^k V$ in eindeutiger Weise linear kombinieren. Bei der Wahl der Basis hat man (wie immer in Vektorräumen) große Freiheiten. Die beliebtesten Basen von $\bigwedge^k V$ sind die, die von einer Basis von V herrühren. Diese Basen haben enge Beziehungen zum Kalkül der Determinanten.

Beispiel.

V sei der Vektorraum der Verschiebungen des dreidimensionalen euklidischen Anschauungsraums. $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ sei eine Orthonormalbasis.

Jedes Dachprodukt $\mathbf{a} \wedge \mathbf{b}$ läßt sich auf eindeutige Weise als Linearkombination der speziellen Dachprodukte

$$\mathbf{e}_2 \wedge \mathbf{e}_3, \mathbf{e}_3 \wedge \mathbf{e}_1, \mathbf{e}_1 \wedge \mathbf{e}_2 \text{ darstellen .}$$

Dies sieht man folgendermaßen

$$\begin{aligned} \mathbf{a} &= a^1 \cdot \mathbf{e}_1 + a^2 \cdot \mathbf{e}_2 + a^3 \cdot \mathbf{e}_3 \\ \mathbf{b} &= b^1 \cdot \mathbf{e}_1 + b^2 \cdot \mathbf{e}_2 + b^3 \cdot \mathbf{e}_3 \\ \mathbf{a} \wedge \mathbf{b} &= (a^1 \cdot \mathbf{e}_1 + a^2 \cdot \mathbf{e}_2 + a^3 \cdot \mathbf{e}_3) \wedge (b^1 \cdot \mathbf{e}_1 + b^2 \cdot \mathbf{e}_2 + b^3 \cdot \mathbf{e}_3) \\ &= \alpha \cdot \mathbf{e}_2 \wedge \mathbf{e}_3 + \beta \mathbf{e}_3 \wedge \mathbf{e}_1 + \gamma \cdot \mathbf{e}_1 \wedge \mathbf{e}_2 \text{ mit} \\ \alpha &= \begin{vmatrix} a^2 & b^2 \\ a^3 & b^3 \end{vmatrix} ; \quad \beta = \begin{vmatrix} a^3 & b^3 \\ a^1 & b^1 \end{vmatrix} ; \quad \gamma = \begin{vmatrix} a^1 & b^1 \\ a^2 & b^2 \end{vmatrix} \end{aligned}$$

Hinweise

- Wir bemerken, dass (α, β, γ) die Koeffizienten im traditionellen Vektorprodukt $\mathbf{a} \times \mathbf{b}$ sind (vorausgesetzt, die Aufzählung der Basisvektoren ist positiv orientiert) . Die Physiker definieren bekanntlich in der Tradition der „Vektoranalysis“:

$$\mathbf{a} \times \mathbf{b} = \alpha \cdot \mathbf{e}_1 + \beta \cdot \mathbf{e}_2 + \gamma \cdot \mathbf{e}_3 .$$

Man beachte aber, dass die Ersetzung

$$\mathbf{e}_2 \wedge \mathbf{e}_3 \longleftrightarrow \mathbf{e}_1, \quad \mathbf{e}_3 \wedge \mathbf{e}_1 \longleftrightarrow \mathbf{e}_2, \quad \mathbf{e}_1 \wedge \mathbf{e}_2 \longleftrightarrow \mathbf{e}_3$$

an die euklidische Struktur (und die Orientierung) gebunden ist. In allgemeinen dreidimensionalen Vektorräumen gibt es keinen irgendwie natürlichen Isomorphismus von V und $\wedge^2 V$.

2. Auch die Physiker wissen natürlich, dass die Elemente von $\wedge^2 V$ etwas anderes sind als die üblichen Vektoren; wenn sie genau sein wollen, nennen sie die Elemente von $\wedge^2 V$ „polare Vektoren“, im Gegensatz zu den „axialen“ Vektoren $v \in V$.

Satz 2.4.1. Spezielle Basen von $\wedge^k V$

Sei $\{v_i; i \in I\}$ eine Basis von V und $\{v_1, \dots, v_n\}$ eine Aufzählung der Basisvektoren. Für $K \subseteq \{1, 2, \dots, n\}$ mit der Mächtigkeit k definieren wir

$$v_K := v_{n_1} \wedge v_{n_2} \wedge \dots \wedge v_{n_k},$$

wobei $1 \leq n_1 < n_2 < \dots < n_k \leq n$ die Elemente von K sind. Jedes $p \in \wedge^k V$ besitzt dann eine Darstellung $p = \sum v_K \cdot a^K$, wobei die Summe über alle aufsteigend angeordneten k -Teilmengen $K \subseteq \{1, 2, \dots, n\}$ zu erstrecken ist.

Beweis

1. Es genügt, die Dachprodukte $p = w_1 \wedge \dots \wedge w_k$ zu untersuchen. Die w_j stellen wir mit Hilfe der gegebenen Basis dar: $w_j = \sum_i v_i \cdot a^i_j$ für $j = 1, \dots, k$. Die Koeffizienten bilden eine $n \times k$ -Matrix A mit angeordneten Reihen.
2. $A^K = A^{\{n_1, \dots, n_k\}}$ sei die $k \times k$ -Matrix, die wir aus A gewinnen, indem wir alle diejenigen Zeilen streichen, die nicht zu K gehören und nur die Zeilen n_1, \dots, n_k stehen lassen. Wir notieren für $K = \{n_1 < n_2 < \dots < n_k\}$

v_1	$a^1_{n_1}$	\dots	$a^1_{n_k}$
	\vdots		\vdots
	//// //// ////		
	\vdots		\vdots
	\vdots		\vdots
	//// //// ////		
	//// //// ////		
v_n	$a^{n_1}_{n_1}$	\dots	$a^{n_1}_{n_k}$
	\vdots		\vdots
	$a^{n_k}_{n_1}$	\dots	$a^{n_k}_{n_k}$
	$= w_1$	\dots	$= w_k$

$$a^K = \det A^K = \sum_{\sigma} (\text{sign} \sigma) \cdot a^{n_1}_{\sigma_1} \cdot a^{n_2}_{\sigma_2} \cdot \dots \cdot a^{n_k}_{\sigma_k}$$

wobei die Summe für alle Permutationen von K zu erstrecken ist.

