Matrixgruppen

Wahlmodul im Bachelorstudium, SoSe 2022

Andreas Čap

Fakultät für Mathematik, Universität Wien, Oskar–Morgenstern–Platz 1, 1090 Wien

Email address: `Andreas.Cap@univie.ac.at`

Inhaltsverzeichnis

Vorwort

In der Theorie der Matrixgruppen werden auf unerwartete Weise Themen aus den beiden Grundvorlesungsyzklen über Analysis und lineare Algebra in Beziehung gebracht und ergänzen einander. Die Verwendung der Methoden der Differentialrechnung in diesem ungewohnten Kontext bietet die Gelegenheit, einen anderen Blick zu entwickeln, der auf die Sichtweise der Analysis auf Mannigfaltigkeiten vorbereitet. Zugleich ergeben sich "echte" Beispiele für viele Konzepte der Analysis, bei denen viele der vereinfachenden Annahmen, die in den Beispielen der Grundvorlesungen meist gemacht werden, wegfallen, was ebenfalls zu einem erweiterten Blick beiträgt.

Andererseits bieten die Matrizengruppen nichttriviale und substantielle Beispiele für Teilmannigfaltigkeiten euklidischer Räume. Während durch die Exponentialfunktion die Analysis auf diesen Teilmannigfaltigkeiten relativ nahe bei den Inhalten der Grundvorlesungen bleibt, zeigt sich doch auch die Notwendigkeit und Nützlichkeit der allgemeineren Version. Durch die Einbindung einer algebraischen Struktur erhält man durch Betrachtung der Ableitung in einem Punkt ungewöhnlich viel Information über Abbildungen (in diesem Fall Homomorphismen) sodass die Stärke der analytischen Methoden deutlich sichtbar wird.

Matrixgruppen spielen, vor allem als Symmetriegruppen, eine wichtige Rolle in vielen Bereicher der Mathematik und der theoretischen Physik. Sie bilden den wichtigsten Spezialfall von Lie Gruppen, deren Theorie in weiten Teilen parallel zur der Theorie der Matrixgruppen verläuft, wobei allerdings andere Methoden verwendet werden. Somit bietet die Vorlesung auch eine Vorbereitung auf die diese allgemeine Theorie.

Die erste Version dieses Skriptums ist im Sommersemester 2018 entstanden, in dem ich dieses Wahlmodul erstmals angeboten habe. Ich danke den Höhrer*innen dieser ersten Vorlesung für verschiedene Korrekturen und Verbesserungsvorschläge. Für das Sommersemester 2022 habe ich das Skriptum geringfügig überarbeitet und einige Druckfehler korrigiert.

KAPITEL 1

Einleitung – Analysis und Matrizen

Die Theorie der Matrixgruppen verbindet Elemente der beiden Grundvorlesungszyklen über Analysis und lineare Algebra. Insbesondere wollen wir Funktionen auf Räumen von Matrizen differenzieren und dabei die aus der linearen Algebra bekannten Operationen auf Matrizen effizient benutzen. Dazu wird ein etwas anderer Blick auf Teile der Analysis hilfreich sein, der auf fortgeschrittenere Teile der Analysis, insbesondere Analysis auf Mannigfaltigkeiten, vorbereitet. Wir werden also einige Teile der Analysis wiederholen und für unsere Zwecke adaptieren, was den ersten Teil dieses Kapitels bildet. Im zweiten Teil werden wir die Exponentialfunktion für Matrizen besprechen, was zum Teil ebenfalls Bekanntes wiederholt.

1.1. Differentialrechnung. Das übliche Setting der Differentialrechnung in mehreren Variablen studiert Funktionen $f : U \to \mathbb{R}^m$, wobei $U \subset \mathbb{R}^n$ eine *offene Teilmenge* ist. Die letztere Bedingung bedeutet, dass es für jeden Punkt $x \in U$ einen offenen Ball um x gibt, der noch ganz in U liegt. Für eine derartige Funktion kann man dann *Differenzierbarkeit in einem Punkt* $x_0 \in U$ definieren. Ist f im Punkt x_0 differenzierbar, dann erhält man die *Ableitung* $Df(x_0)$ von f in x_0, die eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ist. Man kann die Ableitung in x_0 folgendermaßen interpretieren: Man betrachtet die Funktion $\tilde{f}(v) := f(x_0 + v) - f(x_0)$, die auf einer offenen Umgebung von 0 in $\mathbb{R}^n$ definiert ist und 0 auf $0 \in \mathbb{R}^m$ abbildet. Die Ableitung $Df(x_0)$ ist dann die beste lineare Approximation an diese Funktion (und Differenzierbarkeit bedeutet genau, dass eine beste lineare Approximation existiert).

Formal arbeitet man in der Analysis meist mit *partiellen Ableitungen*. Schreibt man die Funktion f in Komponenten, also $f(x) = (f_1(x), \dots, f_m(x))$, dann ist jedes f_i eine Funktion $U \to \mathbb{R}$. Zerlegt man nun auch x in Komponenten, dann erhält man $f_i(x) = f_i(x_1, \dots, x_n)$, also eine reellwertige Funktion von n Variablen. Hält man nun alle dieser Variablen außer einer fest, dann ergibt sich eine reellwertige Funktion die (wegen der Offenheit von U) auf einem Intervall in $\mathbb{R}$ definiert ist. Ist diese Funktion differenzierbar und ist die laufende Variable x_j, dann nennt man ihre Ableitung die *j–te partielle Ableitung von f_i* und bezeichnet sie mit $\frac{\partial f_i}{\partial x_j}(x)$. Existieren alle partiellen Ableitungen in einem Punkt x_0, dann ist f differenzierbar in x_0 und die Ableitung $Df(x_0)$ ist durch die Matrix (a_{ij}) gegeben, wobei $a_{ij} = \frac{\partial f_i}{\partial x_j}(x_0)$ gilt.

Ist f in allen Punkten $x \in U$ differenzierbar, dann sagt man f ist *differenzierbar* auf U und hängen zusätzlich die partiellen Ableitungen stetig von x ab, dann ist f *stetig differenzierbar* auf U. Sind die partiellen Ableitungen selbst wieder differenzierbar, dann kann man zweite partielle Ableitungen bilden und so Schritt für Schritt höhere Differenzierbarkeitsstufen für f definieren. Es wird für uns kaum nötig sein, die Differenzierbarkeit von Funktionen direkt zu verifizieren, sie wird meist aus allgemeinen Resultaten folgen. Auch spielt für unsere Zwecke die genaue Differenzierbarkeitsklasse kaum eine Rolle, wir werden einfach *glatt* als Synonym für "hinreichend oft stetig differenzierbar" benutzen. Tatsächlich werden wir nur C^∞–Funktionen (also beliebig oft differenzierbare Funktionen) antreffen, für die Sätze genügt aber üblicherweise C^1 oder C^2 (also ein– oder zwei mal stetig differenzierbar) als Voraussetzung.

Für unsere Zwecke werden die partiellen Ableitungen nicht so handlich sein, weil sie sich nicht leicht auf allgemeinere Teilmengen von $\mathbb{R}^n$ übertragen lassen. Es gibt aber eine handliche Alternative, die auch anschaulich gut verständlich ist. Der Ausgangspunkt dafür ist, dass die Ableitungen von Kurven leicht zu behandeln und anschaulich verständlich sind. Sei also $I \subset \mathbb{R}$ ein offenes Intervall und $c : I \to \mathbb{R}^n$ eine stetige Funktion, die wir im weiteren als Kurve bezeichnen werden. Für einen Punkt $t \in I$ kann man dann Differenzierbarkeit von c in t durch Existenz des Limes $\lim_{h\to 0} \frac{1}{h}(c(t+h) - c(t)) =: c'(t) \in \mathbb{R}^n$ definieren. (Man multipliziert den Skalar $\frac{1}{h} \in \mathbb{R}$ mit dem Vektor $c(t+h)-c(t) \in \mathbb{R}^n$ und erhält einen Vektor in $\mathbb{R}^n$.) Ist c differenzierbar, dann ist c' wieder eine Funktion von I nach $\mathbb{R}^n$, also ist auch der Umgang mit höherer Differenzierbarkeit für Kurven ganz einfach. Anschaulich ist $c'(t)$ einfach der Geschwindigkeitsvektor von c im Punkt t, also der "Richtungsvektor" im entsprechenden Punkt.

Der Schlüssel zur Verwendung von Kurven ist nun ein Spezialfall der aus der Analysis bekannten Kettenregel. Ist $f : U \to \mathbb{R}^m$ eine glatte Funktion und $c : I \to \mathbb{R}^n$ eine glatte Kurve mit $c(I) \subset U$, dann ist auch die Komposition $f \circ c : I \to \mathbb{R}^m$ eine glatte Kurve und für ihre Ableitung gilt $(f \circ c)'(t) = Df(c(t))(c'(t))$. Insbesondere kann man für einen Punkt $x \in U$ und einen Vektor $v \in \mathbb{R}^n$ immer die Kurve $c(t) = x + tv$ betrachten, die auf einem offenen Intervall um 0 Werte in U hat. Damit kann $Df(x)(v)$ als die *Richtungsableitung* $\frac{d}{dt}|_{t=0} f(x+tv)$ berechnen. Auch die Richtungsableitungen sind aber für den Übergang auf allgemeinere Teilmengen nicht so gut geeignet.

Beispiel 1.1. Wir wollen die Verwendung von Kurven an einem Beispiel illustrieren. Betrachten wir die Funktion $f : \mathbb{R}^n \to \mathbb{R}$, die durch $f(x) = |x|^2$ gegeben ist. Explizit ist $f(x_1, \dots, x_n) = \sum_{i=1}^n (x_i)^2$, also ein Polynom und damit eine glatte Funktion. Die partiellen Ableitungen sind offensichtlich durch $\frac{\partial f}{\partial x_i} = 2x_i$ gegeben, also ist $Df(x)$ die $1 \times n$–Matrix $(2x_1, \dots, 2x_n)$. Damit ist aber für $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ die Richtungsableitung $Df(x)(v) = \sum_{i=1}^n 2x_i v_i = 2\langle x, v\rangle$ gegeben.

Um das weniger koordinatenlastig zu berechnen kann man beobachten, dass $f(x) = \langle x, x\rangle$ gilt. Damit ist $f(x + tv) = \langle x + tv, x + tv\rangle$ was man wegen Symmetrie und Bilinearität des inneren Produkts als $\langle x, x\rangle + 2t\langle x, v\rangle + t^2\langle v, v\rangle$ schreiben kann. Natürlich ist die Ableitung dieser Funktion bei $t = 0$ gerade $2\langle x, v\rangle$.

Betrachten wir nun die Einheitssphäre $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ und eine glatte Kurve $c : (a, b) \to \mathbb{R}^n$, die Werte in S^{n-1} hat. Dann ist $f \circ c : (a, b) \to \mathbb{R}$ die konstante Funktion 1, also gilt für alle $t \in (a, b)$ die Gleichung

$$0 = (f \circ c)'(t) = Df(c(t))(c'(t)) = 2\langle c(t), c'(t)\rangle.$$

Somit steht $c'(t)$ normal auf $c(t)$ für alle t. Ist umgekehrt $x \in S^{n-1}$ ein Punkt und v ein Vektor, der normal auf x steht, dann gibt es eine Kurve $c : \mathbb{R} \to S^{n-1}$, die $c(0) = x$ und $c'(0) = v$ erfüllt. Für $t \in \mathbb{R}$ ist nämlich $f(x+tv) = |x+tv|^2 = 1 + t^2|v|^2 \geq 1$ und damit ist $\alpha(t) := |x + tv| = \sqrt{1 + t^2|v|^2}$ eine glatte Funktion mit Werten ≥ 1. Man sieht auch sofort, dass die Ableitung dieser Funktion bei $t = 0$ verschwindet (sie hat ja dort ein Minimum). Damit kann man aber eine glatte Kurve c durch $c(t) = |x + tv|^{-1}(x + tv)$ definieren, die offensichtlich Werte in S^{n-1} hat und $c(0) = x$ erfüllt. Außerdem gilt für die Ableitung $c'(t) = -\frac{\alpha'(t)}{\alpha(t)^2}(x + tv) + \frac{1}{\alpha(t)} v$ damit insbesondere $c'(0) = v$. Wir können also als Ableitungen von Kurven in S^{n-1} in $x \in S^{n-1}$ genau jene Vektoren realisieren, die orthogonal auf x stehen. Geometrisch entspricht das der anschaulichen Tatsache, dass die Vektoren “im Punkt x”, die tangential zur Sphäre sind genau die (durch den Punkt x verschobene) Normalebene auf x bilden.

Sei nun $g : \mathbb{R}^n \to \mathbb{R}^m$ eine beliebige glatte Funktion, $x \in S^{n-1}$ ein Punkt und $v \in \mathbb{R}^n$ ein Vektor, der normal auf x steht. Dann kann man die Kurve c von oben benutzen, um $Dg(x)(v)$ als $(g \circ c)'(0)$ zu berechnen. Da aber c Werte in S^{n-1} hat, hängt $g \circ c$ und damit $Dg(x)(v)$ nur von der Einschränkung von g auf S^{n-1} ab.

Die Grundaufgabe der Differentialrechnung ist nun, Eigenschaften einer glatten Funktion f aus Eigenschaften Ihrer Ableitung Df zu folgern. Das prototypische Resultat in dieser Richtung ist der aus der Analysis bekannte inverse Funktionensatz, der ein absolutes Schlüsselresultat der Analysis darstellt. Um den Satz zu formulieren, führen wir etwas Terminologie ein. Ein *Diffeomorphismus* ist eine bijektive glatte Funktion f zwischen offenen Teilmengen, für die auch die inverse Funktion f^{-1} glatt ist. Aus der Kettenregel folgt in diesem Fall sofort, dass für jeden Punkt x im Definitionsbereich von f die Gleichung $D(f^{-1})(f(x)) = (Df(x))^{-1}$ gelten muss, also ist in so einem Fall jede Ableitung von f eine invertierbare lineare Abbildung. Insbesondere folgt daraus, dass die beiden beteiligten Räume die gleiche Dimension haben müssen. Der inverse Funktionensatz sagt nun, dass die offensichtlich notwendige Bedingung der Invertierbarkeit der Ableitung in einem Punkt lokal auch hinreichend ist.

Satz 1.1. *Sei $U \subset \mathbb{R}^n$ eine offene Teilmenge und $f : U \to \mathbb{R}^n$ eine glatte Funktion und $x \in U$ ein Punkt. Ist $Df(x) : \mathbb{R}^n \to \mathbb{R}^n$ invertierbar, dann gibt es eine offene Teilmenge $V \subset U$ mit $x \in V$, sodass $f(V) \subset \mathbb{R}^n$ offen und $f|_V : V \to f(V)$ ein Diffeomorphismus ist.*

1.2. Räume von Matrizen. Wie aus der linearen Algebra bekannt, bilden die Matrizen einer fixen Größe mit Eintragungen aus einem Körper $\mathbb{K}$ eine Vektorraum über $\mathbb{K}$. Die Vektorraumoperationen (Addition und Multiplikation mit einem Skalar aus $\mathbb{K}$) sind hier einfach komponentenweise definiert. Für $A = (a_{ij})$ und $B = (b_{ij})$ ist also $A + B = (c_{ij})$ mit $c_{ij} = a_{ij} + b_{ij}$ und für $r \in \mathbb{K}$ ist $rA = (ra_{ij})$. Betrachten wir insbesondere $\mathbb{K} = \mathbb{R}$ und den Raum $M_n(\mathbb{R})$ der $n \times n$–Matrizen, dann erhalten wir $\mathbb{R}^{n^2}$ wobei die Komponenten eines Vektors genau die Eintragungen der entsprechenden Matrix sind.

Es stellt sich heraus, dass viele Konzepte der linearen Algebra, wenn man sie auf reelle (oder komplexe) Matrizen anwendet, interessante Beispiele für die Konzepte aus der Analysis liefern. Betrachten wir zum Beispiel die Teilmengen $GL(n, \mathbb{R}) \subset M_n(\mathbb{R})$ (die Notation werden wir später noch erklären), die aus allen invertierbaren Matrizen besteht. Für eine invertierbare Matrix A werden wir wie üblich die inverse Matrix mit A^{-1} bezeichnen. Das folgende Resultat sollte aus der Analysis bekannt sein, es spielt eine wichtige Rolle im Beweis des inversen Funktionensatzes.

Proposition 1.2. *Die Teilmenge $GL(n, \mathbb{R}) \subset R^{n^2} = M_n(\mathbb{R})$ ist offen. Außerdem ist die Inversionsabbildung $\nu : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$, die definiert ist durch $\nu(A) := A^{-1}$ glatt.*

Beweis. Wir benutzen die Determinantenfunktion det, sowie die aus der linearen Algebra bekannte Tatsache, dass eine Matrix $A \in M_n(\mathbb{R})$ genau dann invertierbar ist, wenn $\det(A) \neq 0$ gilt. Für die Determinante gibt es die Leibniz Formel, die zwar zur Berechnung der Determinante großer Matrizen nicht handlich, hier aber sehr nützlich ist. Für $A = (a_{ij})$ gilt

$$\det(A) = \textstyle\sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

wobei die Summe über alle Permutationen der Menge $\{1,\dots,n\}$ geht und $\operatorname{sgn}(\sigma)$ das übliche Signum der Permutation ist. Jedenfalls ist $\det(A)$ ein Polynom in den Eintragungen von A und damit eine glatte und insbesondere stetige Funktion $\det : \mathbb{R}^{n^2} \to \mathbb{R}$. Wie oben beobachtet ist aber $GL(n,\mathbb{R})$ gerade das Urbild der Teilmenge $\mathbb{R}\setminus\{0\}$ unter der stetigen Funktion det und damit ebenfalls offen.

Nach Standardresultaten der Analysis folgt aus der Glattheit von $\det : \mathbb{R}^n \to \mathbb{R}$ auch die Glattheit von $\det : GL(n,\mathbb{R}) \to \mathbb{R}\setminus\{0\}$ und daraus wieder die Glattheit von $A \mapsto \det(A)^{-1}$ als Funktion auf $GL(n,\mathbb{R})$. Nach der Cramer'schen Regel kann man aber die Eintragungen b_{ij} der Matrix A^{-1} durch $b_{ij} = (-1)^{i+j}\frac{\det(A_{ji})}{\det(A)}$ berechnen. Dabei ist A_{ji} die $(n-1)\times(n-1)$–Matrix die entsteht, wenn man in A die j–te Zeile und die i–te Spalte streicht. Jedenfalls ist $(-1)^{i+j}\det(A_{ji})$ wieder ein Polynom in den Eintragungen von A und damit hängt b_{ij} glatt von A ab. □

Da $GL(n,\mathbb{R})$ eine offene Teilmenge in $M_n(\mathbb{R})$ ist, können wir für jede glatte Funktion $f : GL(n,\mathbb{R}) \to \mathbb{R}^m$ (für beliebiges m) und eine Matrix $A \in GL(n,\mathbb{R})$ die Ableitung $Df(A)$ als lineare Abbildung $M_n(\mathbb{R}) \to \mathbb{R}^m$ betrachten. Damit besteht die Möglichkeit, solche Ableitungen in Termen von Matrizenoperationen zu schreiben.

Aufbauend auf diese Beobachtung können wir nun die Matrixmultiplikation, die eigentlich noch einfacher ist, als die Inversion, vom Standpunkt der Analysis aus studieren. Ungewohnt ist hier wohl hauptsächlich die Betrachtungsweise an die wir uns aber im Laufe der Vorlesung ohnehin gewöhnen müssen. Man kann die Matrixmultiplikation (für quadratische Matrizen) als Funktion $\mu : M_n(\mathbb{K}) \times M_n(\mathbb{K}) \to M_n(\mathbb{K})$ auffassen, die durch $\mu(A,B) = AB$ definiert ist. Wie in der Analysis üblich kann man bei einer Abbildung in zwei Variablen (die beide aus vielen Komponenten bestehen) eines der Argumente festhalten und sie als Funktion im anderen Argument betrachten. Wendet man das hier an, dann erhält man die *Linksmultiplikation mit A* als Funktion $\lambda_A : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ definiert durch $\lambda_A(B) = AB$ bzw. die *Rechtsmultiplikation mit B* als Funktion $\rho^B : M_n(\mathbb{K}) \to M_n(\mathbb{K})$, die durch $\rho^B(A) = AB$ definiert ist. Oft bezeichnet man diese Abbildungen auch als *Links–* bzw. *Rechtstranslationen.*

Aus der linearen Algebra ist nun einerseits bekannt, dass das Produkt von zwei invertierbaren Matrizen ebenfalls invertierbar ist (es gilt ja sogar $\det(AB) = \det(A)\det(B)$). Damit kann man für $A,B \in GL(n,\mathbb{R})$ die Abbildungen λ_A und ρ^B auf $GL(n,\mathbb{R})$ einschränken und erhält Funktionen $\lambda_A, \rho^B : GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$. Für die Abbildung μ selbst, kann man beobachten, dass $GL(n,\mathbb{R}) \times GL(n,\mathbb{R}) \subset M_n(\mathbb{R}) \times M_n(\mathbb{R})$ wieder eine offene Teilmenge ist, und das die Einschränkung von μ eine Funktion $\mu : GL(n,\mathbb{R}) \times GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ definiert. Andererseits wissen wir, dass die Matrizenmultiplikation assoziativ ist, also für quadratische Matrizen $A,B,C \in M_n(\mathbb{R})$ immer $A(BC) = (AB)C$ gilt. In Termen der einseitigen Multiplikationen kann man diese Gleichungen als $\lambda_A(\lambda_B(C)) = \lambda_{AB}(C)$ und als $\rho^{BC}(A) = \rho^C(\rho^B(A))$ lesen. Damit gilt aber immer $\lambda_A \circ \lambda_B = \lambda_{AB}$ und $\rho^{BC} = \rho^C \circ \rho^B$ (was auch erklärt, warum wir die Matrix einmal unten und einmal oben schreiben).

Satz 1.2. *(1) Die Matrixmultiplikation ist eine glatte Abbildung*

$$\mu : GL(n,\mathbb{R}) \times GL(n,\mathbb{R}) \to GL(n,\mathbb{R}).$$

Für jede Matrix $A \in GL(n,\mathbb{R})$ sind die Linkstranslation und die Rechtstranslation $\lambda_A, \rho^A : GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ sogar Diffeomorphismen.

(2) Für die Ableitungen gilt $D\lambda_A(B)(X) = AX$, $D\rho^A(B)(X) = XA$ und

$$D\mu(A,B)(X,Y) = AY + XB$$

für $A, B \in GL(n, \mathbb{R})$ und $X, Y \in M_n(\mathbb{R})$.

(3) Die Ableitung der Inversionsabbildung $\nu : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ ist gegeben durch $D\nu(A)(X) = -A^{-1}XA^{-1}$. Insbesondere gilt für die Einheitsmatrix $\mathbb{I}$ die Gleichung $D\nu(\mathbb{I}) = -\operatorname{id} : M_n(\mathbb{R}) \to M_n(\mathbb{R})$.

Beweis. (1) Sind $A = (a_{ij})$ und $B = (b_{ij})$ dann ist aus der linearen Algebra bekannt, dass für $AB = (c_{ij})$ die Gleichung $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$ gilt. Damit folgt wieder, dass die Eintragungen von $\mu(A, B)$ Polynome in den Eintragungen von A und B sind, also ist μ glatt als Funktion $M_n(\mathbb{R}) \times M_n(\mathbb{R}) \to M_n(\mathbb{R})$. Damit folgt auch die Glattheit der Einschränkung von μ, und damit direkt die Glattheit von λ_A und ρ^A sowohl auf $M_n(\mathbb{R})$ als auch auf $GL(n, \mathbb{R})$. Ist $\mathbb{I}$ die $n \times n$–Einheitsmatrix, dann gilt natürlich $\lambda_{\mathbb{I}} = \operatorname{id} = \rho^{\mathbb{I}}$ auf $M_n(\mathbb{R})$ und damit auch auf $GL(n, \mathbb{R})$. Für $A \in GL(n, \mathbb{R})$ ist auch $A^{-1} \in GL(n, \mathbb{R})$ und $AA^{-1} = A^{-1}A = \mathbb{I}$. Damit folgt aus den Überlegungen von oben sofort $\lambda_A \circ \lambda_{A^{-1}} = \lambda_{AA^{-1}} = \lambda_{\mathbb{I}} = \operatorname{id}$ und analog für $\lambda_{A^{-1}} \circ \lambda_A$. Damit ist die glatte Funktion $\lambda_{A^{-1}}$ invers zu λ_A, also λ_A ein Diffeomorphismus. Das Argument für ρ^A ist ganz analog.

(2) Wir bemerken zunächst, dass wir nur auf offenen Teilmengen arbeiten und daher Richtungsableitungen benutzen können. Aus der Formel für die Matrixmultiplikation von oben folgt aber sofort, dass die Matrixmultiplikation bilinear ist, also $(A_1+rA_2)B = A_1B + rA_2B$ und $A(B_1 + rB_2) = AB_1 + rAB_2$ erfüllt. Damit erhalten wir aber sofort $\lambda_A(B+tX) = AB+tAX$ und differenzieren bei $t = 0$ liefert $D\lambda_A(B)(X) = AX$. Analog folgt $D\rho^A(B)(X) = XA$. Für die Abbildung μ beobachten wir, dass

$$(A, B) + t(X, Y) = (A + tX, B + tY)$$

gilt und wendet man darauf μ an, dann erhält man $AB + t(AY + XB) + t^2XY$. Differenzieren bei $t = 0$ liefert dann $D\mu(A, B)(X, Y) = AY + XB$.

(3) Hier können wir auf raffinierte Weise die Kettenregel benutzen. Betrachten wir die Funktion $\varphi : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})\times GL(n, \mathbb{R})$, die durch $\varphi(A) := (A, A^{-1})$ gegeben ist. Ihre Ableitung ist natürlich durch $D\varphi(A)(X) = (X, D\nu(A)(X))$ gegeben. Nun ist aber nach Definition $\mu \circ \varphi$ die konstante Abbildung $\mathbb{I}$, also gilt $0 = D(\mu \circ \varphi)(A)(X)$ für alle $A \in GL(n, \mathbb{R})$ und $X \in M_n(\mathbb{R})$. Nach der Kettenregel erhalten wir

$$0 = D\mu(\varphi(A))(D\varphi(A)(X)) = D\mu(A, A^{-1})(X, D\nu(A)(X)) = A(D\nu(A)(X)) + XA^{-1},$$

wobei wir im letzten Schritt Teil (2) verwendet haben. Also gilt $A(D\nu(A)(X)) = -XA^{-1}$ und die Behauptung folgt sofort. □

1.3. Gruppen von invertierbaren Matrizen. Für jeden Körper $\mathbb{K}$ ist die Multiplikation auf $GL(n, \mathbb{K})$ assoziativ und besitzt ein neutrales Element $\mathbb{I}$. Für jede Matrix $A \in GL(n, \mathbb{K})$ ist die inverse Matrix A^{-1} ein multiplikativ inverses Element. Damit macht die Matrixmultiplikation $GL(n, \mathbb{K})$ zu einer Gruppe ("general linear group") und man kann Untergruppen davon betrachten. So eine Untergruppe ist eine Teilmenge $G \subset GL(n, \mathbb{K})$, sodass $\mathbb{I} \in G$ gilt, für $A \in G$ auch $A^{-1} \in G$ liegt und für zwei Elemente $A, B \in G$ auch das Produkt AB in G liegt. Unter diesen Voraussetzungen kann man die Multiplikation zu einer Abbildung $G \times G \to G$ einschränken, die G zu einer Gruppe macht.

Die lineare Algebra liefert Beispiele für solche Untergruppen, die automatisch eine schöne Interpretation als Symmetriegruppen besitzen. Zum Beispiel können wir für einen beliebigen Körper $\mathbb{K}$ die Teilmenge $SL(n, \mathbb{K}) = \{A : \det(A) = 1\} \subset GL(n, \mathbb{K})$ betrachten. Die Tatsache, dass dies eine Untergruppe ist, folgt direkt aus $\det(AB) = \det(A)\det(B)$ und $\det(\mathbb{I}) = 1$. In den Fällen $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$ die uns hier primär interessieren, liefert die Multiplikativität der Determinante weitere Untergruppen, die zwischen $SL(n, \mathbb{K})$ und $GL(n, \mathbb{K})$ liegen. Für $\mathbb{K} = \mathbb{R}$ kann man $GL^+(n, \mathbb{R}) = \{A : \det(A) >$

$0\}$ betrachten, für $\mathbb{K} = \mathbb{C}$ die Teilmenge $\{A : |\det(A)| = 1\}$. Für $\mathbb{K} = \mathbb{R}$ erhalten wir eine Interpretation als Symmetriegruppe. Man kann eine Matrix $A \in M_n(\mathbb{R})$ als Kollektion $(a_1, \dots, a_n)$ ihrer Spaltenvektoren $a_i \in \mathbb{R}^n$ interpretieren. Dann gilt $\det(A) \neq 0$ genau dann, wenn die a_i eine Basis von $\mathbb{R}^n$ bilden. Da dann (für $\mathbb{K} = \mathbb{R}$!) entweder $\det(A) > 0$ oder $\det(A) < 0$ gelten muss, erhält man zwei Klassen von geordneten Basen, die man als "positiv orientiert" bzw. "negativ orientiert" bezeichnet. Aus den Definitionen folgt nun sofort, dass eine invertierbare Matrix A genau dann in $GL^+(n, \mathbb{R})$ liegt, wenn für jede positiv orientierte geordnete Basis $(v_1, \dots, v_n)$ auch die Basis $(Av_1, \dots, Av_n)$ positiv orientiert ist. Aus dieser Beschreibung ist offensichtlich, dass dies eine Untergruppe definiert.

