3 Normalformen

Dieses Kapitel ist teilweise angelehnt an [A. Vanderbauwhede: *Centre Manifolds, Normal Forms and Elementary Bifurcations*, Dynamics Reported **2**, 1989].
Um ein dynamisches System nach der Reduktion auf eine Zentrumsmannigfaltigkeit weiter zu untersuchen, bedient man sich oft sogenannter Normalformen. Dabei geht es darum, durch eine geschickte Koordinatentransformation das dynamische System in der Nähe einer Ruhelage oder eines Fixpunktes möglichst „einfach“ zu machen. Ziel dabei ist, dass ein dynamisches System entsteht, dessen Taylor-Entwicklung um die Ruhelage (bzw. den Fixpunkt) möglichst wenige Terme enthält.
Da wir Normalformen hauptsächlich über ihre Taylor-Entwicklungen charakterisieren wollen, führen wir die folgende Notation ein.

Definition: Für $g \in C^k(\mathbb{R}^n)$ und $0 \leq m \leq k$ sei

$$T_m g(x) := \sum_{|\alpha| \leq m} \frac{1}{\alpha!} D^\alpha g(0) \cdot x^\alpha$$

das *Taylor–Polynom* von g in 0 bis zur Ordnung m und

$$\tilde{T}_m g(x) := \sum_{|\alpha| = m} \frac{1}{m!} D^\alpha g(0) \cdot x^\alpha$$

der homogene Anteil vom Grad m.

Hierbei ist $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$ ein *Multiindex*, d.h. es bedeutet wie üblich

$$\begin{aligned} |\alpha| &= \alpha_1 + \alpha_2 + \ldots + \alpha_n, \\ \alpha! &= (\alpha_1!)(\alpha_2!)\ldots(\alpha_n!) \\ D^\alpha &= \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}} \quad \text{und} \\ x^\alpha &= x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot \ldots \cdot x_n^{\alpha_n}. \end{aligned}$$

Sei weiter $H_m(\mathbb{R}^n)$ der Vektorraum aller Abbildungen $h : \mathbb{R}^n \to \mathbb{R}^n$, so dass h_i ein homogenes Polynom vom Grad m in $x_1, x_2, \ldots, x_n$ ist für $i = 1, 2, \ldots, n$.
Dabei ist natürlich

$$(h + \tilde{h})(x_1, x_2, \ldots, x_n) = h(x_1, x_2, \ldots, x_n) + \tilde{h}(x_1, x_2, \ldots, x_n)$$

und

$$(c \cdot h)(x_1, x_2, \ldots, x_n) = c \cdot h(x_1, x_2, \ldots, x_n)$$

für $c \in \mathbb{R}$.

Bemerkung: Für jedes $g \in C^m(\mathbb{R}^n)$ ist $\tilde{T}_m g \in H_m(\mathbb{R}^n)$ ein homogenes Polynom vom Grad m.

3.1 Normalformen für Diffeomorphismen

Dieser Teil wird im Tutorium behandelt.
Zunächst wollen wir Normalformen von Diffeomorphismen

$$x_{n+1} = F(x_n)$$

mit $x \in \mathbb{R}^N$ mit $F(0) = 0$ und $DF(0) = A$, anschließend werden wir sehen, wie man bei Differentialgleichungen weitgehend analog vorgehen kann.
Durch eine Koordinatentransformation $x = \Phi(y)$ ergibt sich die neue Abbildung

$$y_{n+1} = \Phi^{-1}(F(\Phi(y_n))) = G(y_n)$$

bzw. es gilt

$$\Phi \circ G = F \circ \Phi$$

Sei weiter $\Phi(0) = 0$ und $D\Phi(0) = Id$, dann ist weiterhin $y = 0$ ein Fixpunkt, denn $G(0) = \Phi^{-1}(F(\Phi(y))$ und die Linearisierung $DG(0) = A$ bleibt unverändert. (Nachrechnen !)
Nun kann man versuchen die Taylor-Reihe von G soweit wie möglich zu vereinfachen. Analog wie in Lemma 3.1. (?) der Vorlesung zeigt man

$$\begin{aligned} G &= F \circ \Phi - \Phi \circ G + G \\ &= (F - A) \circ \Phi + A \circ \Phi - (\Phi - Id) \circ (G - A) - \Phi \circ A + A. \end{aligned}$$

woraus sich ergibt, dass durch Wahl von Φ genau diejenigen Terme aus der Taylor-Entwicklung wegtransformiert werden können, die im Bild der Abbildung $\Phi \mapsto A \circ \Phi - \Phi \circ A$ liegen, denn für $m \geq 2$ ist

$$\tilde{T}_m G = \tilde{T}_m(A \circ \Phi - \Phi \circ A) + \mathcal{R}_m$$

wobei

$$\mathcal{R}_m = \tilde{T}_m((F - A) \circ \Phi - (\Phi - Id) \circ (G - A))$$

sich aus $T_m F$, $T_{m-1}\Phi$ und $T_{m-1}G$ berechnen lässt.
(Vielleicht sollte man vorher dieser Abbildung $\Phi \mapsto A \circ \Phi - \Phi \circ A$ noch einen Namen geben und zeigen, dass sie $H_m(\mathbb{R}^N$ in sich abbildet...)