3. Nach unseren Umrechnungsregeln (i), (ii), (iii) gilt

$$\begin{aligned} w_1 \wedge \dots \wedge w_k &= \left(\sum v_{i_1} \cdot a^{i_1}_1 \right) \wedge \left(\sum v_{i_2} \cdot a^{i_2}_2 \right) \wedge \dots \wedge \left(\sum v_{i_k} \cdot a^{i_k}_k \right) \\ &= \sum_{i_1, \dots, i_k} (v_{i_1} \wedge v_{i_2} \wedge \dots \wedge v_{i_k}) \cdot a^{i_1}_1 \cdot a^{i_2}_2 \cdot \dots \cdot a^{i_k}_k. \end{aligned}$$

Die Summanden, in welchen zwei Summationsindizes gleich sind, liefern keinen Beitrag. Die anderen sortieren wir nach den k -Teilmengen K ; und wir bringen die Indizes in aufsteigenden Reihenfolge.

4. Somit ergibt sich

$$w_1 \wedge \dots \wedge w_k = \sum a^K \cdot v_K \quad \text{q.e.d.}$$

Die Determinanten a^K nennt man die $k \times k$ -Minoren der Matrix A . Wer den Verdacht hegt, dass man möglicherweise manche Summen $\sum b^K \cdot v_K$ durch geschickte Anwendung der Rechenregeln in das Nullelement umrechnen kann, wird durch die folgende Überlegung eines Besseren belehrt. Wir werden sehen, dass jede formale Summe von k -fachen Dachprodukten $(w_{\alpha_1} \wedge \dots \wedge w_{\alpha_k})$ genau eine Darstellung

$$p = \sum a^K \cdot v_K \quad \text{besitzt.}$$

Mit anderen Worten: Die v_K sind linear unabhängig im Vektorraum $\bigwedge^k V$.

Definition 2.3 (Alternierende Multilinearformen).

Eine alternierende k -Form (für den n -dimensionalen \mathbb{K} -Vektorraum V) ist eine \mathbb{K} -wertige Funktion $\Phi(\cdot)$ auf dem kartesischen Produkt $V \times V \times \dots \times V$ mit den Eigenschaften

i*) $\Phi(w_{\sigma_1}, \dots, w_{\sigma_k}) = (\text{sign} \sigma) \cdot \Phi(w_1, \dots, w_k)$ für alle k -Tupel w_1, \dots, w_k und alle Permutationen σ .

ii*) $\Phi(\lambda w_1, w_2, \dots, w_k) = \lambda \cdot \Phi(w_1, w_2, \dots, w_k)$ für alle $\lambda \in \mathbb{K}$.

iii*) $\Phi(w'_1 + w''_1, w_2, \dots, w_k) = \Phi(w'_1, w_2, \dots, w_k) + \Phi(w''_1, w_2, \dots, w_k)$.

Den \mathbb{K} -Vektorraum der alternierenden k -Formen bezeichnet man mit $\bigwedge^k V^*$.

Für die Grade $k = 1$ und $k = n$ sind diese Vektorräume wohlbekannt. Die alternierenden 1-Formen sind die Linearformen, es gilt also $\bigwedge^1 V^* = V^*$. Die alternierenden n -Formen sind aus der Theorie der Determinanten bekannt. Man nennt sie die Volumenformen; sie bilden einen eindimensionalen Vektorraum. . Es ist üblich auch $\bigwedge^0 V^*$ zu definieren; man identifiziert die Elemente mit den Skalaren.

Aus dem bekannten Satz, dass die Determinante einer Matrix das Vorzeichen ändert, wenn man Reihen vertauscht, ergibt sich der

Satz 2.4.2 (Faktorierte k -Formen).

Jedes k -Tupel von Linearformen $\ell^{(1)}(\cdot), \dots, \ell^{(k)}(\cdot)$ liefert eine alternierende k -Form Φ vermöge der folgenden Konstruktion:

$$\Phi(w_1, \dots, w_k) = \det(\langle \ell^{(i)}, w_j \rangle) \text{ für alle } w_1, \dots, w_k.$$

Man nennt eine solche k -Form eine faktorisierbare k -Form und notiert

$$\Phi = \ell^{(1)} \wedge \dots \wedge \ell^{(k)}.$$

Dass diese Bezeichnung passend ist, wird sich aus den folgenden Überlegungen ergeben. Sie bedarf ebenso wie die Notation $\bigwedge^k V^*$ einer Rechtfertigung; wenn wir nämlich davon ausgehen, dass die obige Definition von $\bigwedge^k W$ für jeden Vektorraum gültig sein soll, dann ist insbesondere die Bedeutung von $\bigwedge^k V^*$ bereits festgelegt. Demnach ist $\bigwedge^k V^*$ der Vektorraum aller k -Vektoren über V^* . Dass es hier keinen Konflikt gibt, zeigt der

Satz 2.4.3. Wenn V^* der Dualraum von V ist, dann ist in natürlicher Weise $\bigwedge^k(V^*)$ der Dualraum von $\bigwedge^k(V)$.

Beweis. Wir zeigen, dass jeder faktorierte k -Vektor $w_1 \wedge \dots \wedge w_k$ eine Linearform auf $\bigwedge^k(V^*)$ liefert. Die elementaren Bausteine von $\bigwedge^k(V^*)$ sind die Elemente $\ell^{(1)} \wedge \dots \wedge \ell^{(k)}$; für ein solches liefert das Tupel (w_1, \dots, w_k) den Skalar $\det(\langle \ell^{(i)}, w_j \rangle)$. Die Zuordnung setzt sich in linearer Weise auf die Äquivalenzklassen der formalen Linearkombinationen fort. Wir haben also eine Linearform auf $\bigwedge^k V^*$ für jedes Tupel von Vektoren. Die so gewonnenen Linearformen auf $\bigwedge^k(V^*)$ hängen in multilinearer Weise vom Tupel (w_1, \dots, w_k) ab. Sie sind also durch $w_1 \wedge \dots \wedge w_k$ bestimmt. Somit haben wir Linearformen $\langle \cdot, w_1 \wedge \dots \wedge w_k \rangle$. Die Linearkombinationen liefern die Linearformen $\langle \cdot, p \rangle$.