Für $\mathbb{K} = \mathbb{R}$ kann man weiters $|\det(A)|$ als das Volumen des von den Spaltenvektoren von A aufgespannten Parallelepipeds interpretieren. Daraus sieht man leicht, dass eine invertierbare Matrix A genau dann in $SL(n, \mathbb{R})$ liegt, wenn für jede positiv orientierte geordnete Basis $(v_1, \dots, v_n)$ die Basis $(Av_1, \dots, Av_n)$ ebenfalls positiv orientiert ist und das gleiche Volumen wie $(v_1, \dots, v_n)$ aufspannt. Man kann also $SL(n, \mathbb{R})$ als jene Matrizen interpretieren die orientierungserhaltend und volumserhaltend auf $\mathbb{R}^n$ wirken, und daraus ist wieder klar, dass es sich um eine Untergruppe handelt.

Schließlich liefern innere Produkte Beispiele für Untergruppen von $GL(n, \mathbb{K})$ für $\mathbb{K} = \mathbb{R}$ und $\mathbb{C}$. Für $\mathbb{K} = \mathbb{R}$ kann man die Bedingung $\langle Av, Aw \rangle = \langle v, w \rangle$ für das standard innere Produkt und alle Vektoren $v, w \in \mathbb{R}^n$ äquivalent als $A^t = A^{-1}$ formulieren. Die entsprechenden Matrizen heißen *orthogonale* Matrizen und die Menge aller dieser Matrizen wird mit $O(n)$ bezeichnet. Dann ist natürlich $\mathbb{I} \in O(n)$ und aus beiden Sichtweisen folgt leicht, dass mit $A, B \in O(n)$ auch $A^{-1} \in O(n)$ und $AB \in O(n)$ gilt. Analog funktioniert das mit dem üblichen Hermite'schen inneren Produkt auf $\mathbb{C}^n$, man erhält dann die Untergruppe $U(n) = \{A \in M_n(\mathbb{C}) : A^* = A^{-1}\}$. Hier bezeichnet A^* die adjungierte Matrix zu A, also $A^* = (\overline{A})^t$.

Einer der Kernpunkte dieser Vorlesung ist, dass eine große Klasse von Untergruppen von $GL(n, \mathbb{R})$ automatisch besonders schöne Teilmengen von $M_n(\mathbb{R})$ sind, nämlich sogenannte Teilmannigfaltigkeiten, auf die man die Differentialrechnung verallgemeinern kann. Das können wir in unseren Beispielen schon ansatzweise sehen. Ein sehr einfaches Beispiel einer Teilmannigfaltigkeit ist die Einheitssphäre $S^{k-1} \in \mathbb{R}^k$, die schon in Beispiel 1.1 vorgekommen ist und der einfachste Fall davon ist der Einheitskreis $S^1 \subset \mathbb{R}^2$. Tatsächlich kann man leicht sehen, dass zwei unserer Untergruppen wie der Einheitskreis aussehen. Einerseits können wir $U(1) = \{z \in \mathbb{C} : z\bar{z} = 1\}$ betrachten. Für $z \in \mathbb{C} = \mathbb{R}^2$ ist $z\bar{z} = |z|^2$, also ist $U(1)$ der Einheitskreis in $\mathbb{C}$.

Eine etwas komplizierte Realisierung des Einheitskreises liefert die Untergruppe $SO(2) := \{A \in O(2) : \det(A) = 1\} \subset GL(2, \mathbb{R})$. Für eine Matrix $A \in SO(2)$ müssen die Spaltenvektoren (a_1, a_2) eine positiv orientierte Orthonormalbasis bilden. Insbesondere ist $a_1 \in \mathbb{R}^2$ ein Einheitsvektor, also ein Punkt am Einheitskreis. Ist $a_1 = \binom{x}{y}$ mit $x^2 + y^2 = 1$, dann muss wegen $a_2 \perp a_1$ und $|a_2| = 1$ automatisch $a_2 = \pm\binom{-y}{x}$ gelten. Aus $\det(A) = 1$ folgt dann $a_2 = \binom{-y}{x}$ also ist die Matrix A durch a_1 eindeutig bestimmt, also sieht auch $SO(2)$ wie der Einheitskreis S^1 aus (der allerdings diesmal als Teilmenge von $\mathbb{R}^4$ realisiert ist).

1.4. Das Matrizenexponential. Man kann eine quadratische Matrix $A \in M_n(\mathbb{K})$ mit sich selbst multiplizieren und erhält damit $A^2 \in M_n(\mathbb{K})$ und induktiv die Potenzen $A^k \in M_n(\mathbb{K})$ für $k \in \mathbb{N}$. Nachdem es auch kein Problem ist, Linearkombinationen von Matrizen gleicher Größe zu bilden, kann man Matrizen in Polynome einsetzen. Für ein Polynom $p(t) = \sum_{k=0}^N c_k t^k$ definiert man einfach $p(A) := \sum_{k=0}^N c_k A^k$. Das ist eines der zentralen technischen Hilfsmittel in einigen fortgeschrittenen Teilen der linearen Algebra. Betrachtet man im Fall von $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ Polynome als Funktionen $\mathbb{K} \to \mathbb{K}$, dann kann man von diesen auf konvergente Potenzreihen übergehen, was einige der wichtigsten Funktionen der Analysis liefert. Es stellt sich heraus, das man auch in konvergente Potenzreihen Matrizen einsetzen kann, was für unsere Zwecke sehr nützlich sein wird. Wir begnügen uns damit, das für die Exponentialfunktion zu besprechen, das ist der einzige Fall, den wir benötigen werden.

Um über Konvergenz von Potenzreihen sprechen zu können, müssen wir die Topologie auf $M_n(\mathbb{R}) = \mathbb{R}^{n^2}$ benutzen. Die Resultate aus der Analysis lassen sich am einfachsten benutzen, wenn man die Topologie über eine Norm beschreibt. Andererseits wollen wir ja die Matrizenmultiplikation benutzen, also sollte die Norm auch mit der Matrizenmultiplikation verträglich sein. Die eleganteste Lösung für dieses Problem ist ein Resultat aus der Topologie zu benutzen, das besagt, dass jede Norm auf einem endlichdimensionalen reellen Vektorraum die übliche Topologie liefert. Damit kann man als Norm die sogenannte Operatornorm verwenden, die nach Definition mit der Komposition von Operatoren (und daher mit der Matrixmultiplikation) verträglich ist. Als Alternative zeigen wir direkt, dass auch die übliche (euklidische) Norm auf $M_n(\mathbb{K})$ für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ mit der Matrixmultiplikation verträglich ist. Wir betrachten also für eine Matrix $A = (a_{ij}) \in M_n(\mathbb{K})$ die Norm $\|A\| := \sqrt{\sum_{i,j} |a_{ij}|^2}$.

Lemma 1.4. *Für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ seien $A, B \in M_n(\mathbb{K})$ Matrizen und sei $AB \in M_n(\mathbb{K})$ ihr Produkt. Dann ist $\|AB\| \leq \|A\|\|B\|$ und somit $\|A^k\| \leq \|A\|^k$ für alle $k \in \mathbb{N}$.*

Beweis. Für $i = 1, \dots, n$ sei a_i der $i$te Zeilenvektor von A und b^i der $i$te Spaltenvektor von B, die wir beide als Elemente von $\mathbb{K}^n$ betrachten. Dann gilt für das standard innere Produkt $\langle\ ,\ \rangle$ auf $\mathbb{K}^n$ die Gleichung $\langle a_i, a_i\rangle = \sum_{j=1}^n |a_{ij}|^2$, also ist $\|A\|^2 = \sum_{i=1}^n \langle a_i, a_i\rangle$. Ganz analog folgt auch $\|B\|^2 = \sum_{j=1}^n \langle b^j, b^j\rangle$. Schreiben wir das Produkt AB als (c_{ij}), dann ist nach Definition $c_{ij} = \langle a_i, b^j\rangle$.

Damit gilt $|c_{ij}|^2 \leq \langle a_i, a_i\rangle\langle b^j, b^j\rangle$ nach der Cauchy–Schwarz–Ungleichung. Summiert man das über alle i, j auf, dann erhält man sofort $\|AB\|^2 \leq \|A\|^2\|B\|^2$. Damit folgt die erste Behauptung sofort und die zweite durch Induktion. □

Damit können wir grundlegende Eigenschaften des Matrizenexponentials leicht beweisen:

Proposition 1.4. *Sei $X \in M_n(\mathbb{K})$ eine beliebige Matrix mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.*

(1) Die Summe $\sum_{k=0}^\infty \frac{1}{k!} X^k$ ist in $M_n(\mathbb{K})$ absolut konvergent und definiert daher ein Element $\exp(X) = e^X \in M_n(\mathbb{K})$.

(2) Die Matrix e^X ist immer invertierbar, genauer gilt $\det(e^X) = e^{\operatorname{tr}(X)}$, *wobei* $\operatorname{tr}(X)$ *die Spur von X bezeichnet.*

(3) Ist $A \in GL(n, \mathbb{K})$, dann gilt $\exp(AXA^{-1}) = A\exp(X)A^{-1}$.

(4) Sind $X, Y \in M_n(\mathbb{K})$, sodass $XY = YX$ gilt, dann ist $e^{X+Y} = e^X e^Y$.

Beweis. (1) Für jedes $N \in \mathbb{N}$ ist natürlich $Z_N := \sum_{k=0}^N \frac{1}{k!} X^k$ ein wohldefiniertes Element von $M_n(\mathbb{K})$. Für $M > N$ gilt $Z_M - Z_N = \sum_{k=N+1}^M \frac{1}{k!} X^k$ und damit

$$\|Z_M - Z_N\| \leq \sum_{k=N+1}^M \tfrac{1}{k!}\|X^k\| \leq \sum_{k=N+1}^M \tfrac{1}{k!}\|X\|^k.$$

Die rechte Seite ist aber gerade die Differenz der entsprechenden Partialsumme der Potenzreihe für $e^{\|X\|}$. Aus der Analysis ist bekannt, dass diese Potenzreihe konvergiert, also ist $(Z_N)_{N\in\mathbb{N}}$ eine Cauchy Folge in $M_n(\mathbb{R}) = \mathbb{R}^{n^2}$. Damit existieren die Limiten $\lim_{N\to\infty} Z_N$ und $\lim_{N\to\infty} \sum_{k=0}^N \frac{1}{k!}\|X\|^k$, was die absolute Konvergenz der Reihe beweist.

(3) Es gilt $(AXA^{-1})(AXA^{-1}) = AX^2A^{-1}$ und induktiv folgt sofort $(AXA^{-1})^k = AX^kA^{-1}$ für alle $k \in \mathbb{N}$. Da die Abbildung $X \mapsto AXA^{-1}$ linear ist, erhalten wir $\sum_{k=0}^{N} \frac{1}{k!}(AXA^{-1})^k = A(\sum_{k=0}^{N} \frac{1}{k!}X^k)A^{-1}$ für alle $N \in \mathbb{N}$. Bildet man auf beiden Seiten den Limes für $N \to \infty$, dann folgt die Behauptung.

(2) Wir können eine Matrix $X \in M_n(\mathbb{R})$ auch als Element von $M_n(\mathbb{C})$ betrachten. Da dies mit Matrixprodukten und Summen verträglich ist, erhalten wir in beiden Sichtweise das gleiche Element e^X nur einmal als reelle und einmal als komplexe Matrix betrachtet. Also dürfen wir ohne Beschränkung der Allgemeinheit über $\mathbb{C}$ arbeiten und in diesem Fall ist aus der linearen Algebra bekannt, dass es eine Matrix $A \in GL(n,\mathbb{C})$ gibt, sodass $Y := AXA^{-1}$ eine obere Dreiecksmatrix ist. Dann ist aber auch Y^k eine obere Dreiecksmatrix und wenn man die Eintragungen von Y auf der Hauptdiagonale mit $\lambda_1, \dots, \lambda_n$ bezeichnet, dann sind die entsprechenden Eintragungen von Y^k gerade $(\lambda_1)^k, \dots, (\lambda_n)^k$. Damit sehen wir, dass $\sum_{k=0}^{N} \frac{1}{k!}Y^k$ eine obere Dreiecksmatrix ist und kennen ihre Eintragungen auf der Hauptdiagonale. Daraus folgt aber sofort, dass e^Y ein obere Dreiecksmatrix mit Hauptdiagonaleintragungen $e^{\lambda_1}, \dots, e^{\lambda_n}$ ist. Damit ist $\det(e^Y) = e^{\lambda_1 + \dots + \lambda_n} = e^{\mathrm{tr}(Y)}$ und aus der linearen Algebra ist bekannt, dass $\mathrm{tr}(Y) = \mathrm{tr}(X)$ gilt. Nach (3) ist aber $e^Y = Ae^XA^{-1}$, also gilt auch $\det(e^Y) = \det(e^X)$.

(4) Die wesentliche Beobachtung hier ist, dass der aus der Analysis bekannte Beweis für die reelle Exponentialfunktion als wesentlichen Input nur den binomischen Lehrsatz benötigt, der für kommutierende Matrizen natürlich gilt. Wegen der absoluten Konvergenz kann man e^Xe^Y als $\sum_{k,\ell=0}^{\infty} \frac{1}{k!\ell!}X^kY^\ell$ berechnen und die Reihenfolge der Summation spielt keine Rolle. Damit kann man die Summe aber insbesondere als $\sum_{n=0}^{\infty} \sum_{k+\ell=n} \frac{1}{k!\ell!}X^kY^\ell$ berechnen und wegen $XY = YX$ ist die innere Summe gleich $\frac{1}{n!}(X+Y)^n$, also folgt $e^Xe^Y = e^{X+Y}$. □

Damit können wir das Matrizenexponential als Funktion $\exp : M_n(\mathbb{K}) \to GL(n,\mathbb{K})$ betrachten und deren Eigenschaften studieren:

Satz 1.4. *Für* $\mathbb{K} = \mathbb{R}$ *oder* $\mathbb{C}$ *ist die Funktion* $\exp : M_n(\mathbb{K}) \to GL(n,\mathbb{K})$, $\exp(X) = e^X$ *beliebig oft differenzierbar (und sogar reell analytisch) und erfüllt* $\exp(0) = \mathbb{I}$. *Die Ableitung* $D(\exp)(0) : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ *ist die Identitätsabbildung, also gibt es offene Umgebungen* U *von* 0 *in* $M_n(\mathbb{K})$ *und* V *von* $\mathbb{I}$ *in* $GL(n,\mathbb{K})$, *sodass* $\exp : U \to V$ *ein Diffeomorphismus ist. Für jede Matrix* $A \in GL(n,\mathbb{K})$ *definieren* $X \mapsto Ae^X$ *und* $X \mapsto e^XA$ *Diffeomorphismen von* U *auf offene Umgebungen von* A *in* $GL(n,\mathbb{K})$.

Beweis. Wie im Beweis von Proposition 1.4 betrachten wir für $N \in \mathbb{N}$ die Funktion $f_N : M_n(\mathbb{K}) \to M_n(\mathbb{K})$, die definiert ist durch $f_N(X) = \sum_{k=0}^{N} \frac{1}{k!}X^k$. Für jedes N ist das eine glatte Funktion (sogar ein Polynom) und wir haben schon gesehen, dass

$$(*) \qquad \|f_M(X) - f_N(X)\| \leq \sum_{k=N+1}^{M} \tfrac{1}{k!}\|X\|^k$$

gilt. Ist nun $K \subset M_n(\mathbb{K})$ eine kompakte Teilmenge, dann ist K nach dem Satz von Heine–Borel beschränkt und abgeschlossen, also $\|X\| \leq C$ für eine konstante C. Die Konvergenz der Reihe für e^C und die Monotonie der Partialsummen zeigt dann aber, dass für hinreichend großes M und N die rechte Seite von $(*)$ für alle $X \in K$ kleiner als eine beliebige vorgegebene Zahl ε ist. Das bedeutet aber gerade, dass die Exponentialreihe auf K gleichmäßig konvergiert und damit folgt die Glattheit der Exponentialfunktion aus bekannten Resultaten der Analysis. Aus Proposition 1.4 wissen wir schon, dass exp Werte in $GL(n,\mathbb{R})$ hat.

Für die Nullmatrix gilt offensichtlich $\exp(0) = \mathbb{I}$ und für $X \in M_n(\mathbb{K})$ können wir $D\exp(0)(X)$ als die Richtungsableitung $\frac{d}{dt}|_{t=0}\exp(tX)$ berechnen. Aber nach Definition ist $\exp(tX) = \sum_{k=0}^{\infty} \frac{1}{k!}t^kX^k$ und das ist eine Potenzreihe auf $\mathbb{R}$ die absolut und auf jedem kompakten Intervall gleichmäßig konvergiert. Damit ist die Ableitung in 0 aber einfach der Limes der Ableitungen von $\sum_{k=0}^{N} \frac{1}{k!} t^k X^k$ in $t = 0$, die aber offensichtlich alle gleich X sind. Damit erhalten wir $D\exp(0)(X) = X$, also ist $D\exp(0) = \mathrm{id}_{M_n(\mathbb{K})}$ und damit ein linearer Isomorphismus.

Nach dem inversen Funktionensatz (Satz 1.1) gibt es offene Umgebungen U von 0 und V von $\exp(0) = \mathbb{I}$, sodass sich exp zu einem Diffeomorphismus von U nach V einschränkt. Schließlich wissen wir für eine invertierbare Matrix $A \in GL(n, \mathbb{K})$ aus Satz 1.2, dass die Linksmultiplikation λ_A und die Rechtsmultiplikation ρ^A mit A Diffeomorphismen sind, die $\lambda_A(\mathbb{I}) = A = \rho^A(\mathbb{I})$ erfüllen. Damit sind aber $\lambda_A(V)$ und $\rho^A(V)$ offene Umgebungen von A und offensichtlich sind $\lambda_A \circ \exp : U \to \lambda_A(V)$ und $\rho^A \circ \exp : U \to \rho^A(V)$ Diffeomorphismen. □

Bemerkung 1.4. Es ist nicht schwierig, eine lokale glatte Inverse zur Exponentialfunktion explizit anzugeben. Man kann dazu die aus der Analysis bekannte Potenzreihendarstellung des Logarithmus benutzen. Für $z \in \mathbb{C}$ mit $|z| < 1$ gilt nämlich $\log(1+z) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} z^k}{k}$. Nun kann man für eine Matrix $X \in M_n(\mathbb{K})$ mit $\|X\| < 1$ den Logarithmus von $\mathbb{I} + X$ als $\log(\mathbb{I} + X) := \sum_{k=1}^{\infty} (-1)^{k+1} \frac{X^k}{k}$ definieren. Dabei stellt man analog zur Exponentialreihe fest, dass diese Reihe absolut und auf jeder kompakten Teilmenge des Einheitsballs in $M_n(\mathbb{K})$ gleichmäßig konvergiert und somit eine glatte Funktion definiert. Mit etwas Wissen über Potenzreihen kann man dann aus der Tatsache, dass $\log \circ \exp = \exp \circ \log = \mathrm{id}$ auf $\mathbb{K}$ gilt, dass exp und log auch für Matrizen invers zueinander sind, siehe Abschnitt 6.2 von [**Kühnel**].

1.5. Exponentialkurven und Differentialgleichungen. Im Beweis von Satz 1.4 haben wir für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und $X \in M_n(\mathbb{K})$ bereits die glatte Kurve $c_X : \mathbb{R} \to GL(n, \mathbb{K})$ benutzt, die durch $c_X(t) := \exp(tX)$ gegeben ist. Insbesondere haben wir gesehen, dass die Ableitung dieser Kurve in $t = 0$ durch $c_X'(0) = X$ gegeben ist. Nun haben diese *Exponentialkurven* gruppentheoretisch schöne Eigenschaften: Natürlich ist $(t+s)X = tX + sX$ und die Matrizen tX und sX kommutieren miteinander. Damit folgt aber nach Teil (4) von Proposition 1.4, dass $c_X(t+s) = e^{tX+sX} = e^{tX}e^{sX} = c_X(t) \cdot c_X(s)$ gilt. Das bedeutet aber gerade, dass c_X ein Homomorphismus von der Gruppe $(\mathbb{R}, +)$ in die Gruppe $GL(n, \mathbb{K})$ ist, der zusätzlich glatt ist. Man nennt so eine Kurve eine *Einparameter–Untergruppe* von $GL(n, \mathbb{K})$.

Diese Eigenschaft der Exponentialkurven ist eng verwandt mit einer weiteren wichtigen Eigenschaft, die sie mit der reellen Exponentialfunktion gemeinsam haben, nämlich Lösungen für grundlegende Differentialgleichungen zu liefern. Betrachtet man auf $\mathbb{R}$ die Differentialgleichung $f'(t) = af(t)$ für $a \neq 0$, dann ist die allgemeine Lösung durch $f(t) = Ce^{at}$ gegeben, wobei $C \in \mathbb{R}$ eine Konstante ist, die durch die Anfangsbedingung $C = f(0)$ festgelegt werden kann. (Um diese Lösung zu finden, bemerkt man, dass man lokal um $t \in \mathbb{R}$ mit $f(t) \neq 0$ die Gleichung als $\frac{f'(t)}{f(t)} = a$ schreiben kann. Der Ansatz $u(t) = \log(f(t))$ liefert dann $u'(t) = \frac{f'(t)}{f(t)} = a$ was als allgemeine Lösung natürlich $u(t) = u_0 + at$ hat. Damit erhält man $f(t) = e^{u(t)} = e^{u_0}e^{at}$ wie behauptet.)

Die natürliche Verallgemeinerung der Gleichung $f'(t) = af(t)$ sind *lineare Differentialgleichungen mit konstanten Koeffizienten*. Man gibt sich eine fixe $n \times n$–Matrix $X \in M_n(\mathbb{K})$ vor und sucht eine Kurve $c : \mathbb{R} \to \mathbb{K}^n$, sodass $c'(t) = Xc(t)$ für alle $t \in \mathbb{R}$ (oder für t nahe bei $0 \in \mathbb{R}$) gilt. Zusätzlich gibt man sich noch einen Anfangswert $c(0) = v_0 \in \mathbb{K}^n$ vor. Aus der Analysis ist bekannt, dass die Lösung so einer Differentialgleichung eindeutig bestimmt ist, soferne sie existiert. (Für den Beweis der Existenz und Eindeutigkeit mit Hilfe des Banach'schen Fixpunktsatzes ist nur eine lokale Lipschitz

Bedingung nötig, die für Gleichungen mit stetig differenzierbaren Koeffizienten immer erfüllt ist.)

PROPOSITION 1.5. *Sei $X \in M_n(\mathbb{K})$ eine $n \times n$–Matrix mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$.*

(1) Für $v_0 \in \mathbb{K}^n$ ist die eindeutige Lösung der Differentialgleichung $c'(t) = Xc(t)$ für eine Kurve $c : \mathbb{R} \to \mathbb{K}^n$ mit Anfangsbedingung $c(0) = v_0$ durch $c(t) = e^{tX}v_0$ gegeben.

(2) Die eindeutige Lösung der Differentialgleichung $Y'(t) = XY(t)$ für eine Kurve $Y : \mathbb{R} \to M_n(\mathbb{K})$ mit Anfangsbedingung $Y(0) = A$ ist durch $Y(t) = e^{tX}A$ gegeben.

(3) Die eindeutige Lösung der Differentialgleichung $Z'(t) = Z(t)X$ für eine Kurve $Z : \mathbb{R} \to M_n(\mathbb{K})$ mit Anfangsbedingung $Z(0) = A$ ist durch $Z(t) = Ae^{tX}$ gegeben.

BEWEIS. (2) und (3): Aus dem Beweis von Satz 1.4 wissen wir schon, dass $B(t) := e^{tX}$ eine glatte Kurve in $M_n(\mathbb{K})$ ist, die außerdem $B'(0) = X$ erfüllt. Außerdem wissen wir von oben, dass $B(t_0 + t) = B(t_0)B(t) = B(t)B(t_0)$ gilt, was man als $B(t_0 + t) = \lambda_{B(t_0)}(B(t)) = \rho^{B(t_0)}(B(t))$. Differenziert man diese Gleichung bei $t = 0$, dann erhält $B'(t_0) = D\lambda_{B(t_0)}(B(0))(B'(0)) = D\rho^{B(t_0)}(B(0))(B'(0))$. Nach Satz 1.2 ist das gleich $B(t_0)X$ beziehungsweise $XB(t_0)$. Das sagt aber genau, dass $B(t)$ sowohl die Differentialgleichung $B'(t) = XB(t)$ aus (2) als auch die Differentialgleichung $B'(t) = B(t)X$ aus (3) mit Anfangsbedingung $B(0) = \mathbb{I}$ erfüllt.

Für $A \in M_n(\mathbb{K})$ ist aber dann die Ableitung der Kurve $Z(t) = AB(t) = \lambda_A(B(t))$ durch $Z'(t) = D\lambda_A(B(t))(B'(t)) = AB'(t) = AB(t)X = Z(t)X$. Analog folgt für $Y(t) = B(t)A$ aus $B'(t) = XB(t)$ die Gleichung $Y'(t) = XY(t)$. Da natürlich $Z(0) = A = Y(0)$ gilt, sind damit (2) und (3) bewiesen.

(1) Für $B(t) = e^{tX}$ ist nach Definition $c(t) = B(t)v_0$ also gilt insbesondere $c(0) = v_0$. Andererseits ist offensichtlich $c'(t) = B'(t)v_0 = XB(t)v_0 = Xc(t)$ und die Behauptung folgt. □

BEMERKUNG 1.5. (1) Man sieht aus dem Beweis, dass die Charakterisierung der Exponentialkurven als Lösungen einer Differentialgleichung eng mit der Tatsache zusammenhängt, dass sie Einparameter–Untergruppen sind.

(2) Gemeinsam mit Proposition 1.4 liefert Proposition 1.5 explizite Formeln für die Lösungen einer beliebigen linearen Differentialgleichung mit konstanten Koeffizienten. Für eine beliebige Matrix $X \in M_n(\mathbb{C})$ können wir eine invertierbare Matrix A so wählen, dass $Y := AXA^{-1}$ Jordan'sche Normalform hat. Das liefert eine Darstellung $Y = Y_D + Y_N$, wobei Y_D eine Diagonalmatrix und Y_N eine strikte obere Dreiecksmatrix ist, was insbesondere $Y_DY_N = Y_NY_D$ impliziert. Nach Teil (4) von Proposition 1.4 folgt $e^{tY} = e^{tY_D}e^{tY_N}$ und nach Teil (3) dieser Proposition gilt $e^{tX} = A^{-1}e^{tY}A$. Hat Y_D die Eintragungen $\lambda_1, \dots, \lambda_n$ auf der Hauptdiagonale, dann ist e^{tY_D} natürlich die Diagonalmatrix mit Eintragungen $e^{t\lambda_1}, \dots, e^{t\lambda_n}$ auf der Hauptdiagonale. Andererseits ist $(Y_N)^{n+1} = 0$, und damit ist e^{tY_N} durch eine endliche Summe gegeben, die man leicht berechnen kann.

KAPITEL 2

Matrixgruppen und ihre Lie Algebren

Matrixgruppen sind als spezielle Untergruppen der Gruppen $GL(n, \mathbb{R})$ definiert. Das wichtigste Prinzip in der Theorie dieser Gruppen ist, dass man einer solchen Untergruppe $G \subset GL(n, \mathbb{R})$ einen linearen Teilraum $\mathfrak{g} \subset M_n(\mathbb{R})$ zuordnen kann. Dieser Teilraum erhält zusätzlich eine bilineare Operation, die ihn zu einer sogenannten Lie Algebra macht. Es stellt sich heraus, das diese relativ einfache algebraische Struktur unerwartet viel Information über die zugehörige Gruppe G enthält und die Verbindung zwischen G und $\mathfrak{g}$ ist der Kern der ganzen Theorie.