3.2 Normalformen für Differentialgleichungen

Für eine gewöhnliche Differentialgleichung der Form

$$\dot{x} = f(x), \quad f \in C^k(\mathbb{R}^n) \tag{1}$$

mit $f(0) = 0$ entwickelt man f in eine Taylorreihe und versucht, durch Koordinatentransformationen möglichst viele Terme wegzutransformieren. Dabei kann man entweder mit einer polynomialen Transformation eine Normalform „endlicher Ordnung“ + ein Restglied polynomialer Ordnung erhalten, oder durch einen Grenzübergang eine Normalform „unendlicher Ordnung“ mit einem Restglied, das kleiner ist als jede polynomiale Ordnung, beispielsweise von der Form $\mathcal{O}(e^{-1/|x|})$.

Sei

$$A := Df(0).$$

Wir suchen eine Klasse $NF_k(A)$ von Vektorfeldern, so dass sich die Gleichung (1) durch eine Koordinatentransformation auf die Gestalt

$$\dot{y} = g(y)$$

mit $g \in NF_k(A)$ bringen läßt.
Satz 3.2 über Normalformen wird uns eine Charakterisierung der Elemente von $NF_k(A)$ durch ihre Taylor-Entwicklung bis zur Ordnung k im Ursprung liefern.
Wir beginnen mit einigen Definitionen und Vorüberlegungen. Sei dazu $\Phi \in C^\infty(\mathbb{R}^n)$ ein Diffeomorphismus. Mit

$$x = \Phi(y)$$

ergibt sich dann aus der Kettenregel

$$\dot{x} = f(\Phi(y)) = D\Phi(y) \cdot \dot{y}$$

so dass unsere Gleichung (1) in den neuen Koordinaten die Form

$$\dot{y} = (D\Phi(y))^{-1} \cdot f(\Phi(y)), \quad y \in \mathbb{R}^n$$

lautet.

Definition: Für zwei Abbildungen $\phi, \psi \in C^\infty(\mathbb{R}^n)$ definieren wir die **Lie-Ableitung** $L_\psi \phi \in C^\infty(\mathbb{R}^n)$ von ϕ in Richtung ψ als

$$\begin{aligned} L_\psi \phi(x) &:= D\phi(x) \cdot \psi(x) \qquad \forall x \in \mathbb{R}^n \\ &= \left. \frac{d}{dt} \right|_{t=0} \phi(x + t \cdot \psi(x)). \end{aligned}$$

Die **Lie-Klammer** von ϕ und ψ ist die bilineare Abbildung

$$[\phi, \psi] := L_\psi \phi - L_\phi \psi.$$

Für eine $n \times n$-Matrix A sei die **adjungierte Darstellung**

$$\text{ad } A : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$$

gegeben durch

$$\text{ad } A \cdot \phi := [A, \phi] \qquad \forall \phi \in C^\infty(\mathbb{R}^n)$$

beziehungsweise ausführlicher

$$(\text{ad } A \cdot \phi)(x) = [A, \phi] = L_\phi A - L_A \phi = A\phi(x) - D\phi(x) \cdot Ax \qquad \forall x \in \mathbb{R}^n.$$

Bemerkung: $\text{ad}_m A := \text{ad } A|_{H_m} : H_m \to H_m$ bildet den Raum H_m in sich ab, denn für $\phi \in H_m$ ist

$$(\text{ad } A \cdot \phi)(x) = \underbrace{A\phi(x)}_{\text{homogen Grad } m} - \underbrace{D\phi(x)}_{\text{hom. Grad } m-1} \cdot \underbrace{Ax}_{\text{hom. Grad } 1}.$$

Bevor wir den Satz über Normalformen formulieren können, zeigen wir noch ein vorbereitendes Lemma.