Ebenso sieht man

1. Jedes Dachprodukt $\ell^{(1)} \wedge \dots \wedge \ell^{(k)}$ kann als eine Linearform auf $\bigwedge^k V$ verstanden werden. In der Tat kann jede alternierende k -Form $\Phi(\cdot)$ auf $V \times \dots \times V$ zu einer Linearform auf $\bigwedge^k V$ fortgesetzt werden.
2. Jede Linearform $\lambda(\cdot)$ auf $\bigwedge^k V$ ist eine Linearkombination solcher spezieller k -Formen

$$\lambda(\cdot) = \sum b_\beta \cdot \ell^{(\beta_1)} \wedge \dots \wedge \ell^{(\beta_k)}.$$

3. Zwei solche Summen sind (als alternierende k -Formen) genau dann gleich, wenn sie sich vermöge der Umrechnungsregeln (i), (ii), (iii) in einem $\bigwedge^k(V^*)$ ineinander umrechnen lassen.

Die Aussagen werden durch die folgenden Konstruktionen noch verdeutlicht. Wir werden dabei auch beweisen:

Satz 2.4.4 (Duale Basen).

Sei $\{v_1, \dots, v_n\}$ eine Basis von V und $\{\ell^{(1)}, \dots, \ell^{(n)}\}$ die dazu duale Basis von V^* .
Dann ist $\{v_K : |K| = k\}$ eine Basis von $\bigwedge^k V$ und
 $\{\ell^{(K)} : |K| = k\}$ die dazu duale Basis für $\bigwedge^k V^*$.

(Dabei bedeutet $\lambda^{(K)}$ für eine k -Teilmenge K von $\{1, \dots, n\}$ das Dachprodukt $\lambda^{(K)} = \ell^{n_1} \wedge \dots \wedge \ell^{n_k}$, wobei wir annehmen, dass die n_i die aufsteigend geordnet sind.

Konstruktion der Grassmann-Algebra

Wir haben Paare zueinander dualer Vektorräume konstruiert :

$\bigwedge^k V$ und $\bigwedge^k V^*$ für $k = 0, 1, \dots, n$ sind zueinander dual.

Man betrachtet dazu die direkten Summen

$$\bigwedge^0 V \oplus \bigwedge^1 V \oplus \dots \oplus \bigwedge^n V \quad \text{und} \quad \bigwedge^0 V^* \oplus \bigwedge^1 V^* \oplus \dots \oplus \bigwedge^n V^* .$$

Dies sind zueinander duale Vektorräume der Dimension $2^n = 1 + n + \binom{n}{2} + \dots + \binom{n}{n+1} + 1$.
Die Elemente von $\bigwedge^k V$ heißen die homogenen Elemente vom Grad k . Andere übliche Namen sind k -Vektoren oder Größen k -ter Stufe.

Satz 2.4.5 (Existenz des Dachprodukts).

Sei $p \in \bigwedge^k V$ und $q \in \bigwedge^\ell V$

$$p = \sum a_\alpha v_{\alpha_1} \wedge \dots \wedge v_{\alpha_k}$$

$$q = \sum b_\beta w_{\beta_1} \wedge \dots \wedge w_{\beta_\ell} .$$

Es existiert dann ein wohlbestimmtes Element $p \wedge q \in \bigwedge^{k+\ell} V$. Anders gesagt:

Das formal gewonnene Produkt $\sum_{\alpha, \beta} a_\alpha b_\beta \cdot w_{\alpha_1} \wedge \dots \wedge w_{\alpha_k} \wedge w_{\beta_1} \wedge \dots \wedge w_{\beta_\ell}$

ist unabhängig von der Darstellung der Faktoren p und q

Der Beweis ergibt sich aus der Feststellung, dass bei jeder der Umformungen (i), (ii), (iii), (auf den ersten oder auf den zweiten Faktor angewandt,) das bilinear ausmultiplizierte Produkt als Element in $\bigwedge^{k+\ell} V$ unverändert bleibt.

Satz 2.4.6.

Das Dachprodukt in der gradierten Algebra $\bigwedge V := \bigwedge^0 V \oplus \bigwedge^1 V \oplus \dots \oplus \bigwedge^n V$ erfüllt

- (i) $(p \wedge q) \wedge r = p \wedge (q \wedge r)$ (assoziativ)
- (ii) $p \wedge (q_1 + q_2) = p \wedge q_1 + p \wedge q_2$ (distributiv)
- (iii) $q \wedge p = (-1)^{k \cdot \ell} \cdot p \wedge q$ für $p \in \bigwedge^k V, q \in \bigwedge^\ell$ (antikommutativ)

Anhang über Unterdeterminanten

Wenn man einer $I \times J$ -Matrix A (mit $|I| = n = |J|$) die Determinante zuordnet, dann geht man davon aus, dass die Reihen angeordnet sind; es gibt eine Zeile i_1 , eine zweite Zeile i_2, \dots sowie eine erste Spalte j_1 , eine zweite Spalte j_2, \dots . Wenn man die Spalten oder der Zeilen mit einer ungeraden Permutation umordnet, dann ändert die Determinante ihr Vorzeichen. (Wenn man es mit $I \times I$ -Matrizen zu tun hat, gibt es kein Problem, weil man davon ausgeht, dass die Anordnungen der Zeilenindizes mit der Anordnung der Spaltenindizes übereinstimmt.) Wenn man es aber beispielweise mit einer Matrix eines Basiswechsels zu tun hat, dann benötigt der Begriff der Determinante eine Festlegung, welche Aufzählungen der Basisvektoren als gleichgerichtet bzw. als gegensinnig gerichtet gelten sollen. Diese Festlegung muß natürlich konsistent sein. Im reellen Fall zeigt sich die Gleichgerichtetheit zweier aufgezählter Basen darin, dass die Matrix des Basiswechsels positive Determinante hat.

Wenn man nun auch für die $k \times k$ -Untermatrizen einer Matrix mit angeordneten Reihen Determinanten haben will, dann ist es eine bequeme Konvention, die Anordnung der Reihen zu übernehmen. Im Kalkül der Grassmann-Algebra erledigt sich das ohne allen Aufwand.

Definition 2.4 (Die Minoren einer Matrix).

Wenn man aus einer $m \times n$ -Matrix A einige Spalten und einige Zeilen streicht, sodass nur noch die Zeilen zu den Indizes $\in K$ und die Spalten zu den Indizes $\in H$ übrig bleiben ($|H| = k = |K|$), so erhält man eine $k \times k$ -Teilmatrix, die wir mit A_H^K bezeichnen. (Mit dem oberen Index werden bei uns die Zeilen indiziert, der untere Index verweist auf Spalten.)

$\det A_H^K$ heißt der **Minor** zu $K \times H$

Satz 2.4.7 (Spezielle Basiswechsel in $\bigwedge^k V$).