2.1. Definition von Matrixgruppen. Die Mischung aus Algebra und Analysis, die die Theorie der Matrixgruppen prägt, kommt schon in der Definition von Matrixgruppen zum Ausdruck, die eine algebraische und eine topologische Bedingung enthält. Um den topologischen Teil zu formulieren, erinnern wir uns an die wichtigsten Begriffe aus diesem Bereich: Da $M_n(\mathbb{R})$ einfach ein endlichdimensionaler $\mathbb{R}$–Vektorraum ist, können wir die üblichen topologischen Begriffe verwenden. Eine Teilmenge $U \subset M_n(\mathbb{R})$ heißt *offen*, wenn es zu jedem $A \in U$ ein Zahl $\varepsilon > 0$ gibt, sodass für jede Matrix $B \in M_n(\mathbb{R})$ mit $\|B - A\| < \varepsilon$ auch $B \in U$ gilt. Andererseits nennt man eine Teilmenge $F \subset M_n(\mathbb{R})$ *abgeschlossen*, wenn das Komplement $M_n(\mathbb{R}) \setminus F$ offen ist.

Eine Teilmenge eines topologischen Raumes erbt von diesem eine Topologie, die sogenannte *Spurtopologie* oder *Teilraumtopologie*. Spezialisiert auf eine Teilmenge $C \subset M_n(\mathbb{R})$, sieht das so aus, dass man eine Teilmenge $U \subset C$ *offen in* C nennt, wenn es eine offene Teilmenge $\tilde{U} \subset M_n(\mathbb{R})$ gibt, sodass $U = C \cap \tilde{U}$ gilt. Das kann analog wie oben charakterisiert werden. Ist C selbst offen in $M_n(\mathbb{R})$, dann ist $U = C \cap \tilde{U}$ als Durchschnitt von offenen Teilmengen ebenfalls offen in $M_n(\mathbb{R})$. Analog wie zuvor nennt man eine Teilmenge $F \subset C$ *abgeschlossen in* C, wenn ihr Komplement $C \setminus F$ offen in C ist. Wie aus den Vorlesungen über Analysis und Topologie bekannt ist, gibt es dafür äquivalente Charakterisierungen:

Lemma 2.1. *Seien $F \subset C \subset M_n(\mathbb{R})$ Teilmengen. Dann sind die folgenden Bedingungen äquivalent: (1) F ist abgeschlossen in C.*

(2) Es gibt eine abgeschlossene Teilmenge $\tilde{F} \subset M_n(\mathbb{R})$, sodass $F = C \cap \tilde{F}$ gilt.

(3) Für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Elementen von F, die (in $M_n(\mathbb{R})$) gegen ein Element $A \in C$ konvergiert, gilt $A \in F$.

Bedingung (2) zeigt sofort, dass für eine abgeschlossene Teilmenge $C \subset M_n(\mathbb{R})$ eine Teilmenge $F \subset C$ genau dann abgeschlossen in C ist, wenn sie abgeschlossen in $M_n(\mathbb{R})$ ist. Ein zentrales Resultat der Topologie ist, dass Urbilder von offenen bzw. abgeschlossenen Mengen unter stetigen Funktionen selbst offen bzw. abgeschlossen sind.

Definition 2.1. Eine *Matrixgruppe* ist eine Teilmenge $G \subset GL(n, \mathbb{R})$, die abgeschlossen in $GL(n, \mathbb{R})$ ist und außerdem eine Untergruppe bildet, also $\mathbb{I}$ enthält und erfüllt, dass für $A, B \in G$ auch AB und A^{-1} in G liegen.

Beispiel 2.1. Wir können nun einfach zeigen, dass einige der Beispiele von Untergruppen, die wir in Abschnitt 1.3 kennen gelernt haben, tatsächlich Matrixgruppen sind.

(1) Im Beweis von Proposition 1.2 haben wir gezeigt, dass die Determinantenfunktion $\det : M_n(\mathbb{R}) \to \mathbb{R}$ stetig ist. Da $\{1\} \subset \mathbb{R}$ abgeschlossen ist, ist das Urbild $SL(n,\mathbb{R}) = \det^{-1}(\{1\})$ eine abgeschlossene Teilmenge von $M_n(\mathbb{R})$. Da aber $SL(n,\mathbb{R}) \subset GL(n,\mathbb{R})$ gilt, ist diese Teilmenge auch in $GL(n,\mathbb{R})$ abgeschlossen und damit ist $SL(n,\mathbb{R})$ eine Matrixgruppe.

(2) Analog ist die Funktion $f : M_n(\mathbb{R}) \to M_n(\mathbb{R})$, die durch $f(A) := AA^t$ gegeben ist, nach Satz 1.2 stetig und es gilt $A \in O(n)$ genau dann, wenn $f(A) = \mathbb{I}$ gilt. Damit ist $O(n) = f^{-1}(\{\mathbb{I}\})$ ebenfalls abgeschlossen in $M_n(\mathbb{R})$ und in $GL(n,\mathbb{R})$. In diesem Fall gilt aber noch mehr: Aus der linearen Algebra ist bekannt, dass eine Matrix A genau dann orthogonal ist, wenn ihre Spaltenvektoren eine Orthonormalbasis von $\mathbb{R}^n$ bilden. Daraus folgt aber sofort, dass für $A = (a_{ij}) \in O(n)$ immer $|a_{ij}| \leq 1$ für alle i und j gilt. Damit hat aber A als Element von $\mathbb{R}^{n^2}$ sicher Norm kleiner als n^2. Somit ist die Teilmenge $O(n) \subset M_n(\mathbb{R})$ nicht nur abgeschlossen, sondern auch beschränkt, also ist nach dem Satz von Heine–Borel $O(n)$ eine *kompakte* Teilmenge von $M_n(\mathbb{R})$.

(3) Etwas komplizierter ist die Situation mit $GL^+(n,\mathbb{R}) = \{A : \det(A) > 0\}$. Einerseits ist natürlich $(0,\infty) \subset \mathbb{R}$ offen, also ist $GL^+(n,\mathbb{R})$ eine offene Teilmenge von $M_n(\mathbb{R})$ und von $GL(n,\mathbb{R})$. Andererseits ist aber $[0,\infty) \subset \mathbb{R}$ abgeschlossen, also ist $\{A : \det(A) \geq 0\}$ eine abgeschlossene Teilmenge von $M_n(\mathbb{R})$, deren Durchschnitt mit $GL(n,\mathbb{R})$ genau $GL^+(n,\mathbb{R})$ ist. Das bedeutet aber genau, dass $GL^+(n,\mathbb{R})$ als Teilmenge von $GL(n,\mathbb{R})$ sowohl offen, als auch abgeschlossen ist. Damit ist also auch $GL^+(n,\mathbb{R})$ eine Matrixgruppe.

Ganz analog ist die Teilmenge $\{A : \det(A) < 0\} \subset GL(n,\mathbb{R})$ sowohl offen als auch abgeschlossen. Diese beiden Teilmengen liefern eine *Disjunktion* von $GL(n,\mathbb{R})$, was zeigt, dass $GL(n,\mathbb{R})$ nicht zusammenhängend ist, siehe 2.9.

(4) Ähnliches passiert im Fall von $O(n)$. Für eine Matrix A gilt $\det(A) = \det(A^t)$, also folgt aus $A \in O(n)$ sofort, dass $\det(A)^2 = 1$, also $\det(A) = \pm 1$ gilt. Nun definiert man $SO(n) := \{A \in O(n) : \det(A) = 1\}$, also gilt $SO(n) = O(n) \cap GL^+(n,\mathbb{R})$ und $SO(n) = O(n) \cap SL(n,\mathbb{R})$. Damit ist $SO(n)$ eine Matrixgruppe, die als Teilmenge von $O(n)$ sowohl offen als auch abgeschlossen ist. Damit ist einerseits auch $SO(n)$ kompakt, andererseits ist auch $O(n)$ nicht zusammenhängend sondern "zerfällt in zwei Teile".

Die Verifikation der Abgeschlossenheit mit Hilfe stetiger Funktionen, die wir im Beispiel von $O(n)$ verwendet haben, ist zwar sehr handlich, es ist aber nicht so offensichtlich, wie man sie verallgemeinern kann. Mit einem etwas alternativen Argument auf Basis von Lemma 2.1 erhält man aber sofort ein sehr allgemeines Resultat.

Proposition 2.1. *Sei β eine beliebige Bilinearform auf $\mathbb{R}^n$, also eine bilineare Abbildung $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Dann ist*

$$O(\beta) := \{A \in GL(n,\mathbb{R}) : \forall v,w \in \mathbb{R}^n : \beta(Av,Aw) = \beta(v,w)\}$$

eine Matrixgruppe.

Beweis. Zunächst folgen die Untergruppeneigenschaften leicht aus der Definition. Natürlich ist $\mathbb{I} \in O(\beta)$ und für $A, B \in O(\beta)$ ist $\beta(ABv, ABw) = \beta(Bv, Bw) = \beta(v,w)$, wobei wir erst $A \in O(\beta)$ und dann $B \in O(\beta)$ benutzt haben. Schließlich ist $\beta(A^{-1}v, A^{-1}w) = \beta(AA^{-1}v, AA^{-1}w) = \beta(v,w)$, also $A^{-1} \in O(\beta)$.

Zur Verifikation der Abgeschlossenheit von $O(\beta)$ bemerken wir, dass die Funktion $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ stetig ist. Ist $v = (v_i)$ und $w = (w_j)$, dann gilt ja $\beta(v,w) =$

$\sum_{i,j} v_i w_j \beta(e_i, e_j)$, also ist das wieder ein Polynom und damit eine glatte Abbildung. Sei nun $(A_n)_{n\in\mathbb{N}}$ eine Folge in $O(\beta)$, die gegen eine invertierbare Matrix $A \in GL(n,\mathbb{R})$ konvergiert. Dann folgt aus $A = \lim_n A_n$ natürlich $Av = \lim_n A_n v$ und $Aw = \lim_n A_n w$. Also konvergiert die Folge $(A_n v, A_n w)$ in $\mathbb{R}^n \times \mathbb{R}^n$ gegen (Av, Aw) und wegen der Stetigkeit von β kann man $\beta(Av, Av) \in \mathbb{R}$ als Limes der Folge $(\beta(A_n v, A_n w))_{n\in\mathbb{N}}$ berechnen. Aber die letztere Folge ist konstant gleich $\beta(v, w)$, also ist $A \in O(\beta)$ und das Resultat folgt aus Lemma 2.1. □

2.2. Die Lie Algebra einer Matrixgruppe. Der erste Schritt zur Analyse einer Matrixgruppe $G \subset GL(n,\mathbb{R})$ aus Sicht der Analysis ist, eine "lineare Approximation" an G im Punkt $\mathbb{I} \in G$ zu finden. Die Idee dazu haben wir schon in Beispiel 1.1 für die Sphäre kennen gelernt. Für die Definition und den ersten Schritt spielt die Tatsache, dass G abgeschlossen in $GL(n,\mathbb{R})$ ist, noch keine Rolle, sie wird aber schnell wichtig werden.

Definition 2.2. Für eine Matrixgruppe $G \subset GL(n,\mathbb{R})$ betrachten wir glatte Kurven $c : I \to M_n(\mathbb{R})$, die auf einem offenen Intervall $I \subset \mathbb{R}$ definiert sind, das 0 enthält, und die $c(0) = \mathbb{I}$ und $c(t) \in G$ für alle $t \in I$ erfüllen. Für jede solche Kurve ist die Ableitung $c'(0)$ ein Element von $M_n(\mathbb{R})$ und wir definieren $\mathfrak{g} \subset M_n(\mathbb{R})$ als die Menge aller dieser Ableitungen.

Eine analoge Definition macht natürlich für beliebige Teilmengen in $\mathbb{R}^m$ Sinn. In unserem Fall können wir aber mit Hilfe der Resultate aus Kapitel 1 beweisen, dass wir einen Teilraum von $M_n(\mathbb{R})$ erhalten, was für allgemeine Teilmengen nicht gilt.

Proposition 2.2. *Für eine Untergruppe $G \subset GL(n,\mathbb{R})$ ist die in Definition 2.2 definierte Teilmenge $\mathfrak{g} \subset M_n(\mathbb{R})$ ein linearer Teilraum.*

Beweis. Natürlich erfüllt die konstante Kurve $c(t) = \mathbb{I}$ die Bedingungen aus Definition 2.2, also gilt $0 \in \mathfrak{g}$. Sei nun $X \in \mathfrak{g}$ ein Element und $c : I \to M_n(\mathbb{R})$ eine glatte Kurve wie in Definition 2.2 und sei $\lambda \in \mathbb{R}$ mit $\lambda > 0$. Ist $I = (-a, b)$, dann betrachten wir die Kurve $\tilde{c} : (-\frac{a}{\lambda}, \frac{b}{\lambda}) \to M_n(\mathbb{R})$, die durch $\tilde{c}(t) := c(\lambda t)$ definiert ist. Dann erfüllt $\tilde{c}$ die Bedingungen aus Definition 2.2 und natürlich ist $\tilde{c}'(0) = \lambda \cdot c'(0)$, also gilt $\lambda \cdot X \in \mathfrak{g}$. Anderseits können wir auch die Kurve $\hat{c} : (-b, a) \to M_n(\mathbb{R})$ betrachten, die durch $\hat{c}(t) = c(-t)$ gegeben ist und $\hat{c}'(0) = -c'(0)$ erfüllt. Somit liegt $\lambda \cdot X$ für alle $\lambda \in \mathbb{R}$ wiederum in $\mathfrak{g}$.

Nehmen wir schließlich an, dass wir Elemente $X_1, X_2 \in \mathfrak{g}$ gegeben haben. Dann finden wir für $i = 1, 2$ Kurven $c_i : I_i \to M_n(\mathbb{R})$ wie in Definition 2.2, die $c_i'(0) = X_i$ erfüllen. Nun ist aber $I := I_1 \cap I_2$ ein offenes Intervall, das 0 enthält und wir definieren eine Kurve $c : I \to M_n(\mathbb{R})$ als $c(t) := c_1(t) \cdot c_2(t)$ (Matrixprodukt). Da G eine Untergruppe ist, hat diese Kurve Werte in G und weil man sie als $c = \mu \circ (c_1, c_2)$ schreiben kann, ist sie glatt nach Teil (1) von Satz 1.2. Mit dieser Darstellung erhält man die Ableitung nach Teil (2) dieses Satzes als

$$c'(0) = D\mu(c_1(0), c_2(0))(c_1'(0), c_2'(0)) = D\mu(\mathbb{I}, \mathbb{I})(X, Y) = \mathbb{I} \cdot Y + X \cdot \mathbb{I} = X + Y,$$

und damit ist der Beweis vollständig. □

Unter Benutzung der Abgeschlossenheit von G können wir aber eine alternative Beschreibung von $\mathfrak{g}$ mittels der Exponentialfunktion aus den Abschnitten 1.4 und 1.5 erhalten.

Satz 2.2. *Sei $G \subset GL(n,\mathbb{R})$ eine Matrixgruppe und $\mathfrak{g} \subset M_n(\mathbb{R})$ die in Definition 2.2 definierte Teilmenge. Dann gilt* $\exp(tX) \in G$ *für alle $X \in \mathfrak{g}$ und $t \in \mathbb{R}$.*

Beweis. Wir beginnen mit folgender Behauptung: Ist $(X_n)_{n\in\mathbb{N}}$ eine Folge in $M_n(\mathbb{R})$, die gegen eine Matrix $X \in M_n(\mathbb{R})$ konvergiert und $(t_n)_{n\in\mathbb{N}}$ eine Folge positiver reeller Zahlen, die gegen 0 konvergiert, sodass $\exp(t_nX_n) \in G$ für alle $n \in \mathbb{N}$ gilt, dann gilt $\exp(tX) \in G$ für alle $t \in \mathbb{R}$.

Fixiere $t \in \mathbb{R}$ und für $n \in \mathbb{N}$ sei $m_n \in \mathbb{Z}$ die größte ganze Zahl, die $\leq t/t_n$ ist. Dann ist $m_nt_n \leq t$ aber $m_nt_n + t_n > t$, also $|t - m_nt_n| < t_n$, also konvergiert die Folge m_nt_n gegen t. Damit konvergiert aber die Folge $m_nt_nX_n$ in $M_n(\mathbb{R})$ gegen tX und aus der Stetigkeit von exp folgt $\exp(tX) = \lim_n \exp(m_nt_nX_n)$. Aus Abschnitt 1.5 wissen wir, dass $\exp((t_n + t_n)X_n) = \exp(t_nX_n)^2$ und $\exp(-t_nX_n) = \exp(t_nX_n)^{-1}$ gilt. Daraus folgt aber induktiv sofort, dass $\exp(m_nt_nX_n) = \exp(t_nX_n)^{m_n}$ gilt, und dieses Element liegt in G, da G eine Untergruppe ist. Wegen der Abgeschlossenheit von G folgt $\exp(tX) \in G$, was die Behauptung beweist.

Um den Beweis abzuschließen, betrachten wir nun für ein Element $X \in \mathfrak{g}$ eine Kurve $c : I \to M_n(\mathbb{R})$ wie in Definition 2.2 mit $c'(0) = X$. Aus Satz 1.4 wissen wir, dass es offene Umgebungen U von 0 in $M_n(\mathbb{R})$ und V von $\mathbb{I}$ in $GL(n, \mathbb{R})$ gibt, sodass $\exp : U \to V$ ein Diffeomorphismus, also ein glatte Bijektion mit glatter Inverser ist. Indem wir I falls nötig verkleinern, können wir wegen $c(0) = \mathbb{I}$ annehmen, dass $c(I) \subset V$ gilt. Dann ist aber $\tilde{c} := \exp^{-1} \circ c : I \to U$ eine glatte Kurve in $M_n(\mathbb{R})$ sodass $c = \exp \circ \tilde{c}$ gilt. Damit ist $X = c'(0) = D\exp(0)(\tilde{c}'(0)) = \tilde{c}'(0)$ nach Satz 1.4. Für hinreichend großes $n \in \mathbb{N}$ liegt $1/n \in I$ und nach Definition der Ableitung ist $\tilde{c}'(0) = \lim_n n\tilde{c}(1/n)$. Für solche n setzt man nun $t_n = 1/n$ und $X_n = n\tilde{c}(1/n)$, dann konvergiert X_n gegen X und t_n gegen 0 und nach Konstruktion ist $\exp(t_nX_n) = \exp(\tilde{c}(1/n)) = c(1/n) \in G$ für alle solchen n. Damit folgt der Satz aus der Behauptung. □

Da die Exponentialkurve $c(t) = \exp(tX)$ eine glatte Kurve $\mathbb{R} \to GL(n, \mathbb{R})$ ist, die $c'(0) = X$ erfüllt sieht man, dass $\mathfrak{g} = \{X \in M_n(\mathbb{R}) : \forall t \in \mathbb{R} : \exp(tX) \in G\}$ gilt.

2.3. Die Lie Klammer. Beim Studium der Exponentialfunktion in Abschnitt 1.4 hat die Frage, ob zwei Matrizen kommutieren schon eine Rolle gespielt. Allgemein kann man das Maß der Nicht–kommutativität durch den *Kommutator* messen, der für $X, Y \in M_n(\mathbb{K})$ durch $[X, Y] := XY - YX$ definiert ist. Überraschenderweise spielt dieser Kommutator auf dem "infinitesimalen" Niveau der Lie Algebra eine viel wichtigere Rolle als die Matrizenmultiplikation. Zunächst müssen wir beweisen, dass man durch Bilden eines Kommutators die Lie Algebra einer Matrixgruppe nicht verlässt, was für sich schon ein überraschendes und wichtiges Resultat ist. Der Beweis sieht ähnlich aus wie der Beweis von Proposition 2.2, hat aber doch eine etwas andere technische Grundlage, die wir sicherheitshalber explizit formulieren und beweisen:

Lemma 2.3. *Sei I ein Intervall in $\mathbb{R}$ und $c : I \to \mathbb{R}^N$ eine glatte Kurve, die Werte in einem Teilraum $V \subset \mathbb{R}^N$ hat. Dann hat auch die Ableitung $c' : I \to \mathbb{R}^N$ Werte in V.*

Beweis. Ist $k = \dim(V)$, dann findet man eine lineare Abbildung $f : \mathbb{R}^N \to \mathbb{R}^{N-k}$, sodass $\mathrm{Ker}(f) = V$ gilt. Man wählt einfach eine Basis $\{v_1, \dots, v_k\}$ von V und ergänzt sie durch Vektoren $w_1, \dots, w_{N-k}$ zu einer Basis von $\mathbb{R}^N$ und betrachtet die eindeutige lineare Abbildung f, die $f(v_i) = 0$ für alle i und $f(w_j) = e_j$ für alle j erfüllt. Nach Voraussetzung ist dann $f \circ c$ identisch 0, und ableiten ergibt $0 = (f \circ c)'(t)$ aber weil f linear ist, stimmt das mit $f(c'(t))$ überein. Damit hat c' Werte in $\mathrm{Ker}(f) = V$. □

Damit können wir nun das angekündigte Resultat beweisen:

Satz 2.3. *Für eine Matrixgruppe $G \subset GL(n, \mathbb{R})$ und die zugehörige Teilmenge $\mathfrak{g} \subset M_n(\mathbb{R})$ aus Definition 2.2 gilt:*

(1) Für $A \in G$ und $X \in \mathfrak{g}$ ist $AXA^{-1} \in \mathfrak{g}$.
(2) Für $X, Y \in \mathfrak{g}$ gilt auch $[X, Y] = XY - YX \in \mathfrak{g}$.

Beweis. (1) Nach Satz 2.2 gilt $\exp(tX) \in G$ für alle $t \in \mathbb{R}$, also sehen wir, dass auch $A\exp(tX)A^{-1} \in G$ für alle $t \in \mathbb{R}$ gilt. Nach Teil (3) von Proposition 1.4 ist aber $A\exp(tX)A^{-1} = \exp(t(AXA^{-1}))$, also folgt die Behauptung.

(2) Nach Voraussetzung finden wir eine glatte Kurve $c : I \to M_n(\mathbb{R})$ wie in Definition 2.2 sodass $c'(0) = X$ gilt. Nach Proposition und Satz 1.2 definiert dann auch $\tilde{c}(t) := c(t) \cdot Y \cdot c(t)^{-1}$ eine glatte Kurve $I \to M_n(\mathbb{R})$ und nach Teil (1) hat diese Kurve Werte im Teilraum $\mathfrak{g} \subset M_n(\mathbb{R})$. Damit liegt nach Lemma 2.3 insbesondere $\tilde{c}'(0)$ in $\mathfrak{g}$. Nun können wir $\tilde{c}$ als $\mu \circ (\rho^Y \circ c, \nu \circ c)$ schreiben und damit erhalten wir

$$\tilde{c}'(0) = D\mu(Y, \mathbb{I})(D\rho^Y(\mathbb{I})(c'(0)), D\nu(\mathbb{I})(c'(0))) = D\mu(Y, \mathbb{I})(XY, -X) = -YX + XY.$$

□

Damit können wir den Kommutator als Operation auf $\mathfrak{g}$ also als Abbildung $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ betrachten. Die grundlegenden Eigenschaften dieser Operation sind zwar ungewohnt, sie erweisen sich aber als ähnlich nützlich wie die gewohnten Eigenschaften wie Assoziativität und Kommutativität.

Proposition 2.3. *Der Kommutator $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ ist eine bilineare Abbildung, die schiefsymmetrisch ist, also $[Y, X] = -[X, Y]$ erfüllt. Außerdem gilt die sogenannte Jacobi Identität, nämlich*

$$\forall X, Y, Z \in \mathfrak{g} : \quad [X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]].$$

Beweis. Das sind einfache Rechnungen, in denen der konkrete Teilraum $\mathfrak{g}$ keine Rolle spielt. Wir haben bereits in Abschnitt 1.2 gesehen, dass die Matrizenmultiplikation eine bilineare Abbildung ist. Damit gilt $(X_1 + \lambda X_2)Y = X_1Y + \lambda X_2Y$ und $Y(X_1 + \lambda X_2) = YX_1 + \lambda YX_2$ und somit $[X_1 + \lambda X_2, Y] = [X_1, Y] + \lambda[X_2, Y]$. Da die Schiefsymmetrie aus der Definition offensichtlich ist, folgt daraus die Bilinearität.

Auch die Jacobi Identität verifiziert man einfach direkt: Nach Definition und Assoziativität der Matrizenmultiplikation ist $[[X, Y], Z] + [Y, [X, Z]]$ gleich

$$\begin{aligned} XYZ - YXZ - ZXY + ZYX + YXZ - YZX - XZY + ZXY& \\ = XYZ - XZY - YZX + ZYX = [X, [Y, Z]].& \end{aligned}$$

□

Definition 2.3. (1) Eine *Lie Algebra* über einem Körper $\mathbb{K}$ ist ein $\mathbb{K}$–Vektorraum $\mathfrak{g}$ zusammen mit einer schiefsymmetrischen, bilinearen Operation $[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, die die Jacobi Identität erfüllt. Die Operation $[\ ,\ ]$ wird als die *Lie Klammer* (der Lie Algebra) bezeichnet.

(2) Für eine Matrixgruppe $G \subset GL(n, \mathbb{R})$ heißt die Teilmenge $\mathfrak{g}$ aus Definition 2.2 zusammen mit dem Kommutator $[\ ,\ ]$ als Operation *die Lie Algebra von G*.

2.4. Beispiele. (1) Aus Beispiel 2.1 (1) wissen wir, dass $G := SL(n, \mathbb{R}) \subset GL(n, \mathbb{R})$ eine Matrixgruppe ist. Ihre Lie Algebra $\mathfrak{sl}(n, \mathbb{R})$ können wir leicht mit Hilfe von Satz 2.2 bestimmen. Nach diesem Satz gilt für eine Matrix $X \in M_n(\mathbb{R})$, dass $X \in \mathfrak{sl}(n, \mathbb{R})$ äquivalent zu $e^{tX} \in G$ für alle $t \in \mathbb{R}$ gilt. Nach Proposition 1.4 ist aber $\det(e^{tX}) = e^{\operatorname{tr}(tX)} = e^{t\operatorname{tr}(X)}$ und das ist genau dann identisch gleich 1, wenn $\operatorname{tr}(X) = 0$ gilt. Damit ist $\mathfrak{sl}(n, \mathbb{R})$ genau die Menge der spurfreien Matrizen. Diese ist ein Teilraum, weil $\operatorname{tr} : M_n(\mathbb{R}) \to \mathbb{R}$ eine lineare Abbildung ist. Aus der linearen Algebra ist auch bekannt, dass für zwei Matrizen $X, Y \in M_n(\mathbb{R})$ immer $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ und damit $\operatorname{tr}([X, Y]) = 0$

gilt. Damit bewahrt der Kommutator nicht nur den Teilraum $\mathfrak{sl}(n,\mathbb{R})$ sondern es liegen sogar alle Werte von $[\ ,\ ]$ in diesem Teilraum. Man beachte aber, dass $\mathfrak{sl}(n,\mathbb{R})$ nicht abgeschlossen unter der Matrizenmultiplikation ist, was zum Beispiel aus $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)^2 = \mathbb{I}$ folgt. Schließlich folgt aus $\operatorname{tr}(XY) = \operatorname{tr}(YX)$ auch sofort, dass $\operatorname{tr}(AXA^{-1}) = \operatorname{tr}(X)$ für alle $A \in GL(n,\mathbb{R})$ und $X \in M_n(\mathbb{R})$ gilt. Damit gilt für $X \in \mathfrak{sl}(n,\mathbb{R})$ und $A \in GL(n,\mathbb{R})$ immer $AXA^{-1} \in \mathfrak{sl}(n,\mathbb{R})$.

(2) Aus der linearen Algebra ist bekannt, wie sich Transponieren mit Produkten verträgt. Es gilt nämlich $(XY)^t = Y^tX^t$ und damit $[X,Y]^t = [Y^t,X^t]$. Sind also X und Y symmetrische Matrizen, also $X^t = X$ und $Y^t = Y$, dann ist $[X,Y]^t = [Y,X] = -[X,Y]$. Somit ist der Raum der symmetrischen Matrizen nicht abgeschlossen unter dem Kommutator und kann damit nicht Lie Algebra einer Matrixgruppe sein. Andererseits zeigt die obige Rechnung aber, dass für schiefsymmetrische Matrizen X und Y, also $X^t = -X$ und $Y^t = -Y$ auch $[X,Y]^t = [-Y,-X] = -[X,Y]$ gilt also ist der Teilraum der schiefsymmetrischen Matrizen abgeschlossen unter dem Kommutator. Wie wir bald sehen werden ist er die Lie Algebra $\mathfrak{o}(n)$ der orthogonalen Gruppe $O(n)$.