Lemma 3.1 *Sei $f \in C^k$ für ein $k \geq 2$, $f(0) = 0$, $Df(0) = A$ und $\Phi \in C^\infty$ mit $\Phi(0) = 0$, $D\Phi(0) = \mathrm{Id}$*
Sei $g := D\Phi^{-1} \cdot (f \circ \Phi)$.
Dann ist $g \in C^k$ mit $g(0) = 0$, $Dg(0) = A$ und

$$\begin{aligned} \tilde{T}_m g &= \mathrm{ad}_m A \cdot \tilde{T}_m \Phi + \tilde{T}_m[(T_m f - A) \circ T_{m-1}\Phi - L_{(T_{m-1}g - A)}(T_{m-1}\Phi - \mathrm{Id})] \\ &=: \mathrm{ad}_m A \cdot \tilde{T}_m \Phi + R_m \end{aligned}$$

für alle $2 \leq m \leq k$. Insbesondere lässt sich $\tilde{T}_m g$ daher aus $T_m f$, $T_m \Phi$ und $T_{m-1} g$ berechnen.

Beweis: Es ist

$$D\Phi(x) \cdot g(x) = f(\Phi(x)),$$

das heißt, speziell für $x = 0$ ergibt sich sofort $g(0) = 0$. Differenziert man diese Gleichung einmal, erhält man

$$D^2\Phi(x)(g(x), g(x)) + D\Phi(x)Dg(x) = Df(\Phi(x))D\Phi(x) \tag{2}$$

und wieder mit $x = 0$ erhält man $Dg(0) = 0$, also $g \in C^\infty_1(A)$. Wegen

$$L_g \Phi = D\Phi \circ g = f \circ \Phi$$

ist

$$\begin{aligned} g &= f \circ \Phi - L_g(\Phi - \mathrm{Id}) \\ &= A\Phi + (f - A) \circ \Phi - L_{g-A}(\Phi - \mathrm{Id}) - L_A(\Phi - \mathrm{Id}) \\ &= (\mathrm{ad}\, A)\Phi + (f - A) \circ \Phi - L_{g-A}(\Phi - \mathrm{Id}) + L_A \mathrm{Id} \\ \Rightarrow \tilde{T}_m g &= \underbrace{\tilde{T}_m(\mathrm{ad}_m A)\Phi}_{=(\mathrm{ad}_m A)\tilde{T}_m \Phi} + \underbrace{\tilde{T}_m L_A \,\mathrm{Id}}_{=0} + \tilde{T}_m[(f - A) \circ \Phi - L_{g-A}(\Phi - \mathrm{Id})]. \end{aligned}$$

Dabei benutzt man im letzten Schritt die Identität $T_1(f - A) = T_1(g - A) = 0$.
Weiter ist

$$\tilde{T}_m((f - A) \circ \Phi) = \tilde{T}_m((T_m f - A) \circ T_{m-1}\Phi),$$

da $T_m f - A$ nach Voraussetzung mit quadratischen Termen beginnt. □

Der Anteil von $\tilde{T}_m g$, der im Bild von H_m unter $\mathrm{ad}_m A$ liegt, kann also durch eine geeignete Wahl von $\tilde{T}_m \Phi$ wegtransformiert werden. Man kann sich daher Schritt für Schritt die Glieder der Taylor-Reihe von Φ und von g besorgen.
Nun sind wir in der Lage den folgenden grundlegenden Satz zu beweisen:

Satz 3.2 (Satz über Normalformen) *Sei $A \in \mathcal{L}(\mathbb{R}^n)$ und für jedes $m \geq 2$ sei W_m ein Unterraum von $H_m(\mathbb{R}^n)$ mit*

$$H_m(\mathbb{R}^n) = \mathcal{R}(\mathrm{ad}_m\, A) \oplus W_m.$$

Sei $k \geq 2$. Dann existiert für jedes $f \in C^k$ mit $Df(0) = A$ ein C^k–Diffeomorphismus Φ mit $D\Phi(0) = \mathrm{Id}$ so dass für die transformierte Differentialgleichung

$$\dot{y} = g(y)$$

mit $x = \Phi(y)$ gilt:

$$\tilde{T}_m g \in W_m, \qquad 2 \leq m \leq k.$$

Beweis: Wir können sukzessive $\phi_m \in H_m(\mathbb{R}^n)$ wählen, so dass mit $\Phi_m = Id + \sum_{l=2}^m \phi_m$ gilt:

$$\tilde{T}_m g = (\mathrm{ad}_m\, A)\phi_m + R_m \in W_m.$$

Das folgende Lemma von Borel zeigt, dass zu einer solchen Folge $\tilde{T}_m\Phi$ von Taylorgliedern auch wirklich ein Diffeomorphismus gehört, der diese Taylor-Entwicklung hat.
Es zeigt, dass durch die Folge der ϕ_m tatsächlich ein Diffeomorphismus $\Phi \in C^\infty$ mit $\Phi(0) = 0$, $D\Phi(0) = \mathrm{Id}$) und $\tilde{T}_m\Phi = \phi_m$ definiert wird. □