Seien $\{v_1, \dots, v_n\}$ und $\{w_1, \dots, w_n\}$ Basen von V

$$(w_1, \dots, w_n) = (v_1, \dots, v_n) \cdot A, \quad w_j = \sum v_i \cdot a^i_j;$$

$$(v_1, \dots, v_n) = (w_1, \dots, w_n) \cdot B, \quad v_i = \sum w_j \cdot b^j_i.$$

Für k -Teilmengen K und $H \subseteq \{1, 2, \dots, n\}$ definieren wir

$$v_K = v_{m_1} \wedge \dots \wedge v_{m_k} \text{ mit } 1 \leq m_1 < \dots < m_k \leq m = n$$

$$w_H = w_{n_1} \wedge \dots \wedge w_{n_k} \text{ mit } 1 \leq n_1 < \dots < n_k \leq n.$$

Sowohl die $\{v_K : |K| = k\}$ als auch die $\{w_H : |H| = k\}$ sind eine Basis von $\bigwedge^k V$.

Die Einträge in der Matrix des Basiswechsels sind die Minoren. Es gilt

$$\begin{aligned} w_H &= \sum_K v_K \cdot \det(A_H^K) \\ v_K &= \sum_H w_H \cdot \det(B_K^H). \end{aligned}$$

Beweis.

Sei $n_1 < n_2 < \dots < n_k$ mit $\{n_1, \dots, n_k\} = H$. Dann gilt

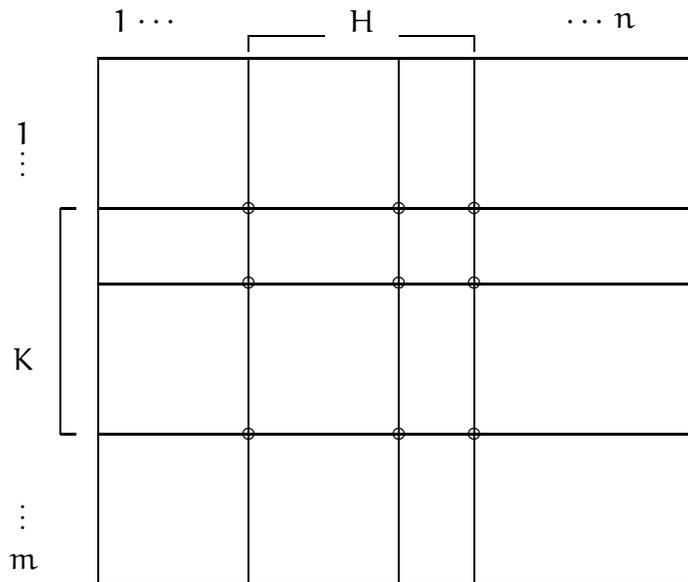
$$\begin{aligned} w_H &= \left(\sum v_{i_1} \cdot a^{i_1}_{n_1} \right) \wedge \dots \wedge \left(\sum v_{i_k} \cdot a^{i_k}_{n_k} \right) \\ &= \sum_{i_1 \dots i_k} v_{i_1} \wedge \dots \wedge v_{i_k} \cdot a^{i_1}_{n_1} \dots a^{i_k}_{n_k} . \end{aligned}$$

Wir sammeln nach den (i_1, \dots, i_k) mit $\{i_1, \dots, i_k\} = K$ und stellen die Faktoren in aufsteigende Reihenfolge. $sign\sigma$ sei das Vorzeichen dieser Permutation.

$$w_{n_1} \wedge \dots \wedge w_{n_k} = \sum_K v_K \cdot \sum sign\sigma \cdot a^{i_1}_{n_1} \dots a^{i_k}_{n_k} = \sum v_K \cdot \det(A_H^K) \quad q.e.d.$$

Zur Konkretisierung ein Diagramm.

$$\begin{aligned} 1 \leq m_1 < m_2 < \dots < m_k \leq m & \text{ mit } \{m_1, \dots, m_k\} = K \\ 1 \leq n_1 < n_2 < \dots < n_k \leq n & \text{ mit } \{n_1, \dots, n_k\} = H . \end{aligned}$$



Cofaktoren und der Laplace'sche Entwicklungssatz

Ein sehr bekannter Sonderfall ergibt sich, wenn man in einer $n \times n$ -Matrix die Unterdeterminanten der Größe $(n - 1) \times (n - 1)$ betrachtet.

Definition 2.5.

$A = (a^{i_j})$ sei eine $I \times J$ -Matrix mit $I = \{1, 2, \dots, n\} = J$.

B^h_k sei die Matrix, die man gewinnt, wenn die h -te Zeile und die k -te Spalte streicht..

Man gewinnt den Cofaktor C^h_k zu a^h_k (in der Matrix A) aus der Determinante dieser Matrix

$$C^h_k := (-1)^{h+k} \cdot \det B^h_k.$$

Der Name wird erklärlich, wenn man an diejenigen Terme denkt, die in der expliziten Formel

$$\det A = \sum (\text{sign} \sigma) a^{i_1} \cdot \dots \cdot a^{i_n}$$

mit dem Faktor a^h_k als Faktor enthalten. Der Cofactor hängt nur von der Matrix A^k_H ab, und es zeigt sich, dass er bis auf das Vorzeichen $(-1)^{h+k}$ die Determinante dieser Matrix ist.

Satz 2.4.8.

Die Matrix der Cofactoren ist bis auf den Faktor $\det A$ die Transponierte der Inversen. Wenn wir die j -te Spalte durch eine der übrigen ersetzen, ergibt sich 0. Für $h = j$ ergibt sich

$$\sum_h a^h_k \cdot C^h_j = \begin{cases} \det A & \text{falls } k = j \\ 0 & \text{falls } k \neq j \end{cases}$$

Die Formel für $k = j$ heisst die Entwicklung der Determinanten nach der k -ten Spalte. Wir betrachten den Fall $j = 1 = k$ für eine 3×3 -Matrix. e_i verstehen wir als die i -te Einheitsspalte.

$$\begin{aligned} & \left(\sum_i a^i e_i \right) \wedge \left(\sum_j b^j e_j \right) \wedge \left(\sum_k c^k e_k \right) = \\ & = a^1 e_1 \wedge \left(\sum_{j \neq 1} b^j e_j \right) \wedge \left(\sum_{k \neq 1} c^k e_k \right) \\ & + a^2 e_2 \wedge \left(\sum_{j \neq 2} b^j e_j \right) \wedge \left(\sum_{k \neq 2} c^k e_k \right) \\ & + a^3 e_3 \wedge \left(\sum_{j \neq 3} b^j e_j \right) \wedge \left(\sum_{k \neq 3} c^k e_k \right) \end{aligned}$$

in der traditionellen Notation

$$\det \begin{pmatrix} a^1 & b^1 & c^1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{pmatrix} = a^1 \cdot \det \begin{pmatrix} b^2 & c^2 \\ b^3 & c^3 \end{pmatrix} - a^2 \cdot \det \begin{pmatrix} b^1 & c^1 \\ b^3 & c^3 \end{pmatrix} + a^3 \cdot \det \begin{pmatrix} b^1 & c^1 \\ b^2 & c^2 \end{pmatrix}.$$

Eine bemerkenswerte Konsequenz der expliziten Formel für die inverse Matrix ist der

Satz 2.4.9.