(3) Aus der linearen Algebra ist der Begriff einer *oberen Dreiecksmatrix* bekannt. Das ist eine Matrix $X = (x_{ij})$ sodass $x_{ij} = 0$ für alle $i > j$ gilt, bei der also alle Eintragungen unterhalb der Hauptdiagonale verschwinden. Natürlich bilden diese Matrizen ein Teilraum $\mathfrak{b}(n,\mathbb{R}) \subset M_n(\mathbb{R})$. Andererseits gilt für eine obere Dreiecksmatrix X natürlich $\det(X) = x_{11}\cdots x_{nn}$ und damit besteht $B(n,\mathbb{R}) := \mathfrak{b}(n,\mathbb{R}) \cap GL(n,\mathbb{R})$ genau aus jenen oberen Dreiecksmatrizen, für die alle Hauptdiagonaleintragungen $\neq 0$ sind. Für zwei obere Dreiecksmatrizen $X = (x_{ij})$ und $Y = (y_{ij})$ ist auch XY eine obere Dreiecksmatrix und die Eintragungen des Produkts auf der Hauptdiagonale sind $x_{11}y_{11}, \dots, x_{nn}y_{nn}$. Damit gilt für $A, B \in B(n,\mathbb{R})$ auch $AB \in B(n,\mathbb{R})$ und mit Hilfe der Cramer'schen Regel verifiziert man leicht, dass für $A \in B(n,\mathbb{R})$ auch $A^{-1} \in B(n,\mathbb{R})$ gilt. Damit ist $B(n,\mathbb{R}) \subset GL(n,\mathbb{R})$ eine Untergruppe und nachdem eine Matrix $A = (a_{ij}) \in GL(n,\mathbb{R})$ genau dann in $B(n,\mathbb{R})$ liegt, wenn $a_{ij} = 0$ für alle $i > j$ gilt, ist $B(n,\mathbb{R})$ abgeschlossen in $GL(n,\mathbb{R})$ und damit eine Matrixgruppe.

Aus dieser Beschreibung können wir auch direkt eine weitere Untergruppe $N(n,\mathbb{R}) \subset B(n,\mathbb{R})$ identifizieren. Betrachten wir nämlich alle jene Matrizen in $B(n,\mathbb{R})$, bei denen alle Eintragungen auf der Hauptdiagonale gleich 1 sind. Dann gilt natürlich $\mathbb{I} \in N(n,\mathbb{R})$ und von oben sehen wir, dass für $A, B \in N(n,\mathbb{R})$ auch $AB \in N(n,\mathbb{R})$ gilt. Schließlich wissen wir von oben auch, dass $A^{-1} \in B(n,\mathbb{R})$ gilt und dann folgt aber auch dass die Eintragungen dieser Matrix auf der Hauptdiagonale gerade die Zahlen $1/a_{ii}$ sein müssen. Damit ist $N(n,\mathbb{R})$ eine Untergruppe von $B(n,\mathbb{R})$ und offensichtlich ist $N(n,\mathbb{R})$ abgeschlossen in $M_n(\mathbb{R})$ und damit in $GL(n,\mathbb{R})$. Also ist auch $N(n,\mathbb{R})$ eine Matrixgruppe.

Zur Bestimmung der Lie Algebren betrachten wir zunächst eine glatte Kurve $c : I \to M_n(\mathbb{R})$ mit $c(0) = \mathbb{I}$, die wir als $c(t) = (a_{ij}(t))$ schreiben. Gilt nun $c(t) \in B(n,\mathbb{R})$ für alle t, dann sind die Eintragungen a_{ij} für $i < j$ identisch Null. Damit liegt aber $c'(0) = (a'_{ij}(0))$ in $\mathfrak{b}(n,\mathbb{R})$. Hat c sogar Werte in $N(n,\mathbb{R})$, dann sind die Kurven $a_{ii}(t)$ konstant gleich 1 für alle t, also gilt auch $a'_{ii}(0) = 0$. Damit liegt aber $c'(0)$ im Teilraum $\mathfrak{n}(n,\mathbb{R}) \subset \mathfrak{b}(n,\mathbb{R})$ der *strikten oberen Dreiecksmatrizen*, die nur oberhalb der Hauptdiagonale Eintragungen haben können, die ungleich Null sind. Somit sind die Lie Algebren von $B(n,\mathbb{R})$ bzw. $N(n,\mathbb{R})$ in $\mathfrak{b}(n,\mathbb{R})$ bzw. $\mathfrak{n}(n,\mathbb{R})$ enthalten.

Ist umgekehrt $X \in \mathfrak{b}(n,\mathbb{R})$, dann folgt nach der Beschreibung des Produkts oberer Dreiecksmatrizen von oben sofort, dass $X^2 \in \mathfrak{b}(n,\mathbb{R})$ gilt. Induktiv folgt sofort, dass $X^k \in \mathfrak{b}(n,\mathbb{R})$ liegt, womit auch $\sum_{k=0}^{N} \frac{t^k}{k!}X^k \in \mathfrak{b}(n,\mathbb{R})$ gilt. Als Teilraum in endlichdimensionalen Vektorraum $M_n(\mathbb{R})$ ist $\mathfrak{b}(n,\mathbb{R})$ abgeschlossen, also liegt auch der Limes e^{tX}

dieser Folge in $\mathfrak{b}(n,\mathbb{R})$. Weil aber e^{tX} immer invertierbar ist, liegt diese Matrix sogar in $B(n,\mathbb{R})$. Damit ist bewiesen, dass $\mathfrak{b}(n,\mathbb{R})$ tatsächlich die Lie Algebra der Matrixgruppe $B(n,\mathbb{R})$ ist. Völlig analog folgt, dass für $X \in \mathfrak{n}(n,\mathbb{R})$ auch alle Potenzen von X in $\mathfrak{n}(n,\mathbb{R})$ liegen und damit, dass $e^{tX} \in N(n,\mathbb{R})$ für alle t gilt. Damit ist $\mathfrak{n}(n,\mathbb{R})$ die Lie Algebra der Matrixgruppe $N(n,\mathbb{R})$.

(4) Analog zu Beispiel (3) können wir auch die Lie Algebra $\mathfrak{o}(\beta)$ der orthogonalen Gruppe $O(\beta)$ einer beliebigen Bilinearform β auf $\mathbb{R}^n$ aus Proposition 2.1 beschreiben:

Proposition 2.4. *Sei $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine Bilinearform und $O(\beta)$ die zugehörige orthogonale Gruppe aus Proposition 2.1. Dann ist die Lie Algebra von $O(\beta)$ gegeben durch*

$$\mathfrak{o}(\beta) := \{X \in M_n(\mathbb{R}) : \forall v, w \in \mathbb{R}^n : \beta(Xv, w) = -\beta(v, Xw)\}.$$

Insbesondere ist die Lie Algebra $\mathfrak{o}(n)$ von $O(n)$ genau der Raum der schiefsymmetrischen Matrizen aus Beispiel (2) von oben.

Beweis. Im Beweis von Proposition 2.1 haben wir gesehen, dass β eine glatte Abbildung $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ ist. Wir können die Ableitung von β als Richtungsableitung berechnen: Für $(v_1, w_1), (v_2, w_2) \in \mathbb{R}^n \times \mathbb{R}^n$ erhalten wir

$$D\beta(v_1, w_1)(v_2, w_2) = \tfrac{d}{dt}|_{t=0}\beta(v_1 + tv_2, w_1 + tw_2).$$

Da β bilinear ist, kann man die rechte Seite einfach berechnen und erhält $\beta(v_1, w_2) + \beta(v_2, w_1)$. Ist nun $c : I \to M_n(\mathbb{R})$ eine glatte Kurve mit Werten in $O(\beta)$, die $c(0) = \mathbb{I}$ erfüllt. Dann ist für beliebige Vektoren $v, w \in \mathbb{R}^n$ die Kurve $\beta(c(t)v, c(t)w)$ konstant gleich $\beta(v, w)$. Nun hat die Kurve $c(t)v$ in $\mathbb{R}^n$ den Wert $\mathbb{I}v = v$ in $t = 0$, während ihre Ableitung in $t = 0$ offensichtlich gerade $c'(0)v$ ist und analog für $c(t)w$. Differenziert man nun die konstante Kurve $\beta(c(t)v, c(t)w)$ bei $t = 0$, dann erhält man

$$0 = D\beta(v, w)(c'(0), c'(0)) = \beta(v, c'(0)w) + \beta(c'(0)v, w).$$

Damit ist die Lie Algebra von $O(\beta)$ im Teilraum $\mathfrak{o}(\beta)$ enthalten und wir können den Beweis abschließen, indem wir zeigen, dass für $X \in \mathfrak{o}(\beta)$ und $t \in \mathbb{R}$ immer $e^{tX} \in O(\beta)$ gilt. Dazu berechnen wir für Vektoren $v, w \in \mathbb{R}^n$ den Wert $\beta(e^{tX}v, e^{tX}w)$, den wir wegen der Stetigkeit und Bilinearität von β als

$$\textstyle\sum_{i,j} \beta(\frac{t^i}{i!}X^i v, \frac{t^j}{j!}X^j w) = \sum_{i,j} \frac{t^{i+j}}{i!j!}\beta(X^i v, X^j w)$$

berechnen können. Für $X \in \mathfrak{o}(n)$ gilt aber natürlich $\beta(X^i v, X^j w) = -\beta(X^{i-1}v, X^{j+1}w)$ und induktiv folgt $\beta(X^i v, X^j w) = (-1)^i\beta(v, X^{i+j}w)$. Da unsere Reihen absolut und auf kompakten Intervallen gleichmäßig konvergieren, kann man sie beliebig umordnen, und damit insbesondere unsere Summe als $\sum_{k=0}^{\infty}\sum_{i+j=k}\frac{t^{i+j}}{i!j!}\beta(X^i v, X^j w)$ berechnen. Dann kann man aber $\frac{t^k}{k!}$ aus der hinteren Summe herausheben und es bleibt nur noch $\sum_{i+j=k}(-1)^i\frac{k!}{i!j!}\beta(v, X^k w)$. Für $k = 0$ liefert das $\beta(v, w)$, für $k > 0$ kann man auch noch $\beta(v, X^k w)$ herausheben und die Summe verschwindet nach dem binomischen Lehrsatz. □

2.5. Teilmannigfaltigkeiten von $\mathbb{R}^n$. Wir machen nun einen kurzen Ausflug in allgemeine Konzepte der Analysis, die vermutlich zum Teil schon aus den Grundvorlesungen bekannt sind. Es geht dabei darum, Konzepte der Analysis, vor allem die Differentialrechnung, auf eine allgemeinere Klasse von Teilmengen der Räume $\mathbb{R}^n$ zu verallgemeinern. Als prototypisches Beispiel kann man immer die Sphäre $S^{n-1} \subset \mathbb{R}^n$ betrachten. Die grundlegende Idee ist, dass S^{n-1} lokal um jeden Punkt wie eine offene Teilmenge von $\mathbb{R}^{n-1}$ als Teilmenge von $\mathbb{R}^n$ "aussieht". Da die Ableitung einer Funktion in einem Punkt nur vom lokalen Verhalten der Funktion um diesen Punkt abhängt, sollte das ausreichen, um eine wohldefinierte Ableitung zu erhalten.

Das liefert die Idee zur allgemeinen Definition einer Teilmannigfaltigkeit, es ist aber doch etwas an technischer Komplikation nötig, um das explizit zu machen. Dafür kann man glatte Funktionen auf solchen Teilmannigfaltigkeiten sehr einfach definieren, man betrachtet einfach Funktionen, die lokal als Einschränkung einer glatter Funktionen von offenen Teilmengen von $\mathbb{R}^n$ auf die Teilmannigfaltigkeit geschrieben werden können.

Definition 2.5. Sei $M \subset \mathbb{R}^n$ eine Teilmenge und für $k \leq n$ betrachte $\mathbb{R}^k \subset \mathbb{R}^n$ als den Teilraum all jener Punkte, deren letzte $n - k$ Koordinaten 0 sind.

(1) Die Menge M heißt eine *k–dimensionale Teilmannigfaltigkeit* von $\mathbb{R}^n$, wenn es für jeden Punkt $x \in M$ eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $x \in U$ und einen Diffeomorphismus $\Phi : U \to V$ auf eine offene Teilmenge $V \subset \mathbb{R}^n$ gibt, sodass $U \cap M = \Phi^{-1}(V \cap \mathbb{R}^k)$ gilt.

(2) Für eine Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ nennt man eine Funktion $f : M \to \mathbb{R}^m$ *glatt*, wenn es für jeden Punkt $x \in M$ eine offene Teilmenge $U \subset \mathbb{R}^n$ und eine glatte Funktion $\tilde{f} : U \to \mathbb{R}^m$ gibt, sodass $\tilde{f}$ auf $U \cap M$ mit f übereinstimmt.

(3) Sind $M \subset \mathbb{R}^n$ und $N \subset \mathbb{R}^m$ Teilmannigfaltigkeiten, dann nennt man eine Funktion $f : M \to N$ *glatt*, wenn sie als Funktion $f : M \to \mathbb{R}^m$ glatt im Sinne von (2) ist.

Beispiel 2.5. (1) Wir können eine offene Teilmenge $U \subset \mathbb{R}^n$ als n–dimensionale Teilmannigfaltigkeit betrachten (wobei man $V = U$ und $\Phi = \mathrm{id}_U$ setzen kann) und werden das auch tun. Teil (2) der Definition liefert dann einfach die üblichen glatten Funktionen $U \to \mathbb{R}^m$. In diesem Sinn verallgemeinerte die Analysis auf Teilmannigfaltigkeiten die aus den Grundvorlesungen bekannte Analysis auf offenen Teilmengen von $\mathbb{R}^n$.

(2) Für die Sphäre $S^{n-1} \subset \mathbb{R}^n$ ist es leicht, Diffeomorphismen wie in der Definition zu finden. Zunächst betrachten wir den Punkt $e_1 \in S^{n-1}$ und schreiben dafür Elemente von $\mathbb{R}^n$ als (t, x) mit $t \in \mathbb{R}$ und $x \in \mathbb{R}^{n-1}$, also $e_1 = (1, 0)$. Die Teilmenge $U := \{(t, x) : t > \|x\|\} \subset \mathbb{R}^n$ ist offen als Urbild von $(0, \infty) \subset \mathbb{R}$ unter der stetigen Funktion $(t, x) \mapsto t - \|x\|$ und natürlich ist $e_1 = (1, 0) \in U$. Andererseits ist auch

$$V := \{(y, s) \in \mathbb{R}^n : y \in \mathbb{R}^{n-1}, \|y\| < 1, s \in \mathbb{R}, s > -1\}$$

eine offene Teilmenge von $\mathbb{R}^n$. Nun definieren wir $\Phi : U \to \mathbb{R}^n$ durch

$$\Phi(t, x) := (\tfrac{1}{t} \cdot x, \|(t, x)\| - 1)$$

und $\Psi : V \to \mathbb{R}^n$ durch $\Psi(y, s) = (\lambda, \lambda y)$ mit $\lambda = \lambda(y, s) := \frac{s+1}{\sqrt{1+\|y\|^2}}$. Dann sind das offensichtlich glatte Funktionen und wegen $\|\frac{1}{t}x\| = \frac{1}{t}\|x\|$ hat Φ Werte in V. Analog hat die zweite Komponente von $\Psi(y, s)$ Norm $\lambda\|y\| < \lambda$, also hat Ψ Werte in U. Nun hat aber $\Phi(\Psi(y, s)) = \Phi(\lambda, \lambda y)$ als erste Komponente offensichtlich y. Die zweite Komponente ist $\lambda\sqrt{1 + \|y\|^2} - 1 = (s + 1) - 1 = s$, also ist $\Phi \circ \Psi = \mathrm{id}_V$. Andererseits gilt für $t > 0$ die Gleichung $\|(t, x)\| = \sqrt{t^2 + \|x\|^2} = t\sqrt{1 + \frac{1}{t^2}\|x\|^2}$ und daraus folgt sofort, dass $\lambda(\frac{1}{t}x, \|(t, x)\| - 1) = t$ und somit $\Psi \circ \Phi = \mathrm{id}_U$ gilt. Also sind $\Phi : U \to V$ und $\Psi : V \to U$ inverse Diffeomorphismen und offensichtlich liegt (t, x) genau dann in S^{n-1} wenn die letzte Koordinate von $\Phi(t, x)$ gleich 0 ist.

Betrachten wir nun einen beliebigen Punkt $x \in S^{n-1}$, dann ist x ein Einheitsvektor in $\mathbb{R}^n$ und kann somit zu einer Orthonormalbasis $\{x, v_2, \dots, v_n\}$ von $\mathbb{R}^n$ ergänzt werden. Definiert man $A \in M_n(\mathbb{R})$ als die Matrix mit den Spaltenvektoren x, $v_2, \dots, v_n$, dann ist A orthogonal, also $A^t = A^{-1}$ und natürlich gilt $Ae_1 = x$. Damit ist aber $U_x := \{Av : v \in U\}$ eine offene Teilmenge von $\mathbb{R}^n$, die x enthält und wir können $\Phi_x : U_x \to V$ durch $\Phi_x(w) := \Phi(A^{-1}w)$ definieren. Analog definiert man $\Psi_x : V \to U_x$ als $\Psi_x(y,s) = A\Psi(y,s)$ und offensichtlich ist dies eine glatte Inverse zu Φ_x. Für die orthogonale lineare Abbildung A^{-1} gilt $\|A^{-1}w\| = \|w\|$ also gilt $w \in U_x \cap S^{n-1}$ genau dann wenn $A^{-1}w \in U \cap S^{n-1}$, also genau dann, wenn die letzte Komponente von $\Phi_x(w)$ gleich 0 ist. Damit erfüllt die Familie $\{(U_x, \Phi_x) : x \in S^{n-1}\}$ alle Eigenschaften aus Definition 2.5 und wir haben verifiziert, dass S^{n-1} eine $(n-1)$–dimensionale Teilmannigfaltigkeit von $\mathbb{R}^n$ ist.

Bemerkung 2.5. Die Definition für eine Teilmannigfaltigkeit, die wir hier gewählt haben, wird als Existenz von "lokalen Trivialisierungen" bezeichnet. Praktisch arbeitet man oft mit äquivalenten Bedingungen, die leichter zu verifizieren sind, insbesondere mit "lokalen Parametrisierungen" und lokalen Darstellungen als "reguläre Nullstellenmenge". Die Äquivalenz dieser Konzepte zu beweisen ist etwas aufwändig, also verschieben wir das zum Teil in die Übungen und zum Teil auf später.

2.6. Tangentialraum und Tangentialabbildung. Der Schlüssel zur Verallgemeinerung der Analysis auf Teilmannigfaltigkeiten ist, dass man für glatte Funktionen in jedem Punkt eine Ableitung definieren kann, die eine lineare Approximation der Funktion in diesem Punkt bildet. Dazu benötigt man zunächst einen Vektorraum (in diesem Fall einen Teilraum von $\mathbb{R}^n$), der die Teilmannigfaltigkeit approximiert. Die Grundidee zur Definition dieses Tangentialraumes haben wir für die Sphäre schon in Beispiel 1.1 und für Matrixgruppen in Definition 2.2 kennen gelernt. In Anbetracht von Abschnitt 2.5 können wir aber unsere Terminologie etwas vereinfachen. Nach Beispiel 2.5 (1) ist ein offenes Intervall $I \subset \mathbb{R}$ eine eindimensionale Teilmannigfaltigkeit von $\mathbb{R}$ und für eine Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ ist eine Funktion $c : I \to M$ nach Definition 2.5 glatt, wenn sie glatt im üblichen Sinn als Funktion nach $\mathbb{R}^n$ ist und Werte in M hat. In diesem Sinn sprechen wir einfach von *glatten Kurven in M*.

Definition 2.6. Für eine Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ und einen Punkt $x \in M$ definiert man den *Tangentialraum T_xM an M bei x* analog zu Definition 2.2. Wir betrachten offene Intervalle $I \subset \mathbb{R}$ mit $0 \in I$ und glatte Kurven $c : I \to M$ mit $c(0) = x$ und definieren $T_xM \subset \mathbb{R}^n$ als die Menge der Vektoren $c'(0)$ für alle solchen Kurven c.

Diese Definition macht natürlich für beliebige Teilmengen von $\mathbb{R}^n$ Sinn, im allgemeinen ist aber T_xM kein linearer Teilraum von $\mathbb{R}^n$. (Man denke zum Beispiel an einen Halbraum in $\mathbb{R}^2$.) Andererseits sagt uns diese Definition des Tangentialraumes eigentlich auch schon, wie man die Ableitung einer glatten Funktion $f : M \to N$ zwischen zwei Teilmannigfaltigkeiten definieren sollte. Zunächst sollte für einen vernünftigen Glattheitsbegriff für eine glatte Kurve $c : I \to M$ und eine glatte Funktion $f : M \to N$ auch $f \circ c : I \to N$ eine glatte Kurve sein. Für einen Ableitungsbegriff, der die Kettenregel erfüllt muss man aber dann die Ableitung $(f \circ c)'(0)$ berechnen können, indem man die Ableitung von f im Punkt $c(0)$ auf $c'(0)$ anwendet. A priori ist nicht klar, dass das tatsächlich Sinn macht, der Beweis ist aber nicht so schwierig:

Satz 2.6. *(1) Für eine k–dimensionale Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ und jeden Punkt $x \in M$ ist der Tangentialraum T_xM ein k–dimensionaler Teilraum von $\mathbb{R}^n$.*

(2) Für eine glatte Funktion $f : M \to N$ zwischen Teilmannigfaltigkeiten und jeden Punkt $x \in M$ gibt es eine eindeutige lineare Abbildung $T_xf : T_xM \to T_{f(x)}N$ die dadurch charakterisiert ist, dass $T_xf(c'(0)) = (f \circ c)'(0)$ für jede glatte Kurve $c : I \to M$ wie in Definition 2.6 gilt.

(3) Sind $f : M \to N$ und $g : N \to P$ glatte Funktionen, dann ist auch die Komposition $g \circ f : M \to P$ glatt und für jeden Punkt $x \in M$ gilt die Kettenregel $T_x(g \circ f) = T_{f(x)}g \circ T_x f : T_x M \to T_{g(f(x))}P$.

Beweis. (1) Nach Definition finden wir offene Teilmengen $U, V \subset \mathbb{R}^n$ mit $x \in U$ und einen Diffeomorphismus $\Phi : U \to V$, sodass $U \cap M = \Phi^{-1}(V \cap \mathbb{R}^k)$ gilt. Dann ist $D\Phi(x) : \mathbb{R}^n \to \mathbb{R}^n$ ein linearer Isomorphismus und wir behaupten, dass $T_x M$ mit der Teilmenge $E := \{v \in \mathbb{R}^n : D\Phi(x)(v) \in \mathbb{R}^k \subset \mathbb{R}^n\}$ übereinstimmt, die offensichtlich ein k–dimensionaler Teilraum von $\mathbb{R}^n$ ist.

Sei zunächst $I \subset \mathbb{R}$ ein offenes Intervall mit $0 \in I$ und $c : I \to M$ eine glatte Kurve mit $c(0) = x$. Weil $U \subset \mathbb{R}^n$ offen ist und $x \in U$ gilt, gibt es ein offenes Teilintervall $\tilde{I} \subset I$ mit $0 \in \tilde{I}$, sodass $c(\tilde{I})$ in U und damit in $U \cap M$ enthalten ist. Damit ist aber $\Phi \circ c : \tilde{I} \to \mathbb{R}^n$ eine glatte Kurve in $\mathbb{R}^n$, die Werte im Teilraum $\mathbb{R}^k \subset \mathbb{R}^n$ hat. Nach Lemma 2.3 folgt $(\Phi \circ c)'(0) \in \mathbb{R}^k$ und nach der Kettenregel ist das $D\Phi(c(0))(c'(0))$, also gilt $c'(0) \in E$, also $T_x M \subset E$.

Umgekehrt können wir für einen Vektor $w \in \mathbb{R}^k \subset \mathbb{R}^n$ und $t \in \mathbb{R}$ den Punkt $\Phi(x) + tw \in \mathbb{R}^n$ betrachten. Da $V = \Phi(U)$ offen in $\mathbb{R}^n$ ist, finden wir ein offenes Intervall $I \subset \mathbb{R}$ mit $0 \in I$, sodass $\Phi(x) + tw \in V$ für all $t \in I$ gilt. Damit ist aber $c : I \to \mathbb{R}^n$, $c(t) := \Phi^{-1}(\Phi(x) + tw)$ eine glatte Kurve, die nach Konstruktion Werte in $U \cap M$ hat. Damit ist aber $c'(0) = D\Phi^{-1}(\Phi(x))(w) \in T_x M$, was sofort $E \subset T_x M$ und damit (1) impliziert.

(2) Nach Definition der Glattheit finden wir eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $x \in U$ und eine Funktion $\tilde{f} : U \to \mathbb{R}^m$ die glatt ist und auf $U \cap M$ mit f übereinstimmt. Für ein Intervall $I \subset \mathbb{R}$ und eine glatte Kurve $c : I \to M$ mit $0 \in I$ und $c(0) = x$ finden wir wie oben ein Teilintervall $\tilde{I} \subset I$ mit $0 \in \tilde{I}$, sodass $c(\tilde{I})$ in U und damit in $U \cap M$ enthalten ist. Damit gilt aber $f \circ c = \tilde{f} \circ c : \tilde{I} \to \mathbb{R}^m$ und aus der Analysis ist bekannt, dass dies eine glatte Kurve ist. Damit existiert $(f \circ c)'(0)$ und stimmt nach der Kettenregel der Analysis mit $D(\tilde{f})(x)(c'(0))$ überein. Daraus folgt, dass $c'(0) \mapsto (f \circ c)'(0)$ wohldefiniert ist und als Abbildung mit $D\tilde{f}(x)|_{T_x M} : T_x M \to \mathbb{R}^m$ übereinstimmt. Das vervollständigt einerseits den Beweis von (2), anderseits bemerken wir, dass $D\tilde{f}(x)(T_x M) \subset T_{f(x)}N$ gilt.

(3) Für einen Punkt $x \in M$ finden wir eine offene Teilmenge $U \subset \mathbb{R}^n$ mit $x \in U$ und eine glatte Funktion $\tilde{f} : U \to \mathbb{R}^m$, die auf $U \cap M$ mit f übereinstimmt. Analog finden wir zum Punkt $f(x) \in N$ eine offene Teilmenge $V \subset \mathbb{R}^m$ mit $f(x) \in V$ und eine glatte Funktion $\tilde{g} : V \to \mathbb{R}^\ell$, die auf $V \cap N$ mit g übereinstimmt. Da $\tilde{f}$ stetig ist, ist $\tilde{f}^{-1}(V)$ offen in $\mathbb{R}^n$ und natürlich ist $x \in \tilde{f}^{-1}(V)$. Indem wir U durch $U \cap \tilde{f}^{-1}(V)$ ersetzen dürfen wir oBdA annehmen, dass $\tilde{f}(U) \subset V$ gilt. Dann ist aber $\tilde{g} \circ \tilde{f} : U \to \mathbb{R}^\ell$ eine glatte Funktion und für $y \in U \cap M$ ist $\tilde{g}(\tilde{f}(y)) = \tilde{g}(f(y))$ und weil $f(y) \in N$ gilt, stimmt das mit $g(f(y))$ über ein. Da der Punkt x beliebig war, ist $g \circ f : M \to P$ eine glatte Funktion.

Aus Teil (2) wissen wir dann aber, dass $T_x(g \circ f)$ mit der Einschränkung von $D(\tilde{g} \circ \tilde{f})(x)$ auf $T_x M \subset \mathbb{R}^n$ übereinstimmt. Nach der Kettenregel der Analysis ist aber $D(\tilde{g} \circ \tilde{f})(x) = D\tilde{g}(\tilde{f}(x)) \circ D\tilde{f}(x)$ und aus Teil (2) wissen wir, dass auf dem Teilraum $T_x M$ die Abbildung $D\tilde{f}(x)$ mit $T_x f$ übereinstimmt. Insbesondere liegen die Werte dann in $T_{f(x)}N$, und darauf stimmt $D\tilde{g}(f(x))$ nach (2) mit $T_{f(x)}g$ überein, was den Beweis vervollständigt. □

Da für unsere Zwecke die Ableitung in einem Punkt ausreicht, skizzieren wir nur kurz, wie man in der Analysis auf Teilmannigfaltigkeiten weiter vorgeht. Zunächst betrachtet man für eine Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ das *Tangentialbündel* $TM := \{(x,v) : x \in M, v \in T_xM\} \subset \mathbb{R}^n \times \mathbb{R}^n$. Für eine glatte Funktion $f : M \to N$ kann man dann die *Tangentialabbildung* $Tf : TM \to TN$ durch $Tf(x,v) := (f(x), T_xf(v))$ definieren. Dann zeigt man, dass TM selbst eine Teilmannigfaltigkeit von $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ ist. Damit ist aber für eine glatte Funktion $f : M \to N$ zwischen Teilmannigfaltigkeiten auch $Tf : TM \to TN$ eine Funktion zwischen Teilmannigfaltigkeiten und man zeigt, dass diese Funktion ebenfalls glatt ist. Die Kettenregel bekommt dann die einfache Form $T(g \circ f) = Tg \circ Tf$. Natürliche kann man die Abbildung Tf selbst wieder differenzieren, also $TTf : TTM \to TTN$ bilden und damit eine zweite Ableitung (und analog höhere Ableitungen) definieren.