Lemma 3.3 ***[Borel]***
Sei $(\phi_m)_{m\geq 2}$ eine Folge von Abbildungen in $C^\infty(\mathbb{R}^n)$ mit $\phi_m \in H_m(\mathbb{R}^n)$. Dann existiert ein Diffeomorphismus $\Phi \in C^\infty$ mit $D\Phi(0) = \mathrm{Id}$ so dass

$$\tilde{T}_m\Phi = \phi_m \qquad \forall m \geq 2.$$

Beweis: Sei χ eine C^∞-Abschneidefunktion mit $\chi(x) = 1$ für $|x| \leq 1$ und $\chi(x) = 0$ für $|x| \geq 2$. Dann hat $\hat{\phi}_k(x) := \phi_k(x)\chi(x)$ kompakten Träger und es ist

$$M_k := \sup_{x\in\mathbb{R}^n, 0\leq m\leq k} \|D^m\hat{\phi}_k(x)\| < \infty.$$

Mit $\rho_k := \min(1, 2^{-k}M_k^{-1})$ definieren wir nun Funktionen

$$\psi_k(x) := \phi_k(x)\chi\left(\frac{x}{\rho_k}\right).$$

Für festes $m \geq 0$ und alle $k > m$, $k \geq 2$ gilt dann

$$\|D^m\psi_k(x)\| \leq \rho_k^{k-m}M_k \leq \rho_k M_k \leq 2^{-k}, \qquad \forall x \in \mathbb{R}^n.$$

Daher konvergiert die Reihe

$$\psi^{(m)}(x) := \sum_{k=2}^{\infty} D^m\psi_k(x)$$

für jedes $m \in \mathbb{N}$ gleichmäßig in $x \in \mathbb{R}^n$. Daraus folgt, dass $\psi := \psi^{(0)}$ eine C^∞-Funktion ist mit $D^m\psi(x) = \psi^{(m)}(x)$. Insbesondere ist dann $\psi(0) = 0$, $D\psi(0) = 0$ und

$$D^m\psi(x)(0) = \psi^{(m)}(0) = \sum_{k=2}^{\infty} D^m\psi_k(0) = \sum_{k=2}^{\infty} D^m\phi_k(0) = D^m\phi_m(0)$$

für alle $m \geq 2$.
Außerdem ist für alle $x \in \mathbb{R}^n$

$$\|D\psi(x)\| \leq \sum_{k=2}^{\infty} 2^{-k} = \frac{1}{2}.$$

Also ist $\Phi := Id + \psi$ ein C^∞-Diffeomorphismus mit den gewünschten Eigenschaften. □

Bemerkungen:

1. Die ϕ_m im Beweis des Normalform-Satzes sind nicht eindeutig bestimmt, wir könnten noch beliebige Elemente aus dem Kern von $\mathrm{ad}_m A$ addieren.

2. Die Klasse von Vektorfeldern

   $$NF_k(A) := \{g \in C^k;\ Dg(0) = A \text{ und } \tilde{T}_m g \in W_m \text{ für alle } 2 \leq m \leq k\}$$

   ist also eine vernünftige Auswahl an Normalformen. Der Beweis hat auch gezeigt, dass es unmöglich ist, diejenigen Terme, die zu W_m gehören, durch eine polynomiale Koordinatentransformation zu beseitigen. In diesem Sinne ist unsere Klasse $NF_k(A)$ „minimal“. Allerdings hängt $NF_k(A)$ noch von der Wahl der W_m ab, über die wir bisher nichts gesagt haben. Die Hoffnung ist natürlich, dass wir durch eine günstige Wahl der Komplemente W_m erreichen können, dass $\dim W_m << \dim H_m$ ist, dass also nur wenige „resonante“ Terme in der Normalform stehenbleiben.

3. Analoge Versionen für den Fall $f \in C^\infty$ bzw. den Fall, dass f noch von einem Parameter abhängt, existieren ebenfalls und werden ähnlich bewiesen.

Bisher haben wir uns noch nicht über die konkrete Wahl der Räume W_m geäußert. In einem Artikel von [*Elphick et al.: A simple global characterization for normal forms of singular vector fields (1987)*] wurde gezeigt, dass es immer möglich ist, als Komplement zu den $\mathcal{R}(\mathrm{ad}_m A)$ die Räume

$$W_m := \mathcal{N}(\mathrm{ad}_m A^T)$$

zu wählen, wobei A^T die Transponierte von A ist, das heißt

$$(A^T x, y) = (x, Ay) \qquad \forall x, y \in \mathbb{R}^n.$$

Diese Behauptung und die sich daraus ergebenden Konsequenzen wollen wir im folgenden Abschnitt behandeln. Wir beginnen wieder mit einem nützlichen Lemma:

Lemma 3.4 *Sei A wieder eine $n \times n$-Matrix und $h \in C^\infty(\mathbb{R}^n)$. Definiert man dann für jedes $t \in \mathbb{R}$ eine Abbildung $h_t \in C^\infty(\mathbb{R}^n)$ durch*

$$h_t(x) := e^{-At} h(e^{At} x)$$

so gilt:

(i) $(\operatorname{ad} A) \cdot h_t(x) = ((\operatorname{ad} A) \cdot h)_t(x) = -\frac{d}{dt} h_t(x)$

(ii) $\mathcal{N}(\operatorname{ad} A) = \{h \in C^\infty(\mathbb{R}^n) \; ; \; h_t = h \;\; \forall t \in \mathbb{R}\}$.