Sei A eine invertierbare $n \times n$ -Matrix mit ganzzahligen Einträgen. Wenn $\det A = \pm 1$, dann hat auch die Inverse ganzzahlige Einträge.

Beweis. Die Einträge der inversen Matrix sind gerade die Cofaktoren.

3 Differentialformen

Es sei (N, \mathcal{D}) n -dimensionale \mathcal{C}^r -Mannigfaltigkeit. Wie wir gesehen haben, liefert jede glatte Funktion f in jedem Punkt P einen Cotangentialvektor $df|_P$, den ‘Anstieg der Funktion f im Punkt P ’. Nach den obigen Ausführungen ist klar, was Summen von Ausdrücken der Art $g \, df^1 \wedge \cdots \wedge df^k$ bedeuten: Jedem Punkt P wird eine alternierende Multilinearform vom Grad k auf dem Tangentialvektorraum T_P zugeordnet, und zwar mit einer Glattheit, über die noch zu reden sein wird.

Wir haben früher 1-Formen $g \, df$ über glatt parametrisierbare Kurven integriert. Im Folgenden geht es nun darum, k -Formen über ‘orientierte k -dimensionale’ Bereiche zu integrieren. Wir brauchen einige Vorbereitungen aus schon früher entwickelten Themenbereichen.

3.1 Vorbereitungen

3.1.1 Topologische Vorbemerkung: Zerlegungen der Eins

Eine Mannigfaltigkeit ist ein HRaB mit einem ausgezeichneten System von Funktionen, welches gewisse Forderungen zu erfüllen hat. Die Hausdorff-Eigenschaft und die Existenz einer abzählbaren Basis wird für das Folgende bedeutsam sein. Wir benötigen diese Eigenschaften, um sog. Zerlegungen der Eins zu konstruieren, ‘partition of unity’ im Englischen. Wir benützen die Gelegenheit, an Begriffe zu erinnern, die wir bereits im Kapitel über allgemeine Punktmengentopologie diskutiert haben.

Für eine Funktion f heisst die abgeschlossene Hülle der Menge $\{P : f(P) \neq 0\}$ der Träger; der Träger von f wird mit $\text{supp}(f)$ bezeichnet (‘support’ im Englischen). Es sei f eine stetige Funktion mit kompakten Träger und Werten zwischen 0 und 1; wir schreiben $f \prec V$, wenn V eine offene Umgebung des Trägers von f ist. Aus dem Lemma von Urysohn ist bekannt: Ist K kompakt und U offen mit $K \subset U$, so existiert eine stetige Funktion mit Werten zwischen 0 und 1, welche auf K den Wert 1 hat und auf U verschwindet.

Wir benützen im Folgenden die üblichen Bezeichnungen: \bar{A} ist die abgeschlossene Hülle A° der offene Kern. A heisst bedingt kompakt, wenn \bar{A} kompakt ist.

Lemma. *Sei K kompakt und $\bigcup_1^n U_i \supset K$ eine offene Überdeckung. Es existieren dann stetige Funktionen f_i mit $f_i \prec U_i$ und $\sum_i f_i = 1$ auf K .*

(Ein solches Tupel von Funktionen heisst eine mit der Überdeckung verträgliche Zerlegung der 1. (‘Partition subordinated to the given covering’))

Beweis. *Zu jedem $P \in K$ wählen wir eine Umgebung W_P , deren abgeschlossene Hülle in einem der U_i enthalten ist, sagen wir $\bar{W}_P \subseteq U_{i(P)}$. Eine endliche Zahl dieser \bar{W}_P überdeckt K . V_i sei die Vereinigung derjenigen, die in U_i enthalten sind. Die V_i sind eine offene Überdeckung von K mit $\bar{V}_i \subseteq U_i$. Seien nun g_i Funktionen mit Werten zwischen 0 und 1,*

die auf $\overline{V_i}$ den Wert 1 haben und ausserhalb U_i verschwinden. die folgenden Funktionen bilden dann eine verträgliche Zerlegung der Eins:

$$f_1 = g_1, \quad f_2 = g_2(1 - g_1), \quad \dots \quad f_n = g_n(1 - g_1) \cdots (1 - g_{n-1})$$

denn es gilt $f_i \prec U_i$ und $f_1 + f_2 + \cdots + f_n = 1 - (1 - g_1) \cdots (1 - g_n)$.

Jede Mannigfaltigkeit M ist lokal kompakt; denn jeder Punkt besitzt eine Koordinaten-Umgebung. Die Mannigfaltigkeit kann mit abzählbar vielen kompakten Mengen ausgeschöpft werden; und die folgende Konstruktion zeigt einiges mehr. Gegeben sei eine Überdeckung $\bigcup U_\alpha = M$ und eine abgezählte Basis B_1, B_2, \dots mit B_i bedingt kompakt. Wir konstruieren induktiv kompakte $A_1 \subset A_2 \subset A_3 \subset \dots$, sodass $A_i \subset A_{i+1}^\circ$, beginnend mit $A_1 = \overline{B_1}$. wenn A_i bereits konstruiert, dann sei j die kleinste Zahl, sodass $A_i \subset B_1 \cup B_2 \cup \cdots \cup B_j$ und $A_{i+1} = \overline{B_1} \cup \overline{B_2} \cup \cdots \cup \overline{B_j} \cup \overline{B_{i+1}}$. Die kompakten Mengen $K_i = A_{i+1} - A_i^\circ$ sind eine Überdeckung der Mannigfaltigkeit und $A_{i+2}^\circ - A_{i-1}$ ist eine offene Umgebung von K_i . Wir bemerken: $(A_{j+2}^\circ - A_{j-1}) \cap (A_{i+2}^\circ - A_{i-1}) = \emptyset$, wenn $j \geq i + 3$.