2.7. Matrixgruppen als Teilmannigfaltigkeiten. Wir können nun das fundamentale Resultat beweisen, dass die Abgeschlossenheit einer Untergruppe in $GL(n,\mathbb{R})$ schon impliziert, dass die Untergruppe eine Teilmannigfaltigkeit von $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ bildet. Damit können Methoden der Analysis auf das Studium von Matrixgruppen angewandt werden.

Satz 2.7. *Sei $G \subset GL(n,\mathbb{R})$ eine Matrixgruppe. Dann ist G eine Teilmannigfaltigkeit in $\mathbb{R}^{n^2}$ und die Lie Algebra $\mathfrak{g}$ von G stimmt mit dem Tangentialraum im neutralen Element $\mathbb{I}$ von G überein.*

Beweis. Der wesentliche Schritt im Beweis ist einen Diffeomorphismus Φ wie in Definition 2.5 rund um das neutrale Element $\mathbb{I} \in G$ zu konstruieren und dafür verwenden wir die Exponentialabbildung. Sei $\mathfrak{g}$ die Lie Algebra von G aus Definition 2.2. Nach Proposition 2.2 ist $\mathfrak{g}$ ein linearer Teilraum von $M_n(\mathbb{R})$ und wir wählen einen komplementären Teilraum $E \subset M_n(\mathbb{R})$, also einen Teilraum sodass $M_n(\mathbb{R}) = \mathfrak{g} \oplus E$ gilt. (Dazu kann man etwa eine Basis für $\mathfrak{g}$ wählen, zu einer Basis von $M_n(\mathbb{R})$ erweitern und E als jenen Teilraum definieren, der von den hinzugekommenen Basiselementen aufgespannt wird.) Dann ist E ein Teilraum der Dimension $\ell := n^2 - \dim(\mathfrak{g})$, den wir mit $\mathbb{R}^\ell$ identifizieren können.

Wir behaupten nun, dass es eine offene Umgebung W von 0 in E gibt, sodass für $X \in W$ mit $\exp(X) \in G$ schon $X = 0$ und damit $\exp(X) = \mathbb{I}$ gilt. Zum Beweis nehmen wir indirekt an, dass wir in jeder Umgebung von 0 in E ein Element $\neq 0$ finden, dessen Exponential in G liegt. Wenden wir das auf die Bälle mit Radius $1/n$ an, dann finden wir zu jedem $n \in \mathbb{N}$ ein Element $\tilde{X}_n \in E \setminus \{0\}$ sodass $\|\tilde{X}_n\| < \frac{1}{n}$ und $\exp(\tilde{X}_n) \in G$ gilt. Setzen wir nun $t_n := \|\tilde{X}_n\|$, dann gilt $0 < t_n < \frac{1}{n}$, also können wir $X_n := \frac{1}{t_n}\tilde{X}_n$ setzen. Dann bilden die t_n natürlich eine Nullfolge in $\mathbb{R}$, während $\|X_n\| = 1$ für alle $n \in \mathbb{N}$ gilt. Somit bilden die X_n eine beschränkte Folge in $E \cong \mathbb{R}^\ell$, die nach dem Satz von Heine–Borel eine konvergente Teilfolge besitzt. Indem wir zu so einer Teilfolge übergehen dürfen wir oBdA annehmen, dass die Folge $(X_n)_{n\in\mathbb{N}}$ gegen ein Element $X \in E$ konvergiert, das dann natürlich $\|X\| = 1$ erfüllen muss. Weil $\exp(t_nX_n) = \exp(\tilde{X}_n) \in G$ gilt, erfüllen die Folgen $(t_n)_{n\in\mathbb{N}}$ und $(X_n)_{n\in\mathbb{N}}$ aber genau die Bedingungen der ersten Behauptung im Beweis von Satz 2.2, also folgt aus dieser, dass $\exp(tX) \in G$ für alle $t \in \mathbb{R}$ und damit $X \in \mathfrak{g}$ gilt. Aber nach Konstruktion ist $\mathfrak{g} \cap E = \{0\}$ also erhalten wir einen Widerspruch zu $\|X\| = 1$ und die Behauptung ist bewiesen.

Damit können wir unseren Diffeomorphismus (eigentlich seine Inverse) konstruieren: Nach Satz 1.4 finden wir offene Umgebungen $\tilde{V}$ von 0 in $M_n(\mathbb{R})$ und $\tilde{U}$ von $\mathbb{I}$ in $GL(n,\mathbb{R})$, sodass sich das Matrizenexponential exp zu einem Diffeomorphismus $\tilde{V} \to \tilde{U}$ einschränkt. Betrachten wir $M_n(\mathbb{R})$ als $\mathfrak{g} \oplus E \cong \mathfrak{g} \times E$, dann können wir offene Nullumgebungen V_1 in $\mathfrak{g}$ und V_2 in E finden, sodass $V_1 \times V_2 \subset \tilde{V}$ gilt. Indem wir V_2 durch den Schnitt von V_2 mit einer Nullumgebung W aus der obigen Behauptung ersetzen, können wir weiters annehmen, dass für $Y \in V_2$ aus $\exp(Y) \in G$ schon $Y = 0$ und damit $\exp(Y) = \mathbb{I}$ folgt. Nun definieren wir $\Psi : V_1 \times V_2 \to GL(n, \mathbb{R})$ durch $\Psi(X, Y) := \exp(X)\exp(Y)$.

Schreibt man Ψ als $\mu \circ (\exp, \exp)$, dann folgt aus Satz 1.2 und Satz 1.4 sofort, dass $D\Psi(0,0)(X,Y) = D\mu(\mathbb{I},\mathbb{I})(X,Y) = X + Y$ für alle $X \in \mathfrak{g}$ und $Y \in E$ gilt. Damit ist aber $D\Psi(0,0)$ ein linearer Isomorphismus und nach dem inversen Funktionensatz finden wir eine Umgebung V von 0 in $V_1 \times V_2 \subset \mathfrak{g} \times E \cong \mathbb{R}^{n^2}$ sodass $U := \Psi(V)$ eine offene Umgebung von $\mathbb{I}$ in $GL(n, \mathbb{R})$ ist und sich Ψ zu einem Diffeomorphismus $V \to U$ einschränkt. Jedes Element von V kann man als (X, Y) mit $X \in V_1$ und $Y \in V_2$ schreiben. Liegt $\Psi(X, Y) \in U \cap G$, dann erhalten wir $\exp(X)\exp(Y) = A$ und damit $\exp(Y) = A\exp(X)^{-1}$ für eine Matrix $A \in G$. Nun ist aber $X \in \mathfrak{g}$, also $\exp(X) \in G$ und damit auch $\exp(Y) = A\exp(X)^{-1} \in G$. Aber daraus folgt nach Konstruktion $Y = 0$, also $(X, Y) \in \mathfrak{g} \cap V$. Das bedeutet aber genau, dass $\Phi := \Psi^{-1} : U \to V$ die in Definition 2.5 geforderten Eigenschaften hat.

Für ein beliebiges Element $A \in G$ ist nach Satz 1.2 $\lambda_A : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$ ein Diffeomorphismus mit Inverser $\lambda_{A^{-1}}$. Damit ist aber $U_A := \lambda_A(U)$ eine offene Teilmenge von $GL(n, \mathbb{R})$ mit $A \in U_A$ und $\Phi_A := \Phi \circ \lambda_{A^{-1}} : U_A \to V$ ein Diffeomorphismus. Ein Punkt $B \in U_A$ liegt genau dann in $U_A \cap G$, wenn $A^{-1}B \in U \cap G$, also genau dann, wenn $\Phi_A(B) = \Phi(A^{-1}B) \in \mathfrak{g} \cap V$ liegt. Damit haben wir verifiziert, dass G ein Teilmannigfaltigkeit ist und die letzte Aussage ist dann nach Definition offensichtlich. □

Mit diesem Resultat erhält man einen natürlichen Begriff für Glattheit von Funktionen zwischen Matrixgruppen. Sind $G \subset GL(n, \mathbb{R})$ und $H \subset GL(m, \mathbb{R})$ Matrixgruppen und ist $f : G \to H$ eine glatte Funktion, die $f(\mathbb{I}) = \mathbb{I}$ erfüllt, dann kann man insbesondere die Ableitung $T_{\mathbb{I}}f : T_{\mathbb{I}}G \to T_{\mathbb{I}}H$ definieren. Nach dem letzten Teil von Satz 2.7 ist diese Ableitung somit eine lineare Abbildung $\mathfrak{g} \to \mathfrak{h}$ zwischen den Lie Algebren der beiden Gruppen.

2.8. Glatte Homomorphismen und ihre Ableitungen. Wir können die Ideen aus dem letzten Abschnitt insbesondere auf Homomorphismen zwischen Matrixgruppen anwenden. Damit erhalten wir den Begriff eines *glatten Homomorphismus* $\varphi : G \to H$ zwischen zwei Matrixgruppen. Das ist einfach eine glatte Funktion, die zusätzlich mit den Gruppenstrukturen verträglich ist, also $\varphi(AB) = \varphi(A)\varphi(B)$ für alle $A, B \in G$ erfüllt. Wie aus der Algebra bekannt ist, erfüllt so ein Homomorphismus automatisch $\varphi(\mathbb{I}) = \mathbb{I}$ (und auch $\varphi(A^{-1}) = \varphi(A)^{-1}$ für alle $A \in G$). Damit können wir insbesondere die *Ableitung von* φ als $\varphi' := T_{\mathbb{I}}\varphi : \mathfrak{g} \to \mathfrak{h}$ definieren.

Zunächst können wir die Glattheit von Homomorphismen leicht charakterisieren:

Lemma 2.8. *Seien $G \subset GL(n, \mathbb{R})$ und $H \subset GL(m, \mathbb{R})$ Matrixgruppen, sei $\mathfrak{g}$ die Lie Algebra von G und sein $\varphi : G \to H$ ein Gruppenhomomorphismus. Dann ist φ genau dann glatt, wenn es eine offene Umgebung V von Null in $\mathfrak{g}$ gibt, sodass $f(X) :=$ $\varphi(\exp(X))$ eine glatte Funktion $f : V \to \mathbb{R}^{m^2}$ definiert.*

Beweis. Ist φ glatt, dann ist auch $\varphi \circ \exp : \mathfrak{g} \to H$ glatt, also auch glatt als Funktion nach $M_m(\mathbb{R}) = \mathbb{R}^{m^2}$. Nehmen wir also umgekehrt an, dass $f(X) = \varphi(\exp(X))$ glatt auf einer offenen Nullumgebung $V \subset \mathfrak{g}$ ist. Aus dem Beweis von Satz 2.7 wissen wir, dass es für ein lineares Komplement $E \subset M_n(\mathbb{R})$ zu $\mathfrak{g}$ und hinreichend kleines V, eine offene Nullumgebung $W \subset E$ gibt, sodass $\Psi(X, Y) := \exp(X)\exp(Y)$ einen Diffeomorphismus von $V \times W$ auf eine offene Teilmenge $U \subset GL(n, \mathbb{R})$ definiert, die $\mathbb{I}$ enthält. Weiters gilt für die Inverse $\Phi : U \to V \times W$ von Ψ, dass $A \in U \cap G$ äquivalent zu $\Phi(A) \in V \times \{0\}$ ist. Schreiben wir die Komponenten von Φ als $\Phi_{\mathfrak{g}}$ und Φ_E, dann sind das natürlich auch glatte Funktionen, also ist insbesondere $F := f \circ \Phi_{\mathfrak{g}} : U \to M_m(\mathbb{R})$ eine glatte Funktion (im üblichen Sinn der Analysis). Für $A \in U \cap G$ ist aber $\Phi(A) = \Phi_{\mathfrak{g}}(A)$ und damit $A = \exp(\Phi_{\mathfrak{g}}(A))$ und damit $F(A) = \varphi(A)$, also ist $F|_{U\cap G} = \varphi|_{U\cap G}$.

Für eine beliebige Matrix $A \in G$ betrachten wir nun die Linkstranslationen λ_A und $\lambda_{A^{-1}}$ auf G und $GL(n, \mathbb{R})$. Da dies inverse Diffeomorphismen sind, ist $U_A := \lambda_A(U)$ eine offene Teilmenge, die A enthält und $F \circ \lambda_{A^{-1}}$ definiert eine glatte Funktion $U_A \to M_m(\mathbb{R})$. Bezeichnet wir mit $\lambda_{\varphi(A)}$ die Linkstranslation in $M_m(\mathbb{R})$, dann ist auch $\lambda_{\varphi(A)} \circ F \circ \lambda_{A^{-1}}$ glatt. Für $B \in U_A \cap G$ ist $\lambda_{A^{-1}}(B) = A^{-1}B \in U \cap G$, also $F(\lambda_{A^{-1}}(B)) = \varphi(A^{-1}B) = \varphi(A^{-1})\varphi(B)$ und $\lambda_{\varphi(A)}$ bildet dieses Element auf $\varphi(B)$ ab. Damit stimmt aber $\lambda_{\varphi(A)} \circ F \circ \lambda_{A^{-1}}$ auf $U_A \cap G$ mit φ überein und damit ist φ glatt nach Definition 2.5. □

Als nächstes beschreiben wir die wichtigsten Eigenschaften der Ableitung eines glatten Homomorphismus und zeigen, dass die Ableitung überraschend viel Information über den Homomorphismus enthält:

Satz 2.8. *Sei $\varphi : G \to H$ ein glatter Homomorphismus zwischen Matrixgruppen mit Ableitung $\varphi' : \mathfrak{g} \to \mathfrak{h}$. Dann gelten folgende Gleichungen für beliebige Elemente $A \in G$ und $X, Y \in \mathfrak{g}$:*
(1) $\varphi(\exp(X)) = \exp(\varphi'(X))$
(2) $\varphi'(AXA^{-1}) = \varphi(A)\varphi'(X)\varphi(A)^{-1}$
(3) $\varphi'([X, Y]) = [\varphi'(X), \varphi'(Y)]$

Beweis. Da $t \mapsto \exp(tX)$ eine glatte Kurve in G ist, ist nach der Glattheit von φ auch $c(t) := \varphi(\exp(tX))$ eine glatte Kurve $c : \mathbb{R} \to GL(m, \mathbb{R})$. Natürlich gilt $c(0) = \varphi(\mathbb{I}) = \mathbb{I}$ und

$$c(t + s) = \varphi(\exp((t + s)X)) = \varphi(\exp(tX) \exp(sX)) = c(t)c(s),$$

vergleiche mit Abschnitt 1.5. Nach Definition ist $c'(0) = T_{c(0)}\varphi(X) = \varphi'(X)$ und wie in Abschnitt 1.5 erhält man daraus und aus $c(t+s) = c(t)c(s)$ sofort, dass $c'(t) = c(t)\varphi'(X)$ gilt. Nach der Eindeutigkeit der Lösung dieser Differentialgleichung in Satz 1.5 erhalten wir $c(t) = \exp(t\varphi'(X))$ und für $t = 1$ erhalten wir (1).

(2) Nach Proposition 1.4 gilt $\exp(tAXA^{-1}) = A \exp(tX)A^{-1}$ für alle $t \in \mathbb{R}$. Wendet man φ auf diese Gleichung an, dann erhält man unter Benutzung von Teil (1) $\exp(t\varphi'(AXA^{-1})) = \varphi(A) \exp(t\varphi'(X))\varphi(A)^{-1}$. Das sind glatte Kurven durch $\mathbb{I}$ und die Ableitung der linken Seite ist klarerweise durch $\varphi'(AXA^{-1})$ gegeben. Für die rechte Seite erhält man als Ableitung unter Benutzung von Satz 1.2 leicht $\varphi(A)\varphi'(X)\varphi(A^{-1})$.

(3) Wendet man (2) auf $A = \exp(tY)$ an und benutzt (1), dann erhält man

$$\varphi'(\exp(tY)X \exp(-tY)) = \exp(t\varphi'(Y))\varphi'(X) \exp(-t\varphi'(Y)). \tag{2.1}$$

Aus dem Beweis von Satz 2.3 wissen wir, das $\exp(tY)X \exp(-tY)$ ein glatte Kurve in $\mathfrak{g}$ definiert, deren Ableitung bei $t = 0$ durch $[Y, X] = YX - XY$ gegeben ist. Da φ' linear ist, ist die linke Seite von (2.1) eine glatte Kurve in $\mathfrak{h}$, deren Ableitung bei $t = 0$ durch $\varphi'([Y, X])$ gegeben ist. Aber für die rechte Seite von (2.1) erhält man als Ableitung bei $t = 0$ analog zum letzten Schritt $[\varphi'(Y), \varphi'(X)]$, also folgt die Behauptung. □

Der erste Teil des Satzes sagt uns, dass der Homomorphismus φ auf den Bild von exp (das insbesondere eine offene Umgebung von $\mathbb{I}$ in G enthält) durch die lineare Abbildung φ' eindeutig bestimmt ist. Teil (3) des Satzes sagt uns zusätzlich auch noch, dass die lineare Abbildung φ' zusätzlich auch noch ein *Homomorphismus von Lie Algebren* sein muss. Dass dies oft eine starke Einschränkung ist, zeigt folgendes einfache Beispiel:

Beispiel 2.8. Sei $G := SL(2,\mathbb{R})$ und sei $\varphi : G \to H$ ein glatter Homomorphismus in eine beliebige Matrixgruppe. Dann behaupten wir, dass die Ableitung $\varphi' : \mathfrak{g} \to \mathfrak{h}$ entweder die Nullabbildung oder injektiv sein muss.

Zum Beweis erinnern wir uns an Beispiel 2.1, wo wir gesehen haben, dass die Lie Algebra $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{R})$ gerade der Raum der spurfreien 2×2–Matrizen ist. Betrachten wir die Basis $\{E, H, F\}$, wobei $E := \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $H := \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ und $F := \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. Dann verifiziert man sofort, dass die Klammern zwischen den Basiselementen durch $[E, F] = H$, $[H, E] = 2E$ und $[H, F] = -2F$ gegeben sind (was wegen der Schiefsymmetrie alle Klammern bestimmt). Nehmen wir nun an, dass wir für einen gegebenen Homomorphismus φ ein Element $0 \neq X \in \mathfrak{g}$ finden, sodass $\varphi'(X) = 0$ gilt. Dann gilt natürlich auch $0 = [\varphi'(X), \varphi'(Y)] = \varphi'([X, Y])$ für alle $Y \in \mathfrak{g}$. Schreiben wir $X = aE + bH + cF$ für $a, b, c \in \mathbb{R}$, dann ist $[X, E] = 2bE - cH$ und $[[X, E], E] = -2cE$. Für $c \neq 0$ liegt somit ein nichttriviales Vielfaches von E im Kern von φ'. Für $c = 0$ erhalten wir schon $[X, E] = 2bE$, also liegt für $b \neq 0$ wieder ein nichttriviales Vielfaches von E im Kern von φ'. Schließlich ist für $b = c = 0$ das Element X selbst ein nichttriviales Vielfaches von E. Somit sehen wir, dass $E \in \mathrm{Ker}(\varphi')$ und damit auch $[E, F] = H \in \mathrm{Ker}(\varphi')$ und $[H, F] = -2F \in \mathrm{Ker}(\varphi')$ gilt, also ist $\varphi' = 0$.

2.9. Zusammenhang. Um genauer zu sehen, wie viel Informationen über einen glatten Homomorphismus man aus seiner Ableitung gewinnen kann, benötigen wir ein weiteres Grundkonzept der Topologie: Eine *Disjunktion* eines topologischen Raumes X ist durch zwei nichtleere, offene Teilmengen $U, V \subset X$ gegeben, die $U \cap V = \emptyset$ und $U \cup V = X$ erfüllen. Ein topologischer Raum heißt *zusammenhängend*, wenn er keine Disjunktion besitzt, oder äquivalent, wenn die einzigen Teilmengen von X, die sowohl offen als auch abgeschlossen sind, die leere Menge $\emptyset$ und der Raum X selbst sind.

Für Teilmengen A eines topologischen Raumes X definiert man Zusammenhang über die Teilraumtopologie. Äquivalent dazu ist, dass $A \subset X$ genau dann zusammenhängend ist, wenn für offene Teilmengen $U, V \subset X$ sodass $U \cap V \cap A = \emptyset$ und $(U \cup V) \cap A = A$ ist, immer $U \cap A = \emptyset$ oder $V \cap A = \emptyset$ gelten muss. Ein einfaches Grundresultat der Topologie ist, dass für eine stetige Funktion $f : X \to Y$ zwischen topologischen Räumen und eine zusammenhängende Teilmenge $A \subset X$ auch $f(A) \subset Y$ zusammenhängend ist. Etwas mühsamer ist der Beweis, dass die zusammenhängenden Teilmengen von $\mathbb{R}$ genau die Intervalle sind, was zusammen mit dem obigen Resultat eine allgemeine Version des Zwischenwertsatzes liefert.

Beispiel 2.9. Erinnern wir uns an Beispiel (3) von 2.1: Dort haben wir bemerkt, dass die Teilmengen $U := \{A : \det(A) > 0\}$ und $V := \{A : \det(A) < 0\}$ offen in $M_n(\mathbb{R})$ sind. Klarerweise ist $GL(n,\mathbb{R}) = U \cup V$ und $U \cap V = \emptyset$, und beide Teilmengen sind nichtleer, also liefern sie eine Disjunktion von $GL(n,\mathbb{R})$. Damit ist $GL(n,\mathbb{R})$ nicht zusammenhängend. Wir haben auch schon bemerkt, dass $U = GL^+(n,\mathbb{R})$ eine abgeschlossene Untergruppe von $GL(n,\mathbb{R})$ ist.

Analog haben wir in Beispiel (4) von 2.1 gesehen, dass für eine Matrix $A \in O(n)$ immer $\det(A) = \pm 1$ gilt. Wie oben folgt $O(n) = (U \cap O(n)) \cup (V \cap O(n))$ also ist auch $O(n)$ nicht zusammenhängend. Hier ist $U \cap O(n)$ die abgeschlossene Untergruppe $SO(n)$ von $O(n)$.

Es gibt noch einen verwandten Begriff, der intuitiv einfacher verständlich ist. Ein topologischer Raum X heißt *bogenzusammenhängend* oder *wegzusammenhängend*, wenn es für je zwei Punkte $x, y \in X$ ein Intervall $[a, b] \subset \mathbb{R}$ und eine stetige Kurve $c : [a, b] \to X$ gibt, sodass $c(a) = x$ und $c(b) = y$ gilt. Man zeigt leicht (siehe Übungen), dass ein bogenzusammenhängender Raum automatisch zusammenhängend ist. Für einen topologischen Raum X und einen Punkt $x \in X$ definiert man die *Bogenkomponente* $\mathcal{B}_x$ von x, als die Menge all jener Punkte, die man durch eine stetige Kurve c wie oben mit x verbinden kann. Man sieht leicht, dass für Punkte $x, y \in X$ entweder $\mathcal{B}_x = \mathcal{B}_y$ oder $\mathcal{B}_x \cap \mathcal{B}_y = \emptyset$ gilt, also ist X disjunkte Vereinigung von Bogenkomponenten.

Nachdem wir die nötigen Begriffe gesammelt haben, können wir nun klären, wie das für Teilmannigfaltigkeiten und Matrixgruppen aussieht.

Satz 2.9. *(1) Für eine Teilmannigfaltigkeit $M \subset \mathbb{R}^N$ und einen Punkt $x \in M$ ist die Bogenkomponente $\mathcal{B}_x$ eine Teilmenge von M, die offen und abgeschlossen ist, und damit selbst eine Teilmannigfaltigkeit von $\mathbb{R}^N$. Insbesondere ist M genau dann zusammenhängend, wenn M bogenzusammenhängend ist.*

(2) Sei $G \subset GL(n, \mathbb{R})$ eine Matrixgruppe und sei $G_0 := \mathcal{B}_{\mathbb{I}} \subset G$ die Bogenkomponente der Einheitsmatrix. Dann ist G_0 eine abgeschlossene, normale Untergruppe von G und damit selbst eine Matrixgruppe, die die gleiche Lie Algebra $\mathfrak{g}$ hat wie G.

Beweis. (1) Sei $x \in M$ ein Punkt und $y \in \mathcal{B}_x$. Dann ist $y \in \mathcal{B}_x \cap \mathcal{B}_y$, also sehen wir von oben, dass $\mathcal{B}_y = \mathcal{B}_x$ gilt. Nach Definition gibt es offene Teilmenge $U, V \subset \mathbb{R}^N$ mit $y \in U$ und einen Diffeomorphismus $\Phi : U \to V$ sodass $\Phi(U \cap M) = V \cap \mathbb{R}^k$ gilt, wobei k die Dimension von M ist. Da V offen ist und $\Phi(y) \in V$ gilt, gibt es ein $\varepsilon > 0$, sodass der Ball $B_\varepsilon(\Phi(y))$ vom Radius ε um $\Phi(y)$ ganz in V liegt. Damit ist auch $W := \Phi^{-1}(B_\varepsilon(\Phi(y)))$ eine offene Teilmenge von $\mathbb{R}^N$ und damit $W \cap M$ eine offene Umgebung von y in M. Für $z \in W \cap M$ gilt $\Phi(z) \in B_\varepsilon(\Phi(y)) \cap \mathbb{R}^k \subset V \cap \mathbb{R}^k$ und dieser Schnitt ist ein Ball in $\mathbb{R}^k$. Damit ist aber $c(t) = \Phi(y) + t(\Phi(z) - \Phi(y))$ eine stetige Kurve in $V \cap \mathbb{R}^k$, die $\Phi(y)$ mit $\Phi(z)$ verbindet, also ist $\Phi^{-1} \circ c$ eine stetige Kurve, die y mit z verbindet. Damit gilt $z \in \mathcal{B}_y$ und damit $W \subset \mathcal{B}_y = \mathcal{B}_x$ und $\mathcal{B}_x$ enthält eine offene Umgebung von y. Da $y \in \mathcal{B}_x$ beliebig war ist $\mathcal{B}_x$ offen in M.

Da der Punkt $x \in M$ beliebig war ist jede Bogenkomponente in M offen und von oben wissen wir, dass M disjunkte Vereinigung von Bogenkomponenten ist. Damit ist aber $M \setminus \mathcal{B}_x$ die Vereinigung jener Bogenkomponenten, die disjunkt zu $\mathcal{B}_x$ sind und damit ebenfalls offen. Ist M zusammenhängend, dann ist das nur für $M = \mathcal{B}_x$ möglich, also ist $M = \mathcal{B}_x$ bogenzusammenhängend. Umgekehrt wissen wir bereits, dass ein bogenzusammenhängender Raum zusammenhängend ist, also ist der Beweis der Äquivalenz in (1) vollständig. Aus der Definition ist offensichtlich, dass eine offene Teilmenge einer Teilmannigfaltigkeit selbst eine Teilmannigfaltigkeit ist.

(2) Für eine Matrix $A \in G_0$ gibt es eine stetige Kurve $c : [0,1] \to G$, die $c(0) = \mathbb{I}$ und $c(1) = A$ erfüllt. Dann ist aber auch $\nu \circ c : [0,1] \to G$ stetig und es gilt $\nu(c(0)) = \mathbb{I}^{-1} = \mathbb{I}$ und $\nu(c(1)) = A^{-1}$, also ist $A^{-1} \in G_0$. Ist $B \in G_0$ eine weitere Matrix, dann finden wir eine stetige Kurve $\tilde{c} : [0,1] \to G$, die $\tilde{c}(0) = \mathbb{I}$ und $\tilde{c}(1) = B$ erfüllt. Dann ist aber $\mu \circ (c, \tilde{c}) : [0,1] \to G$ eine stetige Kurve, die t auf $c(t) \cdot \tilde{c}(t)$ abbildet und damit $\mathbb{I} \cdot \mathbb{I} = \mathbb{I}$ mit AB verbindet. Damit sehen wir, das $G_0 \subset G$ eine Untergruppe ist und aus (1) wissen wir dass G_0 offen und abgeschlossen in G ist. Damit ist G_0 selbst eine Matrixgruppe. Die Aussage über die Lie Algebra folgt sofort aus der Definition, weil jede der in Definition 2.2 vorkommenden Kurven stetig ist und damit Werte in $\mathcal{B}_{\mathbb{I}} = G_0$ hat.