Beweis:

(i) berechnet man leicht

$$\begin{aligned} (\operatorname{ad} A) \cdot h_t(x) &= Ae^{-At} h(e^{At}x) - e^{-At} Dh(e^{At}x) e^{At} Ax \\ ((\operatorname{ad} A) \cdot h)_t(x) &= e^{-At} \left(A\, h(e^{At}x) - Dh(e^{At}x) Ae^{At}x \right) \\ -\frac{d}{dt} h_t(x) &= e^{-At} A\, h(e^{At}x) - e^{-At} Dh(e^{At}x) Ae^{At}x. \end{aligned}$$

Da A mit $e^{\pm At}$ vertauscht, stimmen alle drei Ausdrücke überein.

(ii) Für $h \in \mathcal{N}(\operatorname{ad} A)$ folgt aus (i) sofort, dass $h_t = h$ für alle $t \in \mathbb{R}$. Umgekehrt zeigt die Relation $h_t = h$, dass $\frac{d}{dt} h_t(x) = 0$ in $t = 0$ und damit $h \in \mathcal{N}(\operatorname{ad} A)$.

□

Nun sind wir in der Lage, die folgende Aussage zu beweisen.

Satz 3.5

$$\mathcal{N}(\operatorname{ad}_m A^T) \oplus \mathcal{R}(\operatorname{ad}_m A) = H_m(\mathbb{R}^n)$$

Beweis: Wir werden in Kürze ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf $H_m(\mathbb{R}^n)$ konstruieren, für das die Identität

$$\langle (\operatorname{ad}_m A^T) \cdot g, h \rangle = \langle g, (\operatorname{ad}_m A) \cdot h \rangle \tag{3}$$

gilt. Dann ist für dieses Skalarprodukt

$$\begin{aligned} (\mathcal{R}(\operatorname{ad}_m A))^\perp &= \mathcal{N}((\operatorname{ad}_m A)^T) \\ &= \mathcal{N}(\operatorname{ad}_m A^T) = W_m, \end{aligned}$$

wobei die erste Identität ganz allgemein gilt [1], und die zweite aus (3) folgt. Wir definieren nun ein Skalarprodukt und zeigen anschließend, dass es genau die gewünschte Eigenschaft hat. Es genügt, wenn wir das Skalarprodukt von zwei Monomen festlegen und es dann linear auf alle Polynome fortsetzen. Sei also

$$\langle a_\sigma x^\sigma, b_{\sigma'} x^{\sigma'} \rangle := \delta_{\sigma,\sigma'} \sigma! (a_\sigma, b_{\sigma'})$$

wobei σ und σ' wieder Multiindizes sind und $\delta_{\sigma,\sigma'}$ das Kronecker-δ.

Lemma 3.6 *Um (3) zu beweisen, genügt es zu zeigen, dass für jede $n \times n$-Matrix B gilt:*

$$\langle g, h \circ B \rangle = \langle g \circ B^T, h \rangle \qquad \forall g, h \in H_m(\mathbb{R}^n) \tag{4}$$

[1] $g \in \mathcal{R}(\operatorname{ad}_m A))^\perp \Leftrightarrow \langle g, (\operatorname{ad}_m A) \cdot h \rangle = 0 \; \forall h \Leftrightarrow \langle (\operatorname{ad}_m A)^T g, h \rangle = 0 \; \forall h \Leftrightarrow g \in \mathcal{N}((\operatorname{ad}_m A)^T)$

Beweis: Mit Hilfe von (4) speziell für die Matrix $B := e^{At}$ erhält man zunächst

$$B^T = \left(e^{At}\right)^T = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k\right)^T = \sum_{k=0}^{\infty} \frac{1}{k!} (A^T)^k t^k = e^{A^T t}$$

und damit

$$\langle g, e^{-At} \circ h \circ e^{At} \rangle = \langle e^{-A^T t} \circ g \circ e^{A^T t}, h \rangle.$$