Zu jedem $P \in A_{i+2}^\circ - A_{i-1}$ wählen wir eine offene Umgebung $V_P \subseteq A_{i+2}^\circ - A_{i-1}$, (enthalten in einem der U_α), wie folgt: wir wählen eine Koordinatisierung bei P und einen Radius r , sodass die Kugel $\{x : \|x - x(P)\| \leq r(1 + \delta)\}$ im Bild der Koordinatenumgebung liegt und wir setzen $V_P = \{Q : \|x(Q) - x(P)\| \leq r(1 + \delta)\}$, wobei wir evtl. r noch etwas verkleinern, um $V_P \subseteq A_{i+2}^\circ - A_{i-1}$ zu garantieren. Die offene 'Kugel' $\tilde{V}_P = \{Q : \|x(Q) - x(P)\| < r\}$ ist eine Umgebung von P ; und es reichen endlich viele dieser Kugeln \tilde{V}_P aus, um die kompakte Menge K_i zu überdecken; es seien die Familie der Kugeln $\{\tilde{V}_{P_j} : j \in J_i\}$

Es sei h eine antitone Funktion auf \mathbb{R}_+ mit $h(t) = 1$ für $t \leq 1$ und $h(t) = 0$ für $t \geq 1 + \delta$. Für die P_j betrachten wir die Funktionen $h_j(Q) = h(\frac{1}{r_j} \|x(Q) - x(P_j)\|)$. Die Summe $H(Q) = \sum_{j \in \bigcup J_i} h_j(Q)$ ist in allen Punkten Q endlich und die Funktionen $f_j = \frac{1}{H} \cdot h_j$ bilden eine Zerlegung der 1 auf der Mannigfaltigkeit im Sinne der folgenden

Sprechweise. Eine mit der offenen Überdeckung der Mannigfaltigkeit $M = \bigcup U_\alpha$ verträgliche Zerlegung der Eins ist eine abzählbare Familie von nichtnegativen Funktionen $\{f_j\}$ mit

1. Zu jedem j existiert ein α_j , sodass $f_j \prec U_{\alpha_j}$
2. In jedem Punkt Q sind nur endlich viele der f_j ungleich 0
3. $\sum_j f_j(\cdot) = 1$ auf M .

Wenn M eine C^r -Mannigfaltigkeit ist und alle f_j r -mal stetig differenzierbar sind, spricht man von einer r -glatten Zerlegung der 1.

Wir zeigen noch, dass man durch eine geeignete Wahl der Funktion $h(\cdot)$ eine r -glatte Partition erhält: Es seien $a < b \in \mathbb{R}$ und für $t \in \mathbb{R}$

$$h_{a+}(t) = \begin{cases} \exp\left(-\frac{1}{t-a}\right) & \text{für } t > a \\ 0 & \text{für } t \leq a \end{cases} \quad h_{b-}(t) = \begin{cases} \exp\left(-\frac{1}{b-t}\right) & \text{für } t < b \\ 0 & \text{für } t \geq b \end{cases}$$

Die Funktionen sind unendlich oft stetig differenzierbar und die Funktion

$$h_{(a,b)}(t) = h_{a+}(t) \cdot (h_{a+}(t) + h_{b-}(t))^{-1}$$

ist monoton wachsend mit $h_{(a,b)}(t) = 0$ für $t \leq a$ und $h_{(a,b)}(t) = 1$ für $t \geq b$.

Es liegt auf der Hand, wie man mit Funktionen dieser Art eine unendlich oft differenzierbare Funktion gewinnt, wie sie oben benötigt wird.

3.1.2 Gerichtete Volumina

Nach den obigen Ausführungen ist klar, was der Ausdruck $g \, df^1 \wedge \cdots \wedge df^k$ bedeutet: Jedem Punkt P wird hier eine faktorisierte alternierende Multilinearform vom Grad k auf dem Tangentialvektorraum T_P zugeordnet, Die Linearkombinationen dieser Objekte heissen die k -Formen auf der Mannigfaltigkeit. Der Vektorraum der glatten k -Formen wird mit $\Omega^k N$ bezeichnet, oder kurz mit Ω^k , wenn klar ist, auf welche Mannigfaltigkeit wir uns beziehen.

Eine ganz besondere Rolle spielen nun die n -Formen; man nennt sie auch top-Formen, weil sie die Formen vom maximalen Grad sind. Eine top-Form ω^n , eingeschränkt auf die Koordinatenumgebung zu $\{x^j : j \in J\}$ hat die Form $\omega^n = h \, dx^1 \wedge \cdots \wedge dx^n$. Wir werden Formen integrieren, und zwar über orientierte ‘singuläre’ Simplexes, orientierte ‘singuläre’ Würfel und allgemeinere parametrisierte Bereiche \mathfrak{B} .

(Man den entsprechenden Kenntnissen in der Integrationstheorie kann man auch Ausdrücke $g \cdot \omega^n$ integrieren, wo g eine (in einem geeigneten Sinn) integrable Funktion ist. Das müssen wir hier aber nicht wieder aufgreifen.)

Manchmal wird es günstig sein, auf eine Zerlegung der Eins zurückzugreifen, $1 = \sum_l f_l$, um zuerst einzeln die Formen $(f_l \cdot g) \cdot \omega^n$ zu integrieren. Wenn sie allesamt über dem Bereich \mathfrak{B} integrabel sind und die Summe $\sum_l \int_{\mathfrak{B}} (f_l \cdot g) \cdot \omega^n$ unbedingt summabel ist, dann ergibt das $\int_{\mathfrak{B}} g \cdot \omega^n$.

Bevor wir uns aber der Integrationsthematik zuwenden, machen wir noch einige Anmerkungen über translationsinvariante Volumenformen auf dem orientierten \mathbb{R}^n .

Ausblick: In der Theorie der lokalkompakten Gruppen G mit abzählbarer Basis konstruiert man das (links)-invariante Haar’sche Maß. Dabei handelt es sich um ein σ -endliches translationsinvariantes Maß $\mu_{\mathcal{H}}$:

$$\mu_{\mathcal{H}}(g \cdot M) = \mu_{\mathcal{H}}(M) \quad \text{für alle } g \in G, \quad M \text{ borelsch } \subseteq G$$

Dabei bezeichnet $g \cdot M$ die Menge, die man erhält, wenn man alle Punkte in M von links mit dem festen Gruppenelement g multipliziert. Die Mengen $g \cdot M$ heissen die Linkstranslatierten zur Menge M . Das Haar’sche Maß ist bis auf einen Faktor eindeutig bestimmt. Existenz und Eindeutigkeit (bis auf einen Faktor) eines solchen Maßes wird zurecht als ein großartiges Resultat der Analysis gefeiert. Man kann das Haar’sche Maß als eine Volumenmessung verstehen.