Also bleibt nur noch zu zeigen, dass die Untergruppe $G_0 \subset G$ normal ist. Dazu sei $A \in G_0$ und $c : [0,1] \to G$ eine stetige Kurve, die $\mathbb{I}$ mit A verbindet. Dann ist aber für eine beliebige Matrix $B \in G$ die Kurve $t \mapsto Bc(t)B^{-1}$ eine stetige Kurve mit Anfangspunkt$B\mathbb{I}B^{-1} = \mathbb{I}$. Damit liegt ihr Endpunkt BAB^{-1} in $\mathcal{B}_{\mathbb{I}} = G_0$ was die Normalität beweist. $\square$

Damit können wir nun genau verstehen, wie weit ein glatter Homomorphismus durch seine Ableitung bestimmt ist.

Korollar 2.9. *Sie $G \subset GL(n,\mathbb{R})$ eine Matrixgruppe, $\mathfrak{g}$ ihre Lie Algebra und $G_0 \subset G$ die Bogenkomponente von $\mathbb{I}$.*

(1) Die Untergruppe $G_0 \subset G$ wird (algebraisch) von den Matrizen der Form $\exp(X)$ mit $X \in \mathfrak{g}$ erzeugt.

(2) Sei H eine Matrixgruppe mit Lie Algebra $\mathfrak{h}$ und seien $\varphi, \psi : G \to H$ glatte Homomorphismen. Dann gilt $\varphi' = \psi' : \mathfrak{g} \to \mathfrak{h}$ genau dann, wenn $\varphi|_{G_0} = \psi|_{G_0}$ gilt.

Beweis. (1) Sei $\tilde{G} \subset G$, die Untergruppe, die von der Menge $\{\exp(X) : X \in \mathfrak{g}\}$ erzeugt wird. Für $X \in \mathfrak{g}$ ist $c(t) = \exp(tX)$ eine glatte Kurve, die $c(0) = \mathbb{I}$ und $c(1) = \exp(X)$ erfüllt. Damit ist $\exp(X) \in G_0$ und da $G_0 \subset G$ eine Untergruppe ist, ist $\tilde{G} \subset G_0$.

Aus dem Beweis von Satz 2.7 wissen wir, dass es eine offene Umgebung U von $\mathbb{I}$ in G gibt, die im Bild von exp und damit in $\tilde{G}$ enthalten ist. Für eine beliebige Matrix $A \in G$ ist dann $\lambda_A(U)$ eine offene Umgebung von A in G. Ist $A \in \tilde{G}$, dann ist $\lambda_A(U) \subset \tilde{G}$ und damit folgt, dass die Teilmenge $\tilde{G} \subset G$ offen ist. Ist $A \in G \setminus \tilde{G}$, dann muss $\lambda_A(U) \cap \tilde{G} = \emptyset$ gelten. Wäre nämlich $B \in U$ so, dass $AB \in \tilde{G}$ gilt, dann wäre wegen $B \in \tilde{G}$ auch $(AB)B^{-1} = A \in \tilde{G}$. Damit ist aber auch $G \setminus \tilde{G}$ offen, also $\tilde{G} \subset G_0$ eine nichtleere Teilmenge, die offen und abgeschlossen ist, und weil G_0 zusammenhängend ist, folgt $\tilde{G} = G_0$.

(2) Für $X \in \mathfrak{g}$ können wir $\varphi'(X)$ als $\frac{d}{dt}|_{t=0}\varphi(\exp(tX))$ berechnen. Natürlich gilt $\exp(tX) \in G_0$ für alle $t \in \mathbb{R}$ also hängt φ' nur von $\varphi|_{G_0}$ ab und die eine Implikation folgt. Umgekehrt folgt aus $\varphi' = \psi'$ nach Teil (1) von Satz 2.8 sofort $\varphi(\exp(X)) = \psi(\exp(X))$ für alle $X \in \mathfrak{g}$. Nun sieht man aber sofort, dass Teilmenge $\{A \in G : \varphi(A) = \psi(A)\}$ eine Untergruppe ist. Nach Teil (1) muss diese Untergruppe G_0 enthalten, was den Beweis vervollständigt. □

Ist eine Matrixgruppe G selbst zusammenhängend, dann ist $G = G_0$ und glatte Homomorphismen in beliebige andere Matrixgruppen sind durch ihre Ableitung eindeutig bestimmt. Aus den Resultaten dieses Abschnitts folgt auch, dass die Lie Algebra $\mathfrak{g}$ einer Matrixgruppe G nur von G_0 abhängt, über den "Rest" von G kann man durch $\mathfrak{g}$ und damit durch die Ableitung von Homomorphismen keine Information gewinnen. Daher werden wir im Weiteren meist zusammenhängende Matrixgruppen studieren.

Die Resultate dieses Abschnitts zeigen aber auch, wie man im Fall nicht zusammenhängender Matrixgruppen weiter vorgehen kann. Da $G_0 \subset G$ eine normale Untergruppen ist, erhält man eine induzierte Gruppenstruktur auf G/G_0. Da $G_0 \subset G$ offen und abgeschlossen ist, erbt dieser Quotient die diskrete Topologie, man kann ihn also einfach als Gruppe im Sinn der Algebra betrachten. Aus den Definitionen folgt leicht, dass für ein Element $A \in G$ die Menge $AG_0 = \{AB : B \in G_0\}$ genau die Bogenkomponente von A in G ist. Damit kann man aber die Menge G/G_0 mit der Menge der Bogenkomponenten von G identifizieren. Daher wird G/G_0 üblicherweise als die *Komponentengruppe* von G bezeichnet. In den meisten interessanten Beispielen ist diese Gruppe endlich mit wenigen Elementen. Für einen glatten Homomorphismus $\varphi : G \to H$ ist $\varphi(G_0) \subset H$ eine zusammenhängende Teilmenge, die $\mathbb{I} = \varphi(\mathbb{I})$ enthält, also folgt $\varphi : G_0 \subset H_0$ und natürlich ist $\varphi|_{G_0} : G_0 \to H_0$ ein glatter Homomorphismus zwischen zusammenhängenden Matrixgruppen. Außerdem erhält man einen induzierten Homomorphismus $\underline{\varphi} : G/G_0 \to H/H_0$ zwischen den Komponentengruppen, den man benutzen kann um φ zu studieren.

2.10. Mehr über Homomorphismen und ihre Ableitungen. Für einen glatten Homomorphismus $\varphi : G \to H$ zwischen Matrixgruppen ist die Ableitung $\varphi' : \mathfrak{g} \to \mathfrak{h}$ eine lineare Abbildung, also ein viel einfacheres Objekt als φ selbst. Trotzdem kann man eine ganze Menge an Information über einen glatten Homomorphismus aus seiner Ableitung ablesen. Dazu bemerken wir zunächst das der Kern $\mathrm{Ker}(\varphi') = \{X \in \mathfrak{g} : \varphi'(X) = 0\}$ der linearen Abbildung φ' natürlich ein Teilraum in $\mathfrak{g}$ ist. In Beispiel 2.8 haben wir schon bemerkt, dass für $X \in \mathrm{Ker}(\varphi')$ und $Y \in \mathfrak{g}$ immer $[X, Y] \in \mathrm{Ker}(\varphi')$ gilt, also ist insbesondere $\mathrm{Ker}(\varphi')$ eine Lie Teilalgebra von $\mathfrak{g}$. Man nennt einen Teilraum mit dieser stärkeren Eigenschaft ein *Ideal* in $\mathfrak{g}$.

Weiters benötigen wir den Begriff des *Zentrums* einer Gruppe. Dieses ist als $Z(G) := \{A \in G : \forall B \in G : AB = BA\}$ definiert, was sofort impliziert, dass $Z(G)$ eine kommutative Untergruppe von G ist. Für $A \in Z(G)$ und $B \in G$ gilt aber natürlich $BAB^{-1} = ABB^{-1} = A$, also ist $Z(G)$ eine normale Untergruppe von G. Für eine fixe Matrix $B \in G$ definiert $f_B(A) := ABA^{-1}B^{-1}$ eine glatte Funktion $f_B : G \to G$ also ist $f_B^{-1}(\{\mathbb{I}\})$ eine abgeschlossene Teilmenge von G. Damit ist auch der Durchschnitt dieser Mengen über alle $B \in G$ abgeschlossen in G. Aber aus der Definition folgt sofort, dass dieser Durchschnitt mit $Z(G)$ übereinstimmt, also ist $Z(G)$ selbst eine Matrixgruppe.

Satz 2.10. *Sei $\varphi : G \to H$ ein glatter Homomorphismus zwischen Matrixgruppen mit Ableitung $\varphi' : \mathfrak{g} \to \mathfrak{h}$.*

(1) Der Kern $\mathrm{Ker}(\varphi) = \{A \in G : \varphi(A) = \mathbb{I}\}$ *von φ ist eine abgeschlossene, normale Untergruppe von G und damit selbst eine Matrixgruppe. Die Lie Algebra von* $\mathrm{Ker}(\varphi)$ *ist* $\mathrm{Ker}(\varphi')$.

(2) Ist φ' injektiv, dann ist $\mathrm{Ker}(\varphi)$ *eine diskrete Untergruppe von G. Ist außerdem G zusammenhängend, dann ist* $\mathrm{Ker}(\varphi)$ *im Zentrum $Z(G)$ enthalten.*

(3) Ist φ' surjektiv und H zusammenhängend, dann ist φ surjektiv.

(4) Ist φ' bijektiv, dann gibt es für jede Matrix $A \in G$ eine offene Teilmenge $U \subset G$ mit $A \in U$, sodass $\varphi(U) \subset H$ offen und $\varphi|_U : U \to \varphi(U)$ ein Diffeomorphismus ist.

Beweis. (1) Aus der Algebra ist bekannt, dass der Kern eines Gruppenhomomorphismus immer eine normale Untergruppe ist. Natürlich ist $\{\mathbb{I}\}$ abgeschlossen in $M_m(\mathbb{R})$ und damit auch in H und da $\varphi : G \to H$ stetig ist, ist $\mathrm{Ker}(\varphi)$ als Urbild dieser Teilmenge abgeschlossen in G. Damit ist $\mathrm{Ker}(\varphi)$ abgeschlossen in G und somit auch in $GL(n, \mathbb{R})$, also eine Matrixgruppe.

Für $X \in \mathfrak{g}$ gilt $\varphi(\exp(tX)) = \exp(t\varphi'(X))$ nach Satz 2.8. Somit liegt $\exp(tX)$ genau dann für alle t in $\mathrm{Ker}(\varphi)$, wenn $\exp(t\varphi'(X)) = e$ für alle t gilt. Das ist aber offensichtlich equivalent zu $\varphi'(X) = 0$ also zu $X \in \mathrm{Ker}(\varphi')$. Nach Satz 2.2 sagt das genau, dass $\mathrm{Ker}(\varphi')$ die Lie Algebra der Matrixgruppe $\mathrm{Ker}(\varphi)$ ist.

(2) Nach Teil (1) ist $\mathrm{Ker}(\varphi)$ eine Matrixgruppe mit Lie Algebra $\{0\}$. Also finden wir nach Satz 2.7 eine offene Teilmenge $U \subset GL(n, \mathbb{R})$, sodass $U \cap \mathrm{Ker}(\varphi) = \{\mathbb{I}\}$ gilt. Das bedeutet aber gerade, dass $\{\mathbb{I}\}$ offen in $\mathrm{Ker}(\varphi)$ ist. Analog ist für $A \in \mathrm{Ker}(\varphi)$ die Teilmenge $\lambda_A(U) \subset GL(n, \mathbb{R})$ offen und für $B \in \lambda_A(U) \cap \mathrm{Ker}(\varphi)$ ist $A^{-1}B \in U \cap \mathrm{Ker}(\varphi)$, also $B = A$. Damit ist auch $\{A\} \subset \mathrm{Ker}(\varphi)$ offen, also $\mathrm{Ker}(\varphi)$ eine diskrete Untergruppe von G.

Für $A \in \mathrm{Ker}(\varphi)$ und $X \in \mathfrak{g}$ ist $t \mapsto \exp(tX)A\exp(-tX)$ eine glatte Kurve in G und man sieht sofort, dass die Kurve Werte in $\mathrm{Ker}(\varphi)$ hat. Da $\mathrm{Ker}(\varphi)$ diskret ist, muss die Kurve konstant sein und der Wert bei $t = 0$ ist A. Damit gilt aber $\exp(tX)A\exp(-tX) = A$ also $\exp(tX)A = A\exp(tX)$ für alle $t \in \mathbb{R}$. Damit kommutiert A aber auch mit jedem Element der von $\{\exp(X) : X \in \mathfrak{g}\}$ erzeugten Untergruppe und da G zusammenhängend ist, folgt $A \in Z(G)$.

(3) Nach Satz 2.8 gilt $\varphi(\exp(X)) = \exp(\varphi'(X))$ und weil φ' surjektiv ist, ist die Menge $\{\exp(X) : X \in \mathfrak{h}\}$ im Bild von φ enthalten. Aus der Algebra ist bekannt, dass dieses Bild eine Untergruppe von H ist, also folgt nach Korollar 2.9, dass $H_0 = H$ im Bild von φ enthalten ist.

(4) Betrachten wir zunächst $A = \mathbb{I}$. Im Beweis von Satz 2.7 haben wir eine offene Umgebungen V_G von 0 in $\mathfrak{g}$ und U_G von $\mathbb{I}$ in G konstruiert, sodass exp sich zu einem Diffeomorphismus $\exp|_{V_G} : V_G \to U_G$ einschränkt und analog finden wir $V_H \subset \mathfrak{h}$ und $U_H \subset H$. Da $\varphi' : \mathfrak{g} \to \mathfrak{h}$ als lineare Abbildung stetig ist, ist auch $(\varphi')^{-1}(V_H)$ eine offene Umgebung von 0 in $\mathfrak{g}$ und damit ist auch $V_1 := V_G \cap (\varphi')^{-1}(V_H)$ so eine Umgebung. Weil $V_1 \subset V_G$ gilt, ist $U_1 := \exp(V_1) \subset U_G$ offen in G und $\exp|_{V_1} : V_1 \to U_1$ ist ein Diffeomorphismus. Da φ' bijektiv ist, ist $V_2 := \varphi'(V_1) \subset V_H$ eine offene Umgebung von 0 und natürlich ist $\varphi'|_{V_1} : V_1 \to V_2$ ein Diffeomorphismus. Schließlich folgt wie oben, dass $U_2 := \exp(V_2)$ eine offene Umgebung von $\mathbb{I}$ in H und $\exp|_{V_2} : V_2 \to U_2$ eine Diffeomorphismus ist. Nach Satz 2.8 ist $\varphi(\exp(X)) = \exp(\varphi'(X))$ für alle $X \in \mathfrak{g}$. Für $A \in U_1$ kann man daher $\varphi(A)$ als $\exp(\varphi'((\exp)^{-1}(A)))$ schreiben. Das bedeutet aber gerade, dass man $\varphi|_{U_1} = (\exp|_{V_2}) \circ (\varphi'|_{V_1}) \circ (\exp|_{V_1})^{-1}$ schreiben kann und dass $\varphi|_{U_1} : U_1 \to U_2$ ein Diffeomorphismus ist.

Für eine beliebige Matrix $A \in G$ ist $\lambda_A(U_1)$ eine offene Umgebung von A in G und $\lambda_{\varphi(A)}(U_2)$ eine offene Umgebung von $\varphi(A)$ in H. Für $B \in U_1$ erhalten wir $\varphi(AB) = \varphi(A)\varphi(B)$, also können wir $\varphi|_{\lambda_A(U_1)}$ als $\lambda_{\varphi(A)}|_{U_2} \circ \varphi|_{U_1} \circ \lambda_{A^{-1}}$ schreiben. Damit ist aber $\varphi|_{\lambda_A(U_1)} : \lambda_A(U_1) \to \lambda_{\varphi(A)}(U_2)$ ein Diffeomorphismus. □

Bemerkung 2.10. Mit etwas mehr Theorie von Teilmannigfaltigkeiten liesse sich Teil (4) einfacher beweisen. Aus der Tatsache, dass $\varphi \circ \lambda_A = \lambda_{\varphi(A)} \circ \varphi$ gilt, schließt man leicht, dass aus der Bijektivität von φ' folgt, dass $T_A\varphi : T_AG \to T_{\varphi(A)}H$ für alle $A \in G$ bijektiv ist. Dann kann man einfach den inversen Funktionensatz für Teilmannigfaltigkeiten anwenden, den wir allerdings erst in Korollar 4.2 beweisen werden.

Beispiel 2.10. Ein fundamentales Beispiel eines glatten Homomorphismus mit bijektiver Ableitung ist die Abbildung $t \mapsto e^{it}$ als Funktion von $\mathbb{R}$ nach $\{z \in \mathbb{C} : |z| = 1\}$. Da wir noch nicht ausgiebiger über komplexe Matrixgruppen gesprochen haben, bringen wir das auf etwas andere Art in das Setting der Matrixgruppen. Einerseits betrachten wir $N(2, \mathbb{R}) := \{\left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right) : t \in \mathbb{R}\} \subset GL(2, \mathbb{R})$ wie in Beispiel (3) von 2.4. Bezeichnen wir die Matrix in der Definition mit A_t, dann rechnet man sofort nach, dass $A_sA_t = A_{s+t}$ gilt. Die Lie Algebra $\mathfrak{n}(2, \mathbb{R})$ von $N(2, \mathbb{R})$ besteht aus allen Matrizen der Form $\left(\begin{smallmatrix} 0 & t \\ 0 & 0 \end{smallmatrix}\right)$ mit $t \in \mathbb{R}$. Bezeichnen wir diese Matrix mit X_t, dann gilt $X_t^2 = 0$ und damit $\exp(X_t) = \mathbb{I} + X_t = A_t$. Also können wir $N(2, \mathbb{R})$ einfach mit der additiven Gruppe $(\mathbb{R}, +)$ und ihre Exponentialabbildung mit der Identität auf $\mathbb{R}$ identifizieren.

Die Gruppe $\{z \in \mathbb{C} : |z| = 1\}$ ist natürlich der Einheitskreis, den wir auch anders schön als Matrixgruppe beschreiben können. Dazu betrachten wir die Gruppe $H := SO(2) \subset GL(2, \mathbb{R})$ aus Abschnitt 1.3 und Beispiel (4) von 2.1. Aus der linearen Algebra ist bekannt, dass eine Matrix A genau dann orthogonal ist, wenn ihre Spaltenvektoren eine Orthonormalbasis bilden. In $\mathbb{R}^2$ kann man jeden Einheitsvektor als $\binom{\cos(t)}{\sin(t)}$ für $t \in [0, 2\pi)$ schrieben. Stellen wir den ersten Spaltenvektor von A in dieser Form dar, dann muss der zweite Spaltenvektor einer der beiden Einheitsvektoren sein, die orthogonal auf den gegebenen Vektor stehen, also kommt nur $\pm\binom{-\sin(t)}{\cos(t)}$ in Frage. Für die Determinante ergibt sich dann $\pm(\cos^2(t) + \sin^2(t))$, also sehen wir, dass H aus allen Matrizen der Form $U_t := \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}$ besteht. Diese Darstellung macht natürlich für beliebige $t \in \mathbb{R}$ Sinn, aber um alle solchen Matrizen zu erhalten, genügt es $t \in [0, 2\pi)$ zu wählen.

Wir wissen bereits, dass H eine abgeschlossene Untergruppe von $GL(2,\mathbb{R})$ ist und dass die Lie Algebra $\mathfrak{h}$ von H aus allen schiefsymmetrischen 2×2–Matrizen besteht. So eine Matrix ist von der Form $\left(\begin{smallmatrix} 0 & -t \\ t & 0 \end{smallmatrix}\right)$ und man verifiziert sofort, dass das Exponential dieser Matrix gerade U_t ist. Daraus folgt sofort, dass $U_t U_s = U_{t+s}$ gilt. Nun können wir natürlich einen Funktion $\varphi : G \to H$ durch $\varphi(A_t) := U_t$ definieren und aus unseren Beobachtungen folgt schon, dass φ ein Homomorphismus ist. Außerdem ist in unseren Identifikationen die Funktion $\varphi \circ \exp : \mathfrak{g} \to H$ durch $t \mapsto U_t$ gegeben und aus der Definition von U_t folgt sofort, dass diese Funktion glatt ist. Damit ist φ ein glatter Homomorphismus. Die Ableitung φ' bildet die Matrix X_s auf die Ableitung $\frac{d}{dt}|_{t=0} U_{ts}$ und diese Ableitung ist offensichtlich $\left(\begin{smallmatrix} 0 & -s \\ s & 0 \end{smallmatrix}\right)$ und damit ist φ' ein linearer Isomorphismus.

Wir können nun die in Satz 2.10 bewiesenen Eigenschaften alle explizit sehen. Der Kern von φ ist $\{A_t : t \in 2\pi\mathbb{Z}\}$ und natürlich ist das eine diskrete Untergruppe von $\mathbb{R}$ (die isomorph zu $\mathbb{Z}$ ist). Da G kommutativ ist, sind die anderen beiden Eigenschaften aus Teil (1) und (2) von Satz 2.10 trivialerweise erfüllt. Offensichtlich ist φ surjektiv und für ein offenes Intervall I von Länge höchstens 2π ist die Abbildung $t \mapsto \varphi(A_t)$ offensichtlich ein Diffeomorphismus. Natürlich ist aber φ kein Isomorphismus und es gibt auch gar keinen interessanten glatten Homomorphismus $H \to G$. Ist nämlich $\psi : H \to G$ ein glatter Homomorphismus, dann ist $\psi(H)$ eine Untergruppe von G, die kompakt ist, weil H kompakt und ψ stetig ist. Identifizieren wir G mit $\mathbb{R}$, dann muss nach dem Satz von Heine Borel $\psi(H) \subset G$ beschränkt sein. Aber eine Untergruppe von $(\mathbb{R},+)$ enthält mit jedem Element t auch $t + t = 2t$ und induktiv folgt, dass auch kt für alle $k \in \mathbb{Z}$ in der Untergruppe liegen muss, was für $t \neq 0$ eine unbeschränkte Menge ist. Also muss $\psi(H) = \{0\}$ gelten.

2.11. Eine allgemeine Perspektive. Zum Abschluss des Kapitels wollen wir noch eine kleine Verallgemeinerung der entwickelten Konzepte besprechen. Unsere Beschreibung von Matrixgruppen war auf den Raum $M_n(\mathbb{R})$ der $n \times n$–Matrizen, also auf den Raum der linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^n$, aufgebaut. Aus der linearen Algebra ist bekannt, dass die Wahl einer Basis für einen endlichdimensionalen $\mathbb{R}$–Vektorraum V einen linearen Isomorphismus $\mathbb{R}^n \to V$ definiert, wobei $n = \dim(V)$ ist. Damit kann man auch lineare Abbildungen von V auf sich selbst durch $n \times n$–Matrizen beschreiben. Damit scheint es plausibel, dass sich die Theorie der Matrixgruppen auf allgemeine endlichdimensionale Vektorräume übertragen kann. Dabei muss man aber natürlich darauf achten, dass die Wahl der Basis, die den Isomorphismus geliefert hat, keine Rolle spielt.

Dass ein derartiger Zugang nützlich sein könnte, ergibt sich schon aus Dingen, die wir bereits besprochen haben. Beginnen wir mit einer Matrixgruppe $G \subset GL(n,\mathbb{R})$, dann ist die Lie Algebra $\mathfrak{g}$ von G ein linearer Teilraum von $M_n(\mathbb{R})$ und damit ein endlichdimensionaler Vektorraum. Aus Satz 2.3 wissen wir, dass für $A \in G$ und $X \in \mathfrak{g}$ auch AXA^{-1} in $\mathfrak{g}$ liegt. Damit können wir für fixes $A \in G$, eine Funktion $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ durch $\mathrm{Ad}_A(X) := AXA^{-1}$ definieren. Dann ist Ad_A die Einschränkung der linearen Abbildung $\lambda_A \circ \rho^{A^{-1}} = \rho^{A^{-1}} \circ \lambda_A$ auf den Teilraum $\mathfrak{g} \subset M_n(\mathbb{R})$, also ist $\mathrm{Ad}_A : \mathfrak{g} \to \mathfrak{g}$ eine lineare Abbildung. Aus der Definition folgt auch sofort, dass $\mathrm{Ad}_A \circ \mathrm{Ad}_B = \mathrm{Ad}_{AB}$ und $\mathrm{Ad}_{\mathbb{I}} = \mathrm{id}_{\mathfrak{g}}$ gilt. Damit ist aber Ad_A ein linearer Isomorphismus und die linearen Isomorphismen $\mathfrak{g} \to \mathfrak{g}$ bilden mit der Komposition eine Gruppe $GL(\mathfrak{g})$. Damit sehen wir aber auch, dass $A \mapsto \mathrm{Ad}_A$ einen Gruppenhomomorphismus $\mathrm{Ad} : G \to GL(\mathfrak{g})$ definiert und es stellt sich die Frage, ob wir diesen als glatten Homomorphismus zwischen Matrixgruppen betrachten können.

Sei also V ein n–dimensionaler $\mathbb{R}$–Vektorraum, seien $\mathcal{B}$ und $\mathcal{C}$ Basen von V und $\varphi, \psi : \mathbb{R}^n \to V$ die entsprechenden linearen Isomorphismen. Dann ist $\psi^{-1} \circ \varphi : \mathbb{R}^n \to \mathbb{R}^n$

ein linearer Isomorphismus und damit durch Multiplikation mit einer invertierbaren Matrix T gegeben. Die Matrixdarstellung einer linearen Abbildung $f : V \to V$ zu $\mathcal{B}$ erhält man durch $\varphi^{-1} \circ f \circ \varphi : \mathbb{R}^n \to \mathbb{R}^n$. Analog erhält man für $\mathcal{C}$ die Abbildung $\psi^{-1} \circ f \circ \psi$ und man sieht sofort, dass diese durch $T(\varphi^{-1} \circ f \circ \varphi)T^{-1}$ gegeben ist. Natürlich ist f genau dann invertierbar, wenn die entsprechende Matrix invertierbar ist. Damit liefern die Basen $\mathcal{B}$ und $\mathcal{C}$ auch Isomorphismen $GL(V) \to GL(n, \mathbb{R})$, wobei der zweite Isomorphismus durch Komposition des ersten Isomorphismus mit $A \mapsto TAT^{-1}$ gegeben ist. Bezeichnet man diese Funktion mit $c_T : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$, dann ist das natürlich glatt (als Einschränkung einer linearen Abbildung), invertierbar (mit Inverser $c_{T^{-1}}$) und ein Gruppenhomomorphismus. Damit ist aber c_T sowohl ein Gruppenisomorphismus als auch ein Diffeomorphismus.

Daraus folgt aber sofort, dass man sinnvoll von *abgeschlossenen Untergruppen* von $GL(V)$ sprechen kann. Die Frage, ob das Bild einer Untergruppe $G \subset GL(V)$ unter dem von einer Basis induziertem Isomorphismus zu $GL(n, \mathbb{R})$ abgeschlossen ist, unabhängig von der Wahl der Basis ist. Man kann dann allgemeiner solche Gruppen als Matrixgruppen bezeichnen. Durch die Matrixdarstellung erhält man zugleich einen Isomorphismus $L(V, V) \cong M_n(\mathbb{R})$, der sich bei Änderung der Basis ebenfalls durch Komposition mit c_T ändert. Daraus schließt man leicht, dass man zu G einen wohldefinierten Teilraum $\mathfrak{g} \subset L(V, V)$ erhält, der abgeschlossen unter dem Kommutator $[f, g] := f \circ g - g \circ f$ und damit eine Lie Algebra ist.

Damit können wir nun das obige Beispiel vollständig behandeln. Dazu benötigen wir noch zwei Begriffe.

Definition 2.11. Sei $\mathfrak{g}$ eine Lie Algebra über einem Körper $\mathbb{K}$.

Ein *Automorphismus von* $\mathfrak{g}$ ist eine invertierbare $\mathbb{K}$–lineare Abbildung $f : \mathfrak{g} \to \mathfrak{g}$ sodass $f([X, Y]) = [f(X), f(Y)]$ für alle $X, Y \in \mathfrak{g}$ gilt. Die Menge aller Automorphismen von $\mathfrak{g}$ wird mit $\mathrm{Aut}(\mathfrak{g}) \subset GL(\mathfrak{g})$ bezeichnet.