Differenziert man diesen Ausdruck nach t ergibt sich die Gleichung

$$\begin{aligned} & \langle g', e^{-At} \circ h \circ e^{At} \rangle + \langle g, e^{-At} A \circ h \circ e^{At} + (e^{-At} \circ h' \circ e^{At}) A \rangle \\ = \; & \langle e^{-A^T t} \circ g \circ e^{A^T t}, h' \rangle + \langle -e^{-A^T t} A^T \circ g \circ e^{A^T t} + e^{-A^T t} \circ (g \circ e^{A^T t}) A^T, h \rangle \end{aligned}$$

aus der (3) sofort folgt, indem man bei $t = 0$ auswertet. □

Es bleibt noch, die Gleichung (4) zu verifizieren. Dazu stellen wir zunächst fest, dass sich das Skalarprodukt $\langle \cdot, \cdot \rangle$ noch auf eine andere Art bescheiben läßt. Es ist nämlich

$$\langle g, h \rangle = \left(g(\partial_x)\; h(x)\right)\big|_{x=0}, \tag{5}$$

wobei $\partial_x = (\partial_{x_1}, \partial_{x_2}, \ldots, \partial_{x_n})$ ist.

Beispiel: Seien $g(x,y) = \binom{x^3 - y^3}{x^2 y}$ und $h(x,y) = \binom{2x^3 - xy^2}{xy^2}$ beide in $H_3(\mathbb{R}^2)$. Dann ist

$$g(\partial_x, \partial_y) h(x,y) = \frac{\partial^3}{\partial_x^3}(2x^3 - xy^2) - \frac{\partial^3}{\partial_y^3}(2x^3 - xy^2) + \frac{\partial^3}{\partial_x^2 \partial_y} xy^2 = 6.$$

In der Schreibweise (5) bedeutet (4) nichts anderes als

$$(g(\partial_x)(h \circ B))(0) = ((g \circ B^T)(\partial_x) h)(0). \tag{6}$$

Nun betrachten wir den linearen Koordinatenwechsel $x = B^T y$.

$$\frac{\partial}{\partial y_j} = \sum_{i=1}^{n} b_{ji} \frac{\partial}{\partial x_i} \Rightarrow B \partial_x = \partial_y.$$

Nun gilt

$$\begin{aligned} \langle g \circ B, h \rangle = \sum_{j=1}^{n} \langle g_j \circ B, h_j \rangle &= \sum_{j=1}^{n} g_j(B\partial_x), h_j(x)\big|_{x=0} \\ &= \sum_{j=1}^{n} g_j(\partial_y), h_j(B^T y)\big|_{y=0} \\ &= \sum_{j=1}^{n} \langle g_j, h_j \circ B^T \rangle = \langle g, h \circ B^T \rangle, \end{aligned}$$

das heißt, wir haben (6) gezeigt. □

Insgesamt haben wir damit den folgenden Satz bewiesen.

Satz 3.7 (Normalform eines gegebenen Vektorfelds)
Sei A eine $n \times n$-Matrix und $f \in C^k$ mit $Df(0) = A$. Dann existiert ein lokaler C^∞-Diffeomorphismus Φ, so dass $\Phi^* f = g$ ist, wobei $g \in C^k$ ist mit $Dg(0) = A$ und

$$T_k g \circ e^{A^T t} = e^{A^T t} \circ T_k g \qquad \forall t \in \mathbb{R}.$$

Also reduziert sich die Suche nach der Normalform von f auf die Frage, welche polynomialen Abbildungen mit der Gruppe

$$\Gamma := \{e^{A^T t} \mid t \in \mathbb{R}\} \subseteq GL(n, \mathbb{R})$$

kommutieren.

Bemerkung: Die Restterme von höherer Ordnung als k, die noch in g enthalten sind, können in manchen Fällen wegtransformiert werden, das ist aber wesentlich schwieriger. Einfacher ist es oft zu zeigen, dass diese zusätzlichen Terme die Dynamik nicht ändern.

Bei konkreten Rechnungen erweist sich manchmal eine alternative Charakterisierung der Normalformen als nützlich .

Lemma 3.8 *Sei P ein Polynom. Dann ist*

$$P(e^{A^T t}x) = e^{A^T t}P(x) \qquad \forall x, t \tag{7}$$

äquivalent zu

$$DP(x)A^T x = A^T P(x) \qquad \forall x. \tag{8}$$

Beweis: Wir nehmen zuerst an, dass (7) erfüllt ist. Differenziert man diesen Ausdruck nach t, erhält man

$$\frac{d}{dt}\left(P(e^{A^T t}x) - e^{A^T t}P(x)\right) = DP(e^{A^T t}x)A^T e^{A^T t}x - A^T e^{A^T t}P(x) = 0$$

und durch Einsetzen von $t = 0$ erhält man (8). Falls umgekehrt (2) gilt,

$$\begin{aligned}\frac{d}{dt}\left(e^{-A^T t}P(e^{A^T t}x)\right) &= -A^T e^{-A^T t}P(e^{A^T t}x) + e^{-A^T t}DP(e^{A^T t}x)A^T e^{A^T t}x \\ &= e^{-A^T t}\left(DP(e^{A^T t}x)A^T e^{A^T t}x - A^T P(e^{A^T t}x)\right) = 0\end{aligned}$$

für alle x, t. Insbesondere ist für festes x der Ausdruck $e^{-A^T t}P(e^{A^T t}x)$ nicht von t abhängig. Wegen $\left(e^{-A^T t}P(e^{A^T t}x)\right)\Big|_{t=0} = P(x)$ gilt damit $P(e^{A^T t}x) = e^{A^T t}P(x)$ für alle x und t. □