Im Falle des \mathbb{R}^n ist das Haar'sche Maß (bis auf einen Faktor) das Lebesgue'sche Maß. Das Lebesgue'sche Maß ist bekanntlich nicht nur gegenüber Translationen invariant, sondern auch gegenüber affinen Automorphismen mit der Determinante = 1. Wir werden sehen, welche Konsequenzen das hat. Zur Festlegung einer Volumenmessung geht man folgendermaßen vor: Man wählt ein Haar'sches Maß und eine Orientierung des \mathbb{R}^n , und man ordnet orientierten Mengen ein 'gerichtetes Volumen' zu, zuerst den orientierten Würfeln und den orientierten Simplexes und sodann durch lineare Fortsetzung den 'Ketten', die man als formale Linearkombinationen solcher orientierten Bereiche gewinnt. Dabei erhalten die positiv orientierten offenen Mengen positives 'Volumen'; den umorientierten Mengen ändert man das Vorzeichen. Offenbar bestimmt eine aufgezählte Basis $(\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n)$ eine Volumenmessung, wenn man festlegt, dass das orientierte Parallelepipid $\tilde{\mathfrak{p}}$ mit dieser Kantenfolge das Volumen 1 haben soll; man bezeichnet diese Volumenmessung mit $\tilde{\omega}^n = \tilde{\ell}^1 \wedge \dots \wedge \tilde{\ell}^n$, wo $\{\tilde{\ell}^j : j \in J\}$ die zu $\{\tilde{\mathbf{e}}_j : j \in J\}$ duale Basis ist. Für das Parallelepipid \mathfrak{p} mit der Kantenfolge $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ ergibt sich entsprechend das Volumen

$$\langle \tilde{\omega}^n, \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n \rangle = \det A,$$

wenn $\mathbf{e}_l = \sum_{j=1}^n \tilde{\mathbf{e}}_j a^j_l$, wenn also $\mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n = \det A \cdot \tilde{\mathbf{e}}_1 \wedge \dots \wedge \tilde{\mathbf{e}}_n$.

Ist $[P_0, P_1, \dots, P_n]$ ein Simplex mit aufgezählten Ecken $P_l = P_0 + \mathbf{e}_1 + \dots + \mathbf{e}_l$, ($l = 0, 1, \dots, n$), so ergibt sich das Volumen

$$\text{Vol}[P_0, P_1, \dots, P_n] = \frac{1}{n!} \langle \tilde{\omega}^n, \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n \rangle = \frac{1}{n!} \det A.$$

Fazit: Im orientierten \mathbb{R}^n kann man die alternierenden n -Formen als die natürlichen Volumenmessungen verstehen. Gemessen werden zunächst einmal die formalen Summen von Simplexen; das sind die n -dimensionalen Ketten $B = \sum c_\alpha \cdot |S_\alpha\rangle$ mit $c_\alpha \in \mathbb{Z}$. (Zur Bezeichnung vergleiche den Abschnitt zum Randoperator ∂ .) Man notiert auch

$$\langle \tilde{\omega}^n, B \rangle = \int 1_B \cdot \tilde{\omega}^n = \int_B \tilde{\omega}^n = \sum c_\alpha \cdot \int_{S_\alpha} \tilde{\omega}^n$$

Wir wollen später allgemeinere Volumenformen über geeignete orientierte Bereiche auf einer orientierten Mannigfaltigkeit integrieren. Dazu müssen wir aber mehr über Integrationsbereiche wissen. Wir beginnen mit der orientierten Mannigfaltigkeit $N = \mathbb{R}^n$.

3.1.3 Maßtheoretische Vorbereitung: Wechsel der Integrationsvariablen

Wir wiederholen zunächst einige bekannte Konstruktionen und Bezeichnungen:

Wenn $\psi : (S_1, \mathfrak{B}_1) \rightarrow (S_2, \mathfrak{B}_2)$ eine messbare Abbildung ist und $d\mu(\cdot)$ ein endliches Maß auf \mathfrak{B}_1 , dann konstruiert man dazu das Bildmaß $d\nu(\cdot) = (d\mu)\psi_*(\cdot)$

$$\nu(B_2) = \mu(\psi^{-1}(B_2)), \quad \int f_2(t) d\nu(t) = \int (\psi^* f_2)(s) d\mu(s).$$

Hier bezeichnet $B_1 = \psi^{-1}(B_2)$ das volle Urbild; seine Indikatorfunktion ergibt sich als der Pullback der Indikatorfunktion von B_2 . $\mathbf{1}_{B_1}(s) = (\psi^* \mathbf{1}_{B_2})(s) = \mathbf{1}_{B_2}(\psi(s))$.

Wir lassen die Pullback-Abbildung nicht nur auf die Indikatorfunktionen wirken, sondern (zunächst einmal auch) auf die reellwertigen messbaren Funktionen f_2 , — und später auch auf die Differentialformen. Der Pullback ψ^* bildet also (zunächst einmal) den Raum \mathfrak{F}_2 der messbaren Funktionen in den Raum \mathfrak{F}_1 ab.

Eine wichtige Rolle werden im Folgenden diejenigen ψ spielen, die eine messbare Umkehrung χ besitzen. In diesem Fall der messbaren Invertierbarkeit sind die Bilder messbarer Mengen messbar und das volle ψ -Urbild B_1 von B_2 ist das χ -Bild von B_2 . $\psi^{-1}(B_2) = \chi(B_2)$. Die Pullbackabbildungen sind zueinander invers

$$\begin{aligned} \chi : (S_2, \mathfrak{B}_2) &\longrightarrow (S_1, \mathfrak{B}_1), & \chi \circ \psi(\cdot) &= \text{id}_{S_1}(\cdot), & \psi \circ \chi(\cdot) &= \text{id}_{S_2}(\cdot); \\ \chi^* : \mathfrak{F}_1 &\rightarrow \mathfrak{F}_2, & \psi^* : \mathfrak{F}_2 &\rightarrow \mathfrak{F}_1 & \chi^* \circ \psi^*(\cdot) &= \text{id}_{\mathfrak{F}_2}(\cdot), & \psi^* \circ \chi^*(\cdot) &= \text{id}_{\mathfrak{F}_1}(\cdot). \end{aligned}$$

Wenn $\mathbf{h}_1 \cdot d\mu(\cdot)$ eine bzgl. μ totalstetige Ladungsverteilung ist ($\int |\mathbf{h}_1|(s) d\mu(s) < \infty$), dann ist das Bildmaß bezgl. ν totalstetig mit

$$(\mathbf{h}_1 \cdot d\mu)\varphi_* = \chi^*(\mathbf{h}_1) \cdot (d\mu)\varphi_*.$$

Im Falle der messbar umkehrbaren Abbildungen ist die folgende Darstellungsweise adäquat: Wenn μ der (kleinen) Menge B_1 ein gewisses Gewicht gibt, dann legt das Bildmaß dieses Gewicht in die Bildmenge $B_2 = \psi(B_1)$; wenn man das Maß im Urbildraum mit einer Dichte \mathbf{h}_1 modifiziert, dann modifiziert sich das Bildmaß um die ‘verschobene’ Funktion $\chi^*(\mathbf{h}_1)$.

Satz 3.1.1. *Es sei $\psi : U' \ni s \mapsto t(s) \in U''$ eine stetig differenzierbare Bijektion einer offenen Menge $U' \subseteq \mathbb{R}_{\mathbb{S}^p}^I$ auf eine offene Menge $U'' \subseteq \mathbb{R}_{\mathbb{S}^p}^I$. ($t^j = T^j(s^1, \dots, s^n)$). $\Delta(s) = (\det \frac{\partial t}{\partial s})(s)$ sei die Jacobi-Determinante im Punkt s . Es gilt dann für alle nicht-negativen borelmessbaren f auf U''*

$$\int f(t) dt = \int f(T(s)) \cdot |\Delta(s)| ds.$$

Beweis. *Betrachten wir die speziellen Funktionen $f(t) = \mathbf{1}_{\psi(W)}(t)$, wo W ein Würfel $\subset U'$ ist. Der Satz behauptet dann, dass das Lebesgue-Maß des Bildes von W das Integral der Jacobi-Determinante über W ist. Zum Beweis zerlegen wir W in so kleine Würfel, dass die Jacobi-Determinante in jedem Würfel höchstens um den Faktor $(1 \pm \varepsilon)$ schwankt. Das Volumen des Bildes eines Würfels bzgl. einer affinen Abbildung mit der Determinante Δ ist das Volumen des Würfels multipliziert mit $|\Delta|$. Die Summe ergibt einen Näherungswert des Integrals bis auf einen Faktor $(1 \pm \varepsilon)$.*

Das Lebesgue-Maß $|\Delta(s)| ds$ eingeschränkt auf U' wird durch ψ auf ein Maß auf U'' abgebildet und die Rechnung zeigt, dass das Bildmaß das auf U'' eingeschränkte Lebesgue-Maß ist. In der Tat, wenn $\chi(\cdot)$ die Umkehrabbildung bezeichnet, dann gilt $\mathbf{1}_{\psi(W)}(t) =$

$\chi^*(1_W)(t)$; und wir haben gezeigt

$$\int \chi^*(1_W)(t) dt = \int 1_W(s) \cdot |\Delta(s)| ds \quad \text{für alle Würfel}$$

und das Mengensystem aller Würfel $\subset U'$ ist groß genug, um ein Maß auf der Borelgebra über U' zu identifizieren. Für $f(T(s)) = (\psi^*f)(s)$ haben wir $\chi^* \circ \psi^*(f) = f$, also

$$\int f(T(s)) \cdot |\Delta(s)| ds = \int f(t) dt. \quad \text{für alle nichtnegativen borelmessbaren } f \text{ auf } U'.$$

Normierung und Orientierung

Im Beweis sind wir dem Brauch gefolgt, der das Lebesgue-Maß auf dem Raum der I-Spalten s mit ds bezeichnet und das Lebesgue-Maß auf dem Raum der J-Spalten t mit dt oder mit $dt^1 \cdots dt^n$. In affinen Räumen ist die Darstellung der Punkte und Vektoren durch Tupel nicht zwingend; die Integration der Funktionen auf einem reellaffinen Raum sollte nicht an eine spezielle Koordinatisierung gebunden werden. Allerdings ist das zugrundezuliegende Maß durch seine Eigenschaft der Translationsinvarianz nur bis auf einen Faktor bestimmt. Der gewünschte Normierungsfaktor muß spezifiziert werden.

Wir gehen von einem orientierten reellaffinen Raum aus, und legen fest, dass die positiv orientierten offenen Mengen positives Maß erhalten sollen. Wir brauchen nur eine einzige Vergleichsmenge, um auch den Faktor festzulegen. Der übliche Weg ist der, dass man ein Parallelepiped mit positiv aufgezählten Kanten e_1, e_2, \dots, e_n als eine Menge festlegt, die das Maß = 1 erhalten soll, einen 'Einheitswürfel'. Wenn $\ell^1, \ell^2, \dots, \ell^n$ die duale Basis ist, dann bezeichnet man das jetzt eindeutig festgelegte translationsinvariante Maß mit $\ell^1 \wedge \ell^2 \wedge \dots \wedge \ell^n$. Eine empfohlene Notation für das Integral der Funktion $f(\cdot)$ lautet $\int f dx^1 \wedge \dots \wedge dx^n$, wenn man dasjenige Maß meint, welches den entsprechend orientierten Würfeln $W = \{Q : 0 \leq |x^j(Q) - x^j(\tilde{P})| \leq 1 \text{ für alle } j\}$ das Volumen = 1 zuordnet. Diese Notation passt für die Integration auf orientierten Mannigfaltigkeiten; dort denkt man nicht daran, Funktionen zu integrieren; man integriert Top-Formen ω^n . Und die haben bekanntlich in lokalen Koordinaten die Gestalt

$$f dx^1 \wedge \dots \wedge dx^n \quad \text{oder eben} \quad g dy^1 \wedge \dots \wedge dy^n \quad \text{mit} \quad f = g \cdot \frac{\partial y}{\partial x}.$$

Die Formel vom Wechsel der Integrationsvariablen bekommt so eine übersichtliche Gestalt. Die Frage der Integration über orientierte Bereich soll noch genauer besprochen werden.