Satz 2.11. *Für eine Matrixgruppe $G \subset GL(n, \mathbb{R})$ mit Lie Algebra $\mathfrak{g}$ gilt:*

(1) Die Menge $\mathrm{Aut}(\mathfrak{g})$ *ist eine abgeschlossene Untergruppe von* $GL(\mathfrak{g})$ *und damit eine Matrixgruppe.*

(2) Für jedes $A \in G$ ist $\mathrm{Ad}_A \in \mathrm{Aut}(\mathfrak{g})$ *und* Ad *definiert einen glatten Homomorphismus $G \to$* $\mathrm{Aut}(\mathfrak{g}) \subset GL(\mathfrak{g})$. *Ist G zusammenhängend, dann ist der Kern von* Ad *das Zentrum $Z(G)$ von G.*

(3) Die Ableitung $\mathrm{ad} : \mathfrak{g} \to L(\mathfrak{g}, \mathfrak{g})$ *von* Ad *ist durch* $\mathrm{ad}_X(Y) = [X, Y]$ *für alle $X, Y \in \mathfrak{g}$ gegeben.*

Beweis. (1) Aus der Definition folgt sofort, dass für $f, g \in \mathrm{Aut}(\mathfrak{g})$ auch $g \circ f$ und f^{-1} in $\mathrm{Aut}(\mathfrak{g})$ liegen, also ist $\mathrm{Aut}(\mathfrak{g})$ eine Untergruppe von $GL(\mathfrak{g})$. Um die Abgeschlossenheit zu beweisen können wir $\mathfrak{g}$ durch Wahl einer beliebigen Basis mit $\mathbb{R}^N$ identifizieren, wobei $N = \dim(\mathfrak{g})$ ist. Da die Lie Klammer bilinear ist, definiert für fixe Elemente $v, w \in \mathbb{R}^N$ die Funktion $A \mapsto A[v, w] - [Av, Aw]$ eine stetige Abbildung $M_N(\mathbb{R}) \to \mathbb{R}^N$. Damit ist die Menge jener Matrizen die $A[v, w] = [Av, Aw]$ für diese fixen Vektoren erfüllt, abgeschlossen. Nach Definition ist $\mathrm{Aut}(\mathfrak{g})$ der Durchschnitt über alles diese Mengen und damit ebenfalls abgeschlossen, also eine Matrixgruppe.

(2) Natürlich ist $(AXA^{-1})(AYA^{-1}) = AXYA^{-1}$ und daraus folgt sofort, dass $\mathrm{Ad}_A([X, Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$ und damit $\mathrm{Ad}_A \in \mathrm{Aut}(\mathfrak{g})$ gilt. Aus der Definition folgt sofort, dass $\mathrm{Ad}_A \circ \mathrm{Ad}_B = \mathrm{Ad}_{AB}$ und $\mathrm{Ad}_{\mathbb{I}} = \mathrm{id}_{\mathfrak{g}}$ gilt, also ist $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ ein Gruppenhomomorphismus.

Aus Proposition 1.4 wissen wir, dass $\exp(\mathrm{Ad}_A(X)) = A \exp(X) A^{-1}$ gilt. Für $A \in Z(G)$ folgt $\exp(t \,\mathrm{Ad}_A(X)) = \exp(tX)$ für alle $X \in \mathfrak{g}$ und $t \in \mathbb{R}$. Differenzieren bei

$t = 0$ liefert $\mathrm{Ad}_A(X) = X$ für alle X und damit $A \in \mathrm{Ker}(\mathrm{Ad})$. Umkehrt folgt aus $\mathrm{Ad}_A = \mathrm{id}_{\mathfrak{g}}$, dass $A \exp(X) = \exp(X)A$ für alle $X \in \mathfrak{g}$ gilt. Damit kommutiert A auch mit allen Elementen der von $\{\exp(X) : X \in \mathfrak{g}\}$ erzeugten Untergruppe und weil G zusammenhängend ist, folgt $A \in Z(G)$.

Um die Glattheit zu beweisen, erinnern wir uns daran, dass $\mathfrak{g}$ ein linearer Teilraum von $M_n(\mathbb{R})$ und die Abbildung $X \mapsto AXA^{-1}$ auf ganz $M_n(\mathbb{R})$ definiert ist. Aus den Sätzen 1.1 und 1.2 folgt sofort, dass für fixes $X \in M_n(\mathbb{R})$ die Zuordnung $A \mapsto AXA^{-1}$ eine glatte Funktion $GL(n, \mathbb{R}) \to M_n(\mathbb{R})$ definiert. Damit ist natürlich (für fixes X) auch die Einschränkung auf $G \subset GL(n, \mathbb{R})$ eine glatte Funktion $G \to M_n(\mathbb{R})$. Für $X \in \mathfrak{g}$ hat diese Einschränkung aber Werte in $\mathfrak{g}$ und ist daher glatt als Funktion $G \to \mathfrak{g}$. Wählt man eine Basis $\{Y_1, \dots, Y_N\}$ für $\mathfrak{g}$, dann bedeutet das, dass man AXA^{-1} eindeutig als $\sum_{i=1}^{N} a_i^X(A)Y_i$ schreiben kann und dass die a_i^X glatte Funktionen $G \to \mathbb{R}$ definieren. Die Matrix von Ad_A bezüglich dieser Basis hat aber dann als Eintragungen genau die Zahlen $a_i^{Y_j}(A)$ also hängen alle diese Eintragungen glatt von A ab.

(3) Definieren wir $\mathrm{ad} : \mathfrak{g} \to L(\mathfrak{g}, \mathfrak{g})$ als die Ableitung von Ad. Dann können wir ad_X als Ableitung bei $t = 0$ von $\mathrm{Ad}_{\exp(tX)}$ berechnen und klarerweise ist dann $\mathrm{ad}_X(Y)$ die Ableitung bei $t = 0$ von $\mathrm{Ad}_{\exp(tX)}(Y) = \exp(tX)Y \exp(-tX)$. Wie in (2) macht das wieder auf $M_n(\mathbb{R})$ Sinn und dort funktioniert das Argument aus dem Beweis von Satz 2.3 und liefert $\mathrm{ad}_X(Y) = XY - YX$. Für $Y \in \mathfrak{g}$ stimmt das mit $[X, Y]$ überein. □

Bemerkung 2.11. Man kann auch die Lie Algebra der Automorphismengruppe $\mathrm{Aut}(\mathfrak{g})$ von $\mathfrak{g}$ beschrieben. Durch Differenzieren der Gleichung $c(t)[X, Y] = [c(t)X, c(t)Y]$ erhält man sofort, dass $c'(0)[X, Y] = [c'(0)X, Y] + [X, c'(0)Y]$ gelten muss. Eine lineare Abbildung $\mathfrak{g} \to \mathfrak{g}$ mit dieser Eigenschaft heißt eine *Derivation von* $\mathfrak{g}$. Man kann direkt zeigen, dass für eine Derivation D von $\mathfrak{g}$, die Exponentialkurve $\exp(tD)$ Werte in $\mathrm{Aut}(\mathfrak{g})$ hat. Damit ist die Lie Algebra von $\mathrm{Aut}(\mathfrak{g})$ genau der Raum der Derivationen von $\mathfrak{g}$. Es ist leicht direkt zu verifizieren, dass der Raum der Derivationen abgeschlossen unter dem Kommutator ist. Man bemerke auch, dass die Jacobi Identität genau sagt, dass für jedes $X \in \mathfrak{g}$ die Funktion $Y \mapsto [X, Y]$ eine Derivation von $\mathfrak{g}$ ist.

Die Überlegungen in diesem Abschnitt sind auch dann relevant, wenn man nicht an allgemeinen Vektorräumen interessiert ist. Betrachten wir etwa zwei Bilinearformen $b_1, b_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ auf $\mathbb{R}^n$ und nehmen wir an, dass es eine Matrix $T \in GL(n, \mathbb{R})$ gibt, sodass $b_2(v, w) = b_1(Tv, Tw)$ gilt. Dann können wir wie oben den Isomorphismus $c_T : GL(n, \mathbb{R}) \to GL(n, \mathbb{R})$, $c_T(A) = TAT^{-1}$ betrachten. Dann ist $b_1(c_T(A)v, c_T(A)w) = b_1(TAT^{-1}v, TAT^{-1}w) = b_2(AT^{-1}v, AT^{-1}w)$ und für $A \in O(b_2)$ stimmt das mit $b_2(T^{-1}v, T^{-1}w) = b_1(v, w)$ überein. Das bedeutet aber gerade, dass $c_T(A) \in O(b_1)$ gilt, also können wir c_T als Homomorphismus $O(b_2) \to O(b_1)$ betrachten. Völlig analog folgt aber, dass $c_{T^{-1}}$ die Untergruppe $O(b_1)$ nach $O(b_2)$ abbildet, also sind c_T und $c_{T^{-1}}$ inverse Isomorphismen zwischen $O(b_1)$ und $O(b_2)$. Natürlich erweitert sich das problemlos auf den Fall, dass b_1 und b_2 auf n-dimensionalen Vektorräumen V_1 und V_2 definiert sind und T ein linearer Isomorphismus $V_1 \to V_2$ ist.

Betrachten wir insbesondere denn Fall einer positive definiten Bilinearform b auf einem n-dimensionalen Vektorraum V. Dann ist aus der linearen Algebra bekannt, dass b eine Orhthonormalbasis $\{v_1, \dots, v_n\}$. Ist T die eindeutige lineare Abbildung $\mathbb{R}^n \to V$, die $Te_i = v_i$ für alle i erfüllt, dann gilt $b(Te_i, Te_j) = \langle e_i, e_j \rangle$. Das impliziert aber sofort, dass $b(Tx, Ty) = \langle x, y \rangle$ für alle $x, y \in \mathbb{R}^n$ gilt. Also induziert $c(T)$ einen Isomorphismus $O(b) \to O(n)$ und wir können die orthogonale Gruppe $O(n)$ durch jedes innere Produkt auf einem n-dimensionalen Vektorraum realisieren.

Im Fall symmetrischer Bilinearformen charakterisieren Resultate der linearen Algebra, wann ein solcher Isomorphismus existiert. Nach dem Trägheitssatz von Sylvester gibt es für zwei symmetrische Bilinearformen b_1, b_2 genau dann einen linearen Isomorphismus T wie oben, wenn b_1 und b_2 den gleichen Rang und die gleiche Signatur haben. Das bedeutet natürlich, dass man bis auf Isomorphie nur relativ wenige orthogonale Gruppen betrachten muss. Ähnlich sind für isomorphe Lie Algebren die Automorphismengruppen isomorph und so weiter.

KAPITEL 3

Beispiele von Matrixgruppen

In diesem Kapitel werden wir eine Fülle von Beispielen und Klassen von Beispielen von Matrixgruppen und von glatten Homomorphismen zwischen solchen Gruppen besprechen, die in der Mathematik und/oder der theoretischen Physik eine wichtige Rolle spielen. Dabei werden wir auch jeweils Aspekte klären, die wir im letzten Kapitel als wichtig erkannt haben, also etwa Zusammenhang der entsprechenden Gruppen.

3.1. Komplexe Matrixgruppen. Beim Studium der grundlegenden Operationen auf Matrizen in Kapitel 1 haben wir $\mathbb{R}$ und $\mathbb{C}$ gleichzeitig behandelt, bei der formalen Definition von Matrixgruppen aber dann nur $\mathbb{R}$ betrachtet. Wir wollen nun zeigen, dass $GL(n, \mathbb{C})$ sowie abgeschlossene Untergruppen von $GL(n, \mathbb{C})$ Matrixgruppen sind. Dazu bemerken wir zunächst, dass $\mathbb{C}^n$ natürlich ein reeller Vektorraum der Dimension $2n$ ist. Bezeichnen wir die Standardbasis von $\mathbb{C}^n$ mit $\{e_1, \dots, e_n\}$ dann bildet $\{e_1, ie_1, \dots, e_n, ie_n\}$ eine $\mathbb{R}$–Basis für $\mathbb{C}^n$, die wir benutzen um $\mathbb{C}^n$ mit $\mathbb{R}^{2n}$ zu identifizieren. Die Skalarmultiplikation mit $i \in \mathbb{C}$ definiert eine $\mathbb{R}$–lineare Abbildung $J : \mathbb{C}^n \to \mathbb{C}^n$, deren Matrixdarstellung bezüglich der obigen $\mathbb{R}$–Basis durch eine Blockdiagonalmatrix mit n Blöcken der Form $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ entlang der Hauptdiagonale gegeben ist.

Wir können den Raum $M_n(\mathbb{C})$ der komplexen $n \times n$–Matrizen als $L_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$ betrachten, was natürlich ein linearer Teilraum in $L_{\mathbb{R}}(\mathbb{C}^n, \mathbb{C}^n) \cong M_{2n}(\mathbb{R})$ ist. Die entsprechende lineare Einbettung $M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ können wir sofort explizit angeben. Für eine komplexe Matrix $A = (z_{jk})$ schreiben wir $z_{jk} = a_{jk} + ib_{jk}$ für $a_{jk}, b_{jk} \in \mathbb{R}$. Dann gilt natürlich $Ae_k = \sum_j z_{jk}e_j = \sum_k (a_{jk}e_j + b_{jk}ie_j)$ und $Aie_k = iAe_k = \sum_k(-b_{jk}e_j + a_{jk}ie_j)$. Betrachtet man also A als reelle Matrix der Größe $2n \times 2n$, dann muss man einfach jede der Eintragungen z_{jk} durch einen 2×2–Block der Form $\begin{pmatrix} a_{jk} & -b_{jk} \\ b_{jk} & a_{jk} \end{pmatrix}$ ersetzen. Nachdem die Matrixmultiplikation der Komposition von linearen Abbildungen entspricht folgt sofort, dass unsere Einbettung $M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ mit der Matrizenmultiplikation verträglich ist.

Proposition 3.1. *Schränkt man die oben konstruierte lineare Injektion $M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ auf eine abgeschlossene Untergruppe $G \subset GL(n, \mathbb{C})$ ein, dann ist das Bild eine abgeschlossene Untergruppe von $GL(2n, \mathbb{R})$ und damit eine Matrixgruppe. Insbesondere trifft das auf $GL(n, \mathbb{C})$ selbst zu.*

Die Injektion ist mit der Exponentialabbildung verträglich, also kann man für die Bestimmung der Lie Algebra von G direkt in $M_n(\mathbb{C})$ arbeiten.

Beweis. Ist eine komplex lineare Abbildung $f : \mathbb{C}^n \to \mathbb{C}^n$ invertierbar über $\mathbb{R}$, dann ist $f^{-1} : \mathbb{C}^n \to \mathbb{C}^n$ automatisch $\mathbb{C}$–linear. (Weil f $\mathbb{C}$–linear ist, gilt $f(if^{-1}(z)) = iz = f(f^{-1}(iz))$ also folgt $f^{-1}(iz) = if^{-1}(z)$ aus der Injektivität von f.) In Matrizensprache bedeutet das aber genau, dass das Bild von $GL(n, \mathbb{C})$ in $M_{2n}(\mathbb{R})$ genau der Durchschnitt von $GL(2n, \mathbb{R})$ mit dem Bild von $M_n(\mathbb{C})$ in $M_{2n}(\mathbb{R})$ ist. Das letztere Bild ist ein Teilraum und damit abgeschlossen, also ist das Bild von $GL(n, \mathbb{C})$ eine abgeschlossene Untergruppe von $GL(2n, \mathbb{R})$. Damit ist aber auch jede abgeschlossene Teilmenge von $GL(n, \mathbb{C})$ abgeschlossen in $GL(2n, \mathbb{R})$ und das Resultat für allgemeines G folgt.

Wie oben bemerkt ist die Inklusion $M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ mit der Matrixmultiplikation verträglich. Da das Matrizenexponential durch eine Potenzreihe mit reellen Koeffizienten gegeben ist, folgt sofort, dass sie auch mit der Exponentialfunktion verträglich ist. □

Nachdem wir dieses Resultat beweisen haben, ist es nicht weiter nötig, komplexe Matrizen explizit als reelle Matrizen aufzufassen. Wir können einfach verwenden, dass jede abgeschlossene Untergruppe $G \subset GL(n, \mathbb{C})$ eine Matrixgruppe ist, deren Lie Algebra $\mathfrak{g} \subset \mathfrak{gl}(n, \mathbb{C})$ durch die komplexen Versionen von Definition 2.2 bzw. als $\mathfrak{g} = \{X \in \mathfrak{gl}(n, \mathbb{C}) : \forall t \in \mathbb{R} : \exp(tX) \in G\}$ beschreiben können. Wir sehen auch, dass G eine Teilmannigfaltigkeit des Raumes $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$ ist.

3.2. Unitäre und spezielle unitäre Gruppen. Als erstes Beispiel komplexer Matrixgruppen studieren wir komplexe Analoga der Gruppen $O(n)$ und $SO(n)$, die über Hermite'sche Formen definiert sind. Betrachten wir also die standard Hermite'sche Form auf $\mathbb{C}^n$, die für Vektoren $z = (z_1, \dots, z_n)$ und $w = (w_1, \dots, w_n)$ durch $\langle z, w\rangle := \sum_j z_j\overline{w_j}$. Wie aus der linearen Algebra bekannt ist, ist diese Form komplex linear in der ersten Variable und erfüllt $\langle w, z\rangle = \overline{\langle z, w\rangle}$ (und ist daher konjugiert linear in der zweiten Variable). Daraus folgt, dass $\langle z, z\rangle$ immer reell ist, und die Form ist positiv definit, also gilt $\langle z, z\rangle > 0$ für alle $z \neq 0$. Für eine Matrix $A \in M_n(\mathbb{C})$ und $z, w \in \mathbb{C}^n$ gilt dann $\langle Az, w\rangle = \langle z, A^*w\rangle$, wobei man A^* durch Transponieren und komplexes Konjugieren von A erhält.

Definition 3.2. Für $n \geq 1$ sei $U(n) := \{A \in M_n(\mathbb{C}) : \forall z, w \in \mathbb{C}^n : \langle Az, Aw\rangle = \langle z, w\rangle\}$ und $SU(n) := \{A \in U(n) : \det(A) = 1\}$.

Eine äquivalente Charakterisierung ist, dass $A \in U(n)$ genau dann gilt, wenn $A^*A = \mathbb{I}$, also $A^* = A^{-1}$ gilt. Insbesondere erhalten wir $U(1) = \{z \in \mathbb{C} : |z| = 1\}$, den Einheitskreis in $\mathbb{C}$.

Proposition 3.2. *Für jedes $n \geq 1$ ist $U(n)$ eine kompakte Matrixgruppe mit reeller Dimension n^2, deren Lie Algebra durch $\mathfrak{u}(n) := \{X \in M_n(\mathbb{C}) : X^* = -X\}$ gegeben ist. Für $A \in U(n)$ gilt $|\det(A)| = 1$, also definiert* det *einen glatten Homomorphismus $U(n) \to U(1)$ mit Kern $SU(n)$. Damit ist $SU(n)$ für $n > 1$ ein abgeschlossener Normalteiler von $U(n)$ und damit selbst eine kompakte Matrixgruppe der Dimension $n^2 - 1$. Ihre Lie Algebra ist $\mathfrak{su}(n) := \{X \in M_n(\mathbb{C}) : X^* = -X, \operatorname{tr}(X) = 0\}$.*

Beweis. Der Beweis ist ganz ähnlich wie für die orthogonalen Gruppen. Offensichtlich ist $U(n)$ eine Untergruppe von $GL(n, \mathbb{C})$ und klarerweise ist $A^*A = \mathbb{I}$ eine abgeschlossene Bedingung. Aus der linearen Algebra ist bekannt, dass eine Matrix $A = (a_{ij}) \in M_n(\mathbb{C})$ genau dann in $U(n)$ liegt, wenn ihre Spaltenvektoren eine Orthonormalbasis von $\mathbb{C}^n$ bilden. Damit folgt aber $|a_{ij}| \leq 1$ für alle i, j, also ist $U(n)$ eine beschränkte, abgeschlossene Teilmenge von $M_n(\mathbb{C})$ und damit kompakt.

Zur Bestimmung der Lie Algebra betrachten wir eine Kurve $c : I \to GL(n, \mathbb{C})$ mit $c(0) = \mathbb{I}$, die Werte in $U(n)$ hat. Differenziert man $\langle c(t)z, c(t)w\rangle = \langle z, w\rangle$, dann erhält man sofort $0 = \langle c'(0)z, w\rangle + \langle z, c'(0)w\rangle$ und damit $c'(0) + c'(0)^* = 0$. Umgekehrt zeigt eine direkte Rechnung wie im Beweis von Proposition 2.4, dass aus $X^* = -X$ schon $\exp(tX) \in U(n)$ für alle $t \in \mathbb{R}$ folgt. Zur Bestimmung der Dimension von $\mathfrak{u}(n)$ beobachten wir, dass für $X = (x_{ij}) \in \mathfrak{u}(n)$ die Eintragungen x_{ii} rein imaginär sein müssen, während man oberhalb der Hauptdiagonale beliebige komplexe Eintragungen vorgeben kann, was dann X eindeutig festlegt. Daraus folgt $\dim(\mathfrak{u}(n)) = n^2$ sofort.

Da $\det(A) = \det(A^t)$ gilt, folgt sofort $\det(A^*) = \overline{\det(A)}$. Für $A \in U(n)$ liefert dann aber $A^*A = \mathbb{I}$ direkt $|\det(A)|^2 = 1$. Somit hat die Einschränkung der glatten

Funktion $\det : M_n(\mathbb{C}) \to \mathbb{C}$ auf $U(n)$ Werte in $U(1)$, also ist $\det : U(n) \to U(1)$ eine glatte Funktion und natürlich ist det ein Homomorphismus. Der Kern eines stetigen Homomorphismus ist ein abgeschlossener Normalteiler und als abgeschlossene Teilmenge des kompakten Raumes $U(n)$ ist $SU(n)$ ebenfalls kompakt. Die Beschreibung der Lie Algebra folgt wie in Beispiel (1) von 2.4 und die Aussage über die Dimension ist offensichtlich. □

Neben $U(1)$ hat auch die Gruppe $SU(2)$ eine einfache topologische Beschreibung. Für eine Matrix $A \in SU(2)$ müssen die Spaltenvektoren eine Orthonormalbasis von $\mathbb{C}^2$ bilden. Damit bildet die erste Spalte $\binom{z}{w}$ von A einen Einheitsvektor in $\mathbb{C}^2 \cong \mathbb{R}^4$, also ein Element der Sphäre S^3. Da der zweite Spaltenvektor normal auf $\binom{z}{w}$ steht, muss er die Form $\binom{-\lambda\bar{w}}{\lambda\bar{z}}$ für $\lambda \in \mathbb{C}$ haben und weil auch er ein Einheitsvektor sein muss, folgt $|\lambda| = 1$. Damit ist aber $\det(A) = \lambda(|z|^2 + |w^2|) = \lambda = 1$, also $A = \left(\begin{smallmatrix} z & -\bar{w} \\ w & \bar{z} \end{smallmatrix}\right)$. Somit ist A durch den ersten Spaltenvektor eindeutig bestimmt und wir sehen, dass wir $SU(2)$ mit der Sphäre S^3 identifizieren können.

Bemerkung 3.2. (1) Für Hermite'sche Formen gibt es analog zu reellen symmetrischen Bilinearformen eine Klassifikation durch Rang und Signatur. Insbesondere sind nicht degenerierte Hermite'sche Formen auf $\mathbb{C}^n$ durch ihre Signatur (p, q) mit $p + q = n$ eindeutig bis auf Isomorphie bestimmt. Insbesondere definiert $(z, w) \mapsto \sum_{j=1}^{p} z_j\bar{w}_j - \sum_{j=p+1}^{n} z_j\bar{w}_j$ eine Hermite'sche Form mit Signatur (p, q) und man kann der unitäre und spezielle unitäre Gruppe betrachten, die üblicherweise mit $U(p, q)$ und $SU(p, q)$ bezeichnet werden. Diese sind Matrixgruppen der gleichen Dimension wie $U(n)$ bzw. $SU(n)$, aber nicht kompakt. Auch die Beschreibungen der Lie Algebren sind analog (siehe Übungen).

(2) Es gibt auch eine zweite Version von komplexen orthogonalen Gruppen. Auf jedem endlichdimensionalen komplexen Vektorraum gibt es eine bis auf Isomorphie eindeutig bestimmte, nicht-degenerierte komplexe Bilinearform b. (Weil für so eine Form $b(iz, iz) = i^2 b(z, z) = -b(z, z)$ gilt, gibt es für komplexe Bilinearformen keine Signatur.) Insbesondere kann man auf $\mathbb{C}^n$ die Standardform $(z, w) \mapsto \sum_j z_j w_j$ betrachten und die Matrizen $A \in GL(n, \mathbb{C})$ die diese Form bewahren, bilden eine abgeschlossene Untergruppe $O(n, \mathbb{C}) \subset GL(n, \mathbb{C})$ deren Lie Algebra aus den schiefsymmetrischen, komplexen $n \times n$–Matrizen besteht. Für $A \in O(n, \mathbb{C})$ gilt $\det(A)^2 = 1$, also $\det(A) = \pm 1$ und man definiert $SO(n, \mathbb{C}) := \{A \in O(n, \mathbb{C}) : \det(A) = 1\}$. Auch diese Gruppen sind nicht kompakt, sie sind vor allem für die Strukturtheorie wichtig, weil sie Beispiele für komplexe einfache Lie Gruppen liefern.

3.3. Zusammenhang klassischer Gruppen. Nachdem wir die orthogonalen und unitären Gruppen definiert haben, können wir sie benutzen um die Zusammenhangseigenschaften einiger grundlegender Beispiele von Matrixgruppen zu studieren. Wir beginnen mit einfachen Beispielen, die für sich genommen nicht so wichtig sind, aber im folgenden technisch nützlich sein werden.

Betrachten wir zunächst die Gruppen $B(n, \mathbb{K})$ der invertierbaren oberen Dreiecksmatrizen aus Beispiel 2.4 (3). Für $\mathbb{K} = \mathbb{R}$ ist diese Gruppe sicher nicht zusammenhängend, weil für $A \in B(n, \mathbb{R})$ jede der Hauptdiagonaleintragungen a_{ii} in $\mathbb{R} \setminus \{0\}$ liegt und damit positiv oder negativ sein kann. Betrachten wir $\mathbb{Z}_2$ als die multiplikative Gruppe $\{1, -1\}$ und für $a \in \mathbb{R} \setminus \{0\}$ sei $\operatorname{sgn}(a) \in \mathbb{Z}_2$ das Signum von a also 1 für $a > 0$ und -1 für $a < 0$, sodass $\operatorname{sgn}(ab) = \operatorname{sgn}(a)\operatorname{sgn}(b)$ gilt. Da die Hauptdiagonaleintragungen eines Produkts von oberen Dreiecksmatrizen die Produkte der entsprechenden Eintragungen der beiden Faktoren sind, sehen wir, dass $A = (a_{ij}) \mapsto (\operatorname{sgn}(a_{11}), \dots, \operatorname{sgn}(a_{nn}))$ einen Homomorphismus $B(n, \mathbb{R}) \to (\mathbb{Z}_2)^n$ definiert, der offensichtlich surjektiv ist.

Der Kern dieses Homomorphismus besteht aus allen oberen Dreiecksmatrizen $A = (a_{ij})$, die $a_{ii} > 0$ für alle $i = 1, \dots, n$ erfüllen. Für so eine Matrix A und $t \in [0, 1]$ betrachten wir $(1-t)\mathbb{I} + tA$. Das ist offensichtlich eine obere Dreiecksmatrix mit Hauptdiagonaleintragungen $(1-t) + ta_{ii} > 0$, also definiert $c(t) := (1-t)\mathbb{I} + tA$ eine stetige Kurve $c : [0, 1] \to B(n, \mathbb{R})$, die $c(0) = \mathbb{I}$ und $c(1) = A$ erfüllt. Damit sehen wir aber, dass die Zusammenhangskomponente $B_0(n, \mathbb{R})$ genau aus den Matrizen in $B(n, \mathbb{R})$ besteht, bei denen alle Hauptdiagonaleintragungen positiv sind. Das gleiche Argument zeigt, dass die Gruppe $N(n, \mathbb{R})$ der oberen Dreiecksmatrizen für die alle Hauptdiagonaleintragungen gleich 1 sind, zusammenhängend ist.

Für $\mathbb{K} = \mathbb{C}$ ist die Situation einfacher, weil $\mathbb{C} \setminus \{0\}$ zusammenhängend ist. Man kann nämlich eine komplexe Zahl $z \neq 0$ eindeutig als $z = |z|e^{it_0}$ schreiben wobei $|z| > 0$ und $t \in [0, 2\pi)$ gilt. Für $t \in [0, 1]$ ist nun $(1-t) + t|z| > 0$, also definiert $c(t) := ((1-t) + t|z|)e^{itt_0}$ eine stetige Kurve in $\mathbb{C} \setminus \{0\}$, die $c(0) = 1$ und $c(1) = z$ erfüllt. Gilt $z \in U(1)$, also $|z| = 1$, dann ist $c(t) = e^{itt_0}$, also hat die Kurve Werte in $U(1)$ und wir sehen, dass auch $U(1)$ zusammenhängend ist. Für eine Matrix $A = (a_{ij}) \in B(n, \mathbb{C})$ können wir nun zunächst die Elemente außerhalb der Hauptdiagonale fix lassen, und die Hauptdiagonalelemente mit stetigen Kurven wie oben bewegen. Das zeigt, dass wir eine Kurve finden, die A mit einem Element von $N(n, \mathbb{C})$ verbindet. Für $B \in N(n, \mathbb{C})$ ist aber wieder $c(t) := (1-t)\mathbb{I} + tB$ eine Kurve in $N(n, \mathbb{C})$, die B mit $\mathbb{I}$ verbindet. Somit sind $B(n, \mathbb{C})$ und $N(n, \mathbb{C})$ zusammenhängend.

Als nächsten Schritt können wir das Gram–Schmidt Orthonormalisierungsverfahren in Matrizenschreibweise interpretieren um Informationen über die Gruppen $GL(n, \mathbb{K})$ und $SL(n, \mathbb{K})$ zu erhalten.

Lemma 3.3. *(1) Für eine Matrix $A \in GL(n, \mathbb{R})$ gibt es Matrizen $C \in O(n)$ und $N \in B_0(n, \mathbb{R})$, sodass $A = CN$ gilt. Für $A \in GL^+(n, \mathbb{R})$ gilt $C \in SO(n)$.*

(2) Für eine Matrix $A \in GL(n, \mathbb{C})$ gibt es Matrizen $C \in U(n)$ und $N \in B(n, \mathbb{C})$ mit positiven, reellen Eintragungen auf der Hauptdiagonale, sodass $A = CN$ gilt. Für $A \in SL(n, \mathbb{C})$ gilt $C \in SU(n)$.

Beweis. Der Beweis funktioniert für $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$ gleich, nur die Interpretation des Ergebnisses ist verschieden. Für $A \in GL(n, \mathbb{K})$ bilden die Spaltenvektoren $a_1, \dots, a_n$ von A eine Basis von $\mathbb{K}^n$ und wir können darauf das Gram–Schmidt Verfahren anwenden um eine Orthonormalbasis $c_1, \dots, c_n$ von $\mathbb{K}^n$ zu erhalten. Nach Definition ist $c_1 = \frac{1}{\|a_1\|}a_1$, also ein positives Vielfaches von a_1. Dann setzt man $\tilde{c}_2 := a_2 - \langle a_2, c_1 \rangle c_1$ und erhält c_2 durch Normieren dieses Vektors. Da c_1 ein Vielfaches von a_1 ist, ist c_2 eine Linearkombination von a_1 und a_2, wobei der Koeffizient von a_2 gerade $\frac{1}{\|\tilde{c}_2\|}$ und damit reell und positiv ist. Induktiv folgt für jeden Index k, dass c_k eine Linearkombination von a_1, …, a_k ist, wobei der Koeffizient von a_k reell und positiv ist.

Schreiben wir das explizit als $c_k = \sum_{j=1}^k m_{jk}a_j$ und setzen $m_{jk} = 0$ für $j > k$, dann ist $M = (m_{jk})$ eine obere Dreiecksmatrix mit positiven reellen Eintragungen auf der Hauptdiagonale. In Komponenten sagt das $c_{ik} = \sum_j a_{ij}m_{jk}$, also $C = AM$, wobei C die Matrix ist, die $c_1, \dots, c_n$ als Spaltenvektoren hat. Nun ist $N := M^{-1}$ ebenfalls eine obere Dreiecksmatrix und $n_{jj} = 1/m_{jj}$ also hat auch N positive reelle Eintragungen auf der Hauptdiagonale und natürlich gilt $CN = AMN = A$. Da die c_j eine Orthonormalbasis bilden, gilt $C \in O(n)$ für $\mathbb{K} = \mathbb{R}$ und $C \in U(n)$ für $\mathbb{K} = \mathbb{C}$. Schließlich ist $\det(N)$ reell und positiv, $|\det(C)| = 1$ und $\det(A) = \det(C)\det(N)$. Damit ist für $\mathbb{K} = \mathbb{R}$

$\det(A) > 0$ nur für $\det(C) = 1$ möglich, währen für $\mathbb{K} = \mathbb{C}$ aus $\det(A) = 1$ schon $\det(C) = \det(N) = 1$ folgt. □

Bemerkung 3.3. In dem Beweis haben wir eine bijektive Abbildungen $GL(n, \mathbb{R}) \to O(n) \times B_0(n, \mathbb{R})$ analog eine Bijektion von $GL(n, \mathbb{C})$ auf das Produkt von $U(n)$ mit einer abgeschlossenen Untergruppe von $B(n, \mathbb{C})$ konstruiert. Diese Abbildungen sind keine Gruppenhomomorphismen, aber man kann zeigen, dass sie Diffeomorphismen sind, also sowohl die Funktionen als auch ihre Inversen glatt sein. Damit kann man die Struktur von $GL(n, \mathbb{K})$ als Mannigfaltigkeit besser verstehen. Offensichtlich ist $B_0(n, \mathbb{R})$ als Mannigfaltigkeit sehr einfach, nämlich (isomorph zu) $(0, \infty)^n \times \mathbb{R}^{n(n-1)/2}$, wobei die ersten n Komponenten den positiven Eintragungen auf der Hauptdiagonale und die restlichen Komponenten den beliebigen Eintragungen über der Hauptdiagonale entsprechen. Dieser Raum hat sehr einfache Topologie, man kann ihn auf einen Punkt zusammenziehen. Damit ist die "interessante" Topologie von $GL(n, \mathbb{R})$ ganz in $O(n)$ enthalten.

Ganz analog sieht die Untergruppe von $B(n, \mathbb{C})$ wie $(0, \infty)^n \times \mathbb{C}^{n(n-1)/2}$ aus, also ist wieder die "interessante" Topologie von $GL(n, \mathbb{C})$ in $U(n)$ enthalten. Diese Zerlegungen haben ein Analogon für die Klasse der komplexen einfachen Lie Gruppen, die sogenannte Iwasawa Zerlegung.

Satz 3.3. *Für $n \geq 1$ sind die Gruppen $GL^+(n, \mathbb{R})$, $SL(n, \mathbb{R})$, $SO(n)$, $GL(n, \mathbb{C})$, $U(n)$ und $SU(n)$ zusammenhängend, während die Gruppen $GL(n, \mathbb{R})$ und $O(n)$ genau zwei Zusammenhangskomponenten besitzen.*

Beweis. Nach Lemma 3.3 finden wir für eine Matrix $A \in GL(n, \mathbb{R})$ Matrizen $C \in O(n)$ und $N \in B_0(n, \mathbb{R})$, sodass $A = CN$ gilt. Wir haben auch schon gesehen, dass $c(t) := (1-t)\mathbb{I} + tN$ eine stetige Kurve in $B_0(n, \mathbb{R})$ definiert, die $\mathbb{I}$ mit N verbindet. Damit ist aber $t \mapsto C \cdot c(t)$ eine stetige Kurve in $GL(n, \mathbb{R})$, die $C \cdot c(0) = C$ mit $C \cdot c(1) = CN = A$ verbindet. Damit folgt, dass $GL(n, \mathbb{R})$ höchstens so viele Zusammenhangskomponenten wie $O(n)$ besitzt. Für $A \in GL^+(n, \mathbb{R})$ ist $C \in SO(n)$, also liegt $c(t)$ ganz in $GL^+(n, \mathbb{R})$ und verbindet A mit einem Element von $SO(n)$. Damit besitzt $GL^+(n, \mathbb{R})$ höchstens so viele Zusammenhangskomponenten wir $SO(n)$.

Betrachten wir schließlich $A \in SL(n, \mathbb{R})$, dann gilt wieder $C \in SO(n)$ und außerdem $\det(N) = 1$. Damit ist aber $t \mapsto \det(c(t))$ eine stetige Funktion $[0, 1] \to (0, \infty)$ die 0 und 1 auf 1 abbildet. Damit ist auch $\frac{1}{\det(c(t))}$ eine stetige Funktion, und wir definieren $\tilde{c}(t)$ indem wir die erste Spalte von $c(t)$ mit $\frac{1}{\det(c(t))}$ multiplizieren. Dann folgt nach Konstruktion $\det(\tilde{c}(t)) = 1$ für alle t, sowie $\tilde{c}(0) = \mathbb{I}$ und $\tilde{c}(1) = N$. Damit hat aber die Kurve $t \mapsto C \cdot \tilde{c}(t)$ Werte in $SL(n, \mathbb{R})$ und verbindet A mit C. Damit hat auch $SL(n, \mathbb{R})$ höchstens so viele Zusammenhangskomponenten wie $SO(n)$.

Wir wissen bereits, dass $SO(n)$ und $\{A \in O(n) : \det(A) = -1\}$ disjunkte offene Teilmengen von $O(n)$ sind, womit $O(n)$ mindestens zwei Zusammenhangskomponenten besitzt. Ist $A \in O(n)$ mit $\det(A) = -1$, dann definiert λ_A einen Homöomorphismus von $SO(n)$ auf $\{A \in O(n) : \det(A) = -1\}$. Somit können wir aber alle behaupteten Aussagen für $\mathbb{K} = \mathbb{R}$ beweisen, in dem wir zeigen, dass $SO(n)$ für $n \geq 1$ zusammenhängend ist.

Nach Definition ist $SO(1) = \{1\}$ und damit zusammenhängend. Aus Beispiel 2.10 wissen wir, dass $SO(2)$ aus den Matrizen der Form $\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$ mit $t \in [0, 2\pi)$ besteht und offensichtlich kann man jede dieser Matrizen mit einer stetigen Kurve mit $\mathbb{I}$ verbinden. Damit können wir die Behauptung über $SO(n)$ nun induktiv beweisen. Sei also $A \in SO(n)$ für $n > 2$ und seien $a_1, \dots, a_n$ die Spaltenvektoren von A, die nach Voraussetzung eine positiv orientierte Orthonormalbasis von A bilden. Wir behaupten zunächst, dass es eine stetige Kurve $c(t) = (a_1(t), \dots, a_n(t))$ in $SO(n)$ gibt, die bei A beginnt und bei einer Matrix endet, deren letzter Spaltenvektor der Einheitsvektor e_n ist.

Ist $a_n = e_n$, dann ist nichts zu zeigen und ist $a = -e_n$ dann definieren wir $c(t) := (a_1, \dots, a_{n-2}, \cos t a_{n-1} + \sin t a_n, -\sin t a_{n-1} + \cos t a_n)$. Das ist offensichtlich eine stetige Kurve und man verifiziert sofort, dass $c(t) \in SO(n)$ für alle $t \in \mathbb{R}$ gilt. Außerdem ist $c(0) = A$, während der letzte Spaltenvektor von $c(\pi)$ gerade $-a_n = e_n$ ist. Ist $a_n \neq \pm e_n$, dann spannen diese beiden Vektoren eine Ebene auf und wir benutzen eine Rotation in dieser Ebene. Dazu ergänzen wir a_n durch y_n zu einer Orthonormalbasis dieser Ebene, setzen also $\tilde{y}_n := e_n - \langle e_n, a_n \rangle a_n$ und $y_n := \frac{1}{\|\tilde{y}_n\|}\tilde{y}_n$. Da e_n ein Einheitsvektor in dieser Ebene ist, gibt es eine Eindeutige Zahl $t_0 \in [0, 2\pi)$, sodass $e_n = \cos t_0 a_n + \sin t_0 y_n$ gilt. Für $i < n$ ist $\langle a_i, a_n \rangle = 0$, also steht $b_i := a_i - \langle a_i, y_n \rangle y_n$ normal auf a_n und y_n und damit auf die von diesen beiden Vektoren erzeugte Ebene und $a_i = \langle a_i, y_n \rangle y_n + b_i$. Für $j < n$ erhalten wir damit sofort

$$\langle a_i, a_j \rangle = \langle a_i, y_n \rangle \langle a_j, y_n \rangle + \langle b_i, b_j \rangle = \delta_{ij}. \tag{3.1}$$

Nun setzen wir $a_n(t) := \cos t a_n + \sin t y_n$ und für $i < n$ definieren wir

$$a_i(t) := \langle a_i, y_n \rangle (-\sin t a_n + \cos t y_n) + b_i.$$

Dann gilt $b_i \perp a_n(t)$ und damit $a_i(t) \perp a_n(t)$ für alle t. Andererseits gilt offensichtlich $\|a_n(t)\| = 1$ für alle t und aus aus (3.1) folgt sofort $\langle a_i(t), a_j(t) \rangle = \delta_{ij}$ für $i, j < n$ und alle t. Damit ist $c(t) = (a_1(t), \dots, a_n(t))$ eine stetige Kurve in $O(n)$ mit $c(0) = A \in SO(n)$, also bleibt c in $SO(n)$. Andererseits ist nach Konstruktion $a_n(t_0) = e_n$, also hat die Matrix $c(t_0)$ als letzten Spaltenvektor e_n. Die andere Spaltenvektoren von $c(t_0)$ stehen normal auf e_n, also hat $c(t_0)$ die Form $\begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}$ für eine $(n-1) \times (n-1)$–Matrix B, die nach Konstruktion orthogonal ist und Determinante 1 hat, also in $SO(n-1)$ liegt. Nach Induktionsvoraussetzung ist $SO(n-1)$ zusammenhängend, also nach Satz 2.9 bogenzusammenhängend. Damit finden wir eine stetige Kurve $B(t)$, die B mit der Einheitsmatrix verbindet und indem wir erst durch c und dann durch die Kurve $\begin{pmatrix} B(t) & 0 \\ 0 & 1 \end{pmatrix}$ laufen, können wir A mit der Einheitsmatrix verbinden.

Im komplexen Fall geht man analog vor, die Situation ist aber etwas einfacher: Man kann wieder jede Matrix $A \in GL(n, \mathbb{C})$ mit einer stetigen Kurve mit einer Matrix $C \in U(n)$ verbinden. Für $A \in SL(n, \mathbb{C})$ gilt $C \in SU(n)$ und wie zuvor kann man die Kurve so modifizieren, dass sie Werte in $SL(n, \mathbb{C})$ hat. Für eine Matrix $C \in U(n)$ ist $\det(C) \in U(1)$, also $\det(C) = e^{it_0}$ für ein $t_0 \in [0, 2\pi)$. Ist nun $C(t)$ die Matrix, die man erhält indem man die erste Spalte von C mit e^{-itt_0} multipliziert, dann erhält man eine glatte Kurve $t \mapsto C(t)$ in $U(n)$, die $C(0) = C$ und $C(1) \in SU(n)$ erfüllt. Damit können wir alle behaupteten Aussagen über die komplexen Gruppen beweisen, indem wir zeigen, dass $SU(n)$ für alle $n \geq 1$ zusammenhängend ist.

Nun gilt wieder $SU(1) = \{1\}$ und wir haben in Abschnitt 3.2 bereits $SU(2)$ mit der Einheitssphäre S^3 in $\mathbb{C}^2 \cong \mathbb{R}^4$ identifiziert haben. Man sieht aber ganz leicht (siehe Übungen), dass S^n für $n \geq 1$ bogenzusammenhängend ist, also ist $SU(2)$ zusammenhängend. Nun behaupten wir analog wie vorher, dass es zu einer gegebene Matrix $A \in SU(n)$ eine stetige Kurve in $SU(n)$ gibt, die A mit einer Matrix verbindet, deren letzter Spaltenvektor e_n ist. Ist a_n ein komplexes Vielfaches von e_n, dann gilt $a_n = e^{it_0} e_n$ für $t_0 \in [0, 2\pi)$. In diesem Fall definieren wir $c(t) := (a_1, \dots, a_{n-2}, e^{itt_0} a_{n-1}, e^{-itt_0} a_n)$ und erhalten eine Kurve $c : [0, 1] \to SU(n)$, die $c(0) = A$ erfüllt, sodass $c(1)$ als letzten Spaltenvektor $e^{-it_0} a_n = e_n$ hat.

Ist a_n kein komplexes Vielfaches von e_n, dann spannen die beiden Vektoren eine komplexe Ebene auf und wir konstruieren einen komplexen Normalvektor y_n zu a_n in dieser Ebene wie zuvor. Natürlich gibt es eine spezielle unitäre Abbildung dieser Ebene, die a_n auf e_n abbildet und von oben wissen wird, dass wir eine stetige Kurve c finden, die die Identität mit dieser Abbildung verbindet. Wie oben können wir jeden der anderen Spaltenvektoren a_j als $\lambda_j y_n + b_j$ schreiben, wobei $\lambda_j \in \mathbb{C}$ gilt und der Vektor b_j orthogonal auf die Ebene steht. Nun definiert man $a_n(t) := c(t)a_n$ und $a_j(t) := \lambda_j c(t)y_n + b_j$ und verifiziert leicht, dass das die gewünschten Eigenschaften hat. Damit erhält man wieder eine Blockform mit einer Matrix aus $SU(n-1)$ und der Zusammenhang von $SU(n)$ folgt durch Induktion. □

3.4. Die Struktur von $SO(3)$. Die speziellen unitären Gruppen führen überraschenderweise auch zu einem besseren Verständnis der speziellen orthogonalen Gruppen in niedrigen Dimensionen. Das zeigt auch, wie man die bisher entwickelten Techniken effizient anwenden kann.

Betrachten wir die Lie Algebra $\mathfrak{su}(2) = \{X \in M_2(\mathbb{C}) : X^* = -X\}$. Eine Matrix $X \in \mathfrak{su}(2)$ kann eindeutig in der Form $\begin{pmatrix} ia & -\bar{z} \\ z & -ia \end{pmatrix}$ für $a \in \mathbb{R}$ und $z \in \mathbb{C}$ geschrieben werden. Betrachten wir die Abbildung $\mathfrak{su}(2) \times \mathfrak{su}(2) \to \mathbb{C}$, die durch $(X, Y) \mapsto -\frac{1}{2}\operatorname{tr}(XY) = \frac{1}{2}\operatorname{tr}(XY^*)$ gegeben und damit offensichtlich bilinear über $\mathbb{R}$ ist. Stellen wir unsere Matrizen wie oben durch Paare (a, z) und (b, w) dar, dann sehen wir sofort, dass wir $ab + \operatorname{Re}(z\bar{w}) \in \mathbb{R}$ und für $Y = X$ die Zahl $a^2 + |z|^2$ erhalten, die für $X \neq 0$ positiv ist. Also haben wir ein positiv definites inneres Produkt auf dem 3–dimensionalen reellen Vektorraum $\mathfrak{su}(2)$ definiert, das wir mit $\langle\ ,\ \rangle$ bezeichnen. Wir sehen auch, dass die Matrizen $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$ eine Orthonormalbasis für diesen Raum bilden.

Damit können wir nun den glatten Homomorphismus $\operatorname{Ad} : SU(2) \to GL(\mathfrak{su}(2))$ aus Satz 2.11 betrachten, der durch $\operatorname{Ad}_A(X) = AXA^{-1}$ für $A \in SU(2)$ und $X \in \mathfrak{su}(2)$ gegeben ist. Offensichtlich gilt

$$\langle \operatorname{Ad}_A(X), \operatorname{Ad}_A(Y)\rangle = -\tfrac{1}{2}\operatorname{tr}(AXA^{-1}AYA^{-1}) = -\tfrac{1}{2}\operatorname{tr}(XY) = \langle X, Y\rangle,$$

also gilt $\operatorname{Ad}_A \in O(\mathfrak{su}(2))$ für jedes $A \in SU(2)$. Aus Satz 3.3 wissen wir, dass $SU(2)$ zusammenhängend ist, also ist auch $\operatorname{Ad}(SU(2))$ zusammenhängend. Aus Abschnitt 2.11 wissen wir, dass $O(\mathfrak{su}(2)) \cong O(3)$ ist, was nach Satz 3.3 genau zwei Zusammenhangskomponenten besitzt. Da $\mathbb{I} = \operatorname{Ad}_{\mathbb{I}}$ in $\operatorname{Ad}(SU(2))$ liegt, ist $\operatorname{Ad}(SU(2))$ in der Zusammenhangskomponente von $\mathbb{I}$ enthalten, die isomorph zu $SO(3)$ ist. Damit können wir Ad als glatten Homomorphismus $\operatorname{Ad} : SU(2) \to SO(3)$ betrachten.

Satz 3.4. *Der Homomorphismus* $\operatorname{Ad} : SU(2) \to SO(3)$ *ist surjektiv und* $\operatorname{Ker}(\operatorname{Ad}) = \{\mathbb{I}, -\mathbb{I}\}$*. Damit ist $SO(3)$ als Gruppe isomorph zu $SU(2)/\mathbb{Z}_2$. Als topologischer Raum (und als Mannigfaltigkeit) erhält man $SO(3)$ indem man in der Sphäre S^3 jeden Punkt x mit dem antipodalen Punkt $-x$ identifiziert. Also realisiert $SO(3)$ den dreidimensionalen projektiven Raum $\mathbb{R}P^3$.*

Beweis. Aus Satz 2.11 wissen wir, dass die Ableitung $\operatorname{ad} : \mathfrak{su}(2) \to \mathfrak{gl}(\mathfrak{su}(2))$ von Ad durch $\operatorname{ad}_X(Y) = [X, Y]$ gegeben ist und wir wissen, dass ad Werte in $\mathfrak{so}(\mathfrak{su}(2))$ hat. Nun hat aber sowohl $\mathfrak{su}(2)$ als auch $\mathfrak{so}(3)$ Dimension 3, also sind für ad Injektivität und Surjektivität äquivalent. Nun bedeutet aber $\operatorname{ad}_X(Y) = 0$ genau $XY = YX$, also besteht der Kern von ad aus jenen Matrizen in $\mathfrak{su}(2)$, die mit allen anderen Matrizen in $\mathfrak{su}(2)$ kommutieren. Da wir das später noch benötigen werden, rechnen wir für Matrizen der Form $\left(\begin{smallmatrix} z & -\bar{w} \\ w & \bar{z} \end{smallmatrix}\right)$, die sowohl $\mathfrak{su}(2)$ (als die Matrizen mit $z \in i\mathbb{R}$) als auch $SU(2)$ (als die

Matrizen mit $|z|^2+|w|^2=1$ umfasst:

$$\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}=\begin{pmatrix} iz & i\bar{w} \\ iw & -i\bar{z} \end{pmatrix} \qquad \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}=\begin{pmatrix} iz & -i\bar{w} \\ -iw & -i\bar{z} \end{pmatrix}$$

Also kommutieren diese Matrizen nur für $w=0$. Analog rechnet man, dass $\left(\begin{smallmatrix} z & -\bar{w} \\ w & \bar{z} \end{smallmatrix}\right)$ nur dann mit $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ kommutiert, wenn sowohl z also auch w reell sind. Da sowohl $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ als auch $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ in $\mathfrak{su}(2)$ liegt, sehen wir, dass $\mathrm{ad}_X=0$ nur für $X=0$ gilt.

Da $SO(3)$ nach Satz 3.3 zusammenhängend ist, folgt nach Satz 2.10 aus der Surjektivität von ad die Surjektivität von Ad. Nach dem gleichen Satz folgt aus der Injektivität von ad, dass Ker(Ad) im Zentrum von $SU(2)$ enthalten sein muss. Da aber die Matrizen $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ auch in $SU(2)$ liegen, folgt aus der obigen Rechnung, dass eine Matrix $A\in SU(2)$ nur dann mit allen Elementen von $SU(2)$ kommutieren kann, wenn A ein reelles Vielfaches von $\mathbb{I}$ ist, was wegen $\det(A)=1$ sofort $A=\pm\mathbb{I}$ impliziert. Nachdem Vielfache von $\mathbb{I}$ natürlich mit jeder Matrix kommutieren, sehen wir, dass $Z(SU(2))=\{\pm\mathbb{I}\}=\mathrm{Ker}(\mathrm{Ad})\cong\mathbb{Z}_2$ gilt.

Da ein surjektiver Homomorphismus $\varphi:G\to H$ zwischen zwei Gruppen einen Isomorphismus $G/\mathrm{Ker}(\varphi)\to H$ induziert, folgt die behauptete Beschreibung von $SO(3)$ als Gruppe. Um die Beschreibung als Raum zu erhalten, erinnern wir uns, dass die Identifikation von $SU(2)$ mit $S^3\subset\mathbb{C}^2$ einfach dadurch gegeben war, dass eine Matrix A auf ihren ersten Spaltenvektor abbildet, der ein Einheitsvektor in $\mathbb{C}^2\cong\mathbb{R}^4$ ist. Im Quotienten $SU(2)/\{\pm\mathbb{I}\}$ muss man aber jede Matrix A mit $A\cdot(-\mathbb{I})=-A$ identifizieren und ist $x\in S^3\subset\mathbb{C}^2$ der erste Spaltenvektor von A, dann ist die erste Spalte von $-A$ gerade $-x$. Nachdem ein (reell) 1–dimensionaler Teilraum von $\mathbb{R}^4\cong\mathbb{C}^2$ die Einheitssphäre S^3 in genau zwei Punkten der Form $\pm x$ schneidet, kann man diesen Quotienten der Sphäre mit der Raum $\mathbb{R}P^3$ der eindimensionalen Teilräume von $\mathbb{R}^4$ identifizieren. □

Im einfachen Fall von $SU(2)$ kann man die allgemeinen Resultate über Ad und ad, die wir verwendet haben auch durch direkte Berechnung von Matrixdarstellungen bezüblich der Orthonormalbasis von oben ersetzen. Insbesondere zeigt man leicht, dass für $X=\left(\begin{smallmatrix} ia & -u+iv \\ u+iv & -ia \end{smallmatrix}\right)$ die Matrixdarstellung von ad_X durch $\left(\begin{smallmatrix} 0 & 2v & -2u \\ -2v & 0 & -2a \\ 2u & 2a & 0 \end{smallmatrix}\right)$ gegeben ist. Das zeigt natürlich sofort, dass ad einen linearen Isomorphismus $\mathfrak{su}(2)\to\mathfrak{so}(3)$ definiert.

3.5. Die Struktur von $SO(4)$. In Dimension $n=4$ können wir eine ähnlichen Methode verwenden, um die Gruppe $SO(4)$ zu beschreiben. Wir betrachten die Menge

$$\mathbb{H}:=\left\{\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}: z,w\in\mathbb{C}\right\}\subset M_2(\mathbb{C})$$

von Matrizen aus dem Beweis von Proposition 3.4. Das ist offensichtlich ein 4–dimensionaler, reeller Teilraum von $M_2(\mathbb{C})$ und wir haben schon gesehen, dass $\mathbb{H}$ sowohl $\mathfrak{su}(2)$ (Elemente mit $z\in i\mathbb{R}$) als auch $SU(2)$ (Elemente mit $|z|^2+|w|^2=1$) enthält. Eine kurze Rechnung (siehe Übungen) zeigt, dass $\mathbb{H}$ abgeschlossen unter der Matrixmultiplikation ist, also für $X,Y\in\mathbb{H}$ auch $XY\in\mathbb{H}$ gilt. Für $X\in\mathbb{H}$ ist $\det(X)=|z|^2+|w|^2$ und das ist für $X\neq 0$ positiv. Damit ist jedes Element $X\in\mathbb{H}\setminus\{0\}$ eine invertierbare Matrix und man verifiziert sofort, dass $X^{-1}=\frac{1}{\det(X)}X^*$ gilt und dass auch diese Matrix in $\mathbb{H}$ liegt. Damit folgt, dass $\mathbb{H}$ ein Schiefkörper ist, also alle Körperaxiome außer der Kommutativität der Multiplikation erfüllt.

Für $X,Y\in\mathbb{H}$ definieren wir $\langle X,Y\rangle:=\frac{1}{2}\mathrm{tr}(X^*Y)$ (was auf dem Teilraum $\mathfrak{su}(2)$ offensichtlich mit dem oben definierten inneren Produkt übereinstimmt). Eine einfache direkte Rechnung zeigt sofort, dass dies dem (reellen) standard inneren Produkt entspricht, wenn wir $\mathbb{H}$ analog wie oben über die erste Spalte mit $\mathbb{C}^2 \cong \mathbb{R}^4$ identifizieren. Insbesondere folgt aus der Rechnung von oben sofort, dass $\langle X, X\rangle = \det(X)$ gilt.

Bevor wir die Beschreibung von $SO(4)$ geben können, benötigen wir noch ein Lemma:

Lemma 3.5. *Seien $G \subset GL(n,\mathbb{R})$ und $H \subset GL(m,\mathbb{R})$ Matrixgruppen mit Lie Algebren $\mathfrak{g}$ und $\mathfrak{h}$. Dann kann man die Produktgruppe $G \times H$ (also das Produkt der Mengen mit den komponentenweisen Operationen) in natürlicher Weise als abgeschlossene Untergruppe von $GL(n+m,\mathbb{R})$ betrachten, also ist $G\times H$ wieder eine Matrixgruppe. Die Lie Algebra dieser Gruppe kann man mit $\mathfrak{g}\times\mathfrak{h}$ identifizieren, wobei sowohl die Lie Klammer als auch die Exponentialabbildung komponentenweise gegeben ist.*

Beweis. Betrachten wir $M_{m+n}(\mathbb{R})$ und nennen eine Matrix $X = (x_{ij})$ in diesem Raum block–diagonal mit Blöcken der Größe n und m wenn $x_{ij} = 0$ für alle $i > n$ und $j \leq n$ sowie für alle $i \leq n$ und $j > n$ gilt. Diese Matrizen bilden offensichtlich einen Teilraum