Beispiele:

1. $A =$ **Diagonalmatrix**

   Falls

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist, dann läßt sich $\mathrm{ad}_m A$ explizit berechnen und ist ebenfalls diagonal. Dazu betrachten wir eine Basis von $H_m(\mathbb{R}^n)$, nämlich

$$\{x^\sigma e_j \mid \sigma \in \mathbb{N}^n,\ |\sigma| = m,\ 1 \le j \le n\}$$

Wir berechnen für ein Element dieser Basis

$$\begin{aligned} \mathrm{ad}_m A(x^\sigma e_j) &= x^\sigma A e_j - \sum_i \underbrace{\frac{\sigma_i}{x_i} x^\sigma}_{=\frac{\partial}{\partial x_i} x^\sigma} \underbrace{(Ax)_i}_{=\lambda_i x_i} e_j \\ &= (\lambda_j - \sum_i \sigma_i \lambda_i) x^\sigma e_j, \end{aligned}$$

also sind alle unsere Basisvektoren Eigenvektoren von $\mathrm{ad}_m A$ mit Eigenwerten $\lambda_j - \sum_i \sigma_i \lambda_i$. Falls alle diese Eigenwerte von 0 verschieden sind, das heißt, wenn keine *Resonanzen* auftreten, dann ist $\mathrm{ad}_m A$ bijektiv für beliebiges m. Nach dem Satz über Normalformen existiert dann eine Koordinatentransformation $\Phi \in C^\infty$, so dass in der transformierten Gleichung

$$\dot{y} = Ay + \tilde{g}(y)$$

alle Ableitungen von $\tilde{g}$ in $y = 0$ verschwinden. Sternberg hat bewiesen, dass man sogar ein Φ finden kann, so dass lokal in einer Umgebung von $y = 0$ dieser Rest völlig wegtransformiert wird, also $\tilde{g} \equiv 0$.

2. **Zwei rein imaginäre Eigenwerte**
Als nächstes interessiert uns der Fall rein imaginärer Eigenwerte, d.h. der Linearteil A ist von der Form

$$A = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}.$$

Es wird sich zeigen, dass es viele Rechnungen erleichtert, wenn man statt mit der Differentialgleichung

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} + g(x, y)$$

mit komplexen Variablen $z = x + iy$ und $\bar{z} = x - iy$ rechnet. Dann ist

$$\begin{aligned} \dot{z} &= -\omega y + i\omega x \\ &= i\omega(x + iy) \\ &= i\omega z \end{aligned}$$

und analog

$$\dot{\bar{z}} = -i\omega \bar{z}.$$

Genau genommen geht man zur Komplexifizierung $\mathbb{C}^2$ von $\mathbb{R}^2$ über und identifiziert dabei $\mathbb{R}^2$ mit dem Unterraum $U := \{(z, \bar{z}) \mid z \in \mathbb{C}\} \subseteq \mathbb{C}^2$.

Dort wird eine lineare Abbildung induziert, die beschrieben wird durch die Matrix

$$\hat{A} = \begin{pmatrix} i\omega & 0 \\ 0 & -i\omega \end{pmatrix}.$$

Nach Satz 3.7 enthält die Normalform von f nur Terme, die mit den Matrizen

$$e^{\hat{A}^T t} = \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix}$$

kommutieren. Damit der Unterraum U invariant bleibt, muss die Normalform die Gestalt $(g(z,\bar{z}), \overline{g(z,\bar{z})})$ haben. Wir wollen nun die Normalform N-ter Ordnung für beliebiges $N \in \mathbb{N}$ bestimmen. Dazu schreiben wir das Polynom g als

$$g(z,\bar{z}) = \sum_{k+\ell \le N} a_{k\ell} z^k \bar{z}^\ell.$$

Die Bedingung aus Satz 3.7 lautet hier konkret

$$\begin{aligned} e^{-it} g(z,\bar{z}) &= g(e^{-it}z, e^{it}\bar{z}) \\ \sum_{k+\ell \le N} a_{k\ell} e^{-it} z^k \bar{z}^\ell &= \sum_{k+\ell \le N} a_{k\ell} e^{-ikt} z^k e^{i\ell t} \bar{z}^\ell \\ \sum_{k+\ell \le N} a_{k\ell} (e^{-it} - e^{it(\ell-k)} z^k \bar{z}^\ell &= 0 \end{aligned}$$

Daraus folgt sofort $k = \ell + 1$, die Gleichung in Normalform lautet also

$$\dot{z} = g(z) = \left(\sum_{\ell=0}^{[N/2]} |z|^{2\ell} \right) z.$$

Kehren wir zu reellen Koordinaten $(x_1, x_2) = ((z+\bar{z})/2, (z-\bar{z})/2i)$ zurück, ergibt sich als Normalform N-ter Ordnung

$$\begin{aligned} \dot{x}_1 &= x_1 Q_1(x_1^2 + x_2^2) - x_2 Q_2(x_1^2 + x_2^2) \\ \dot{x}_2 &= x_1 Q_2(x_1^2 + x_2^2) + x_2 Q_1(x_1^2 + x_2^2) \end{aligned}$$

mit Polynomen Q_1 und Q_2.

Sei weiter $P(x_1, x_2) = (P_1(x_1, x_2), P_2(x_1, x_2))$ das Taylor-Polynom $T_k g$ der Normalform g.

Nach Lemma 3.8 erfüllt P dann die Gleichung $DP(x)A^T x = A^T P(x)$, bzw. ausführlicher

$$\begin{pmatrix} \frac{\partial P_1}{\partial x_1} & \frac{\partial P_1}{\partial x_2} \\ \frac{\partial P_2}{\partial x_1} & \frac{\partial P_2}{\partial x_2} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_1(x_1, x_2) \\ P_2(x_1, x_2) \end{pmatrix}$$

3. **Jordan-Block**
   Betrachte als nächstes den Fall, dass

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

ein Jordan-Block zum Eigenwert 0 ist. Sei eine Differentialgleichung

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

mit $f(0) = 0$ und $Df(0) = A$ gegeben. Sei weiter $P(x_1, x_2) = (P_1(x_1, x_2), P_2(x_1, x_2))$ das Taylor-Polynom $T_k g$ der Normalform g.

Nach Lemma 3.8 erfüllt P dann die Gleichung $DP(x)A^T x = A^T P(x)$, bzw. ausführlicher

$$\begin{pmatrix} \frac{\partial P_1}{\partial x_1} & \frac{\partial P_1}{\partial x_2} \\ \frac{\partial P_2}{\partial x_1} & \frac{\partial P_2}{\partial x_2} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_1(x_1, x_2) \\ P_2(x_1, x_2) \end{pmatrix}$$

Daraus erhält man sofort die beiden Gleichungen

$$\begin{aligned} x_1 \frac{\partial P_1}{\partial x_2} &= 0 \\ x_1 \frac{\partial P_2}{\partial x_2} &= P_1 \end{aligned}$$

Insbesondere hängt damit P_1 nicht von x_2 ab, d.h. $P_1(x_1, x_2) = p_1(x_1)$, wobei p_1 wieder ein Polynom ist. Setzt man dies in die zweite Gleichung ein, so folgt

$$\frac{\partial P_2}{\partial x_2}(x_1, x_2) = \frac{p_1(x_1)}{x_1}.$$

Da auf der linken Seite ein Polynom steht, muss auch die rechte Seite polynomial sein, daher muss p_1 von der Form $p_1(x_1) = x_1 q_1(x_1)$ sein. Dann lässt sich P_2 leicht durch Integration bestimmen:

$$P_2(x_1, x_2) = q_1(x_1)x_2 + q_2(x_1).$$

Setzt man alles ein und behält nur die ersten Terme bis zur 2. Ordnung, kommt man auf diese Weise zur Normalform

$$\begin{aligned} \dot{x}_1 &= x_2 + a x_1^2 + \mathcal{O}(x_1^3) \\ \dot{x}_2 &= a x_1 x_2 + b x_1^2 + \mathcal{O}(x_1^3 + x_1^2 x_2) \end{aligned}$$

wobei a und b Parameter sind, die sich aus der vorgegebenen Nichtlinearität f ermitteln lassen.

Hier kann man übrigens durch eine andere Wahl der Komplemente auch andere Normalformen erreichen. Takens hat 1974 die Normalform

$$\begin{aligned} \dot{x}_1 &= x_2 + a x_1^2 + \mathcal{O}(x_1^3) \\ \dot{x}_2 &= b x_1^2 + \mathcal{O}(x_1^3 + x_1^2 x_2) \end{aligned}$$

hergeleitet. Sie spielt eine wichtige Rolle bei der Untersuchung der *Takens-Bogdanov-Verzweigung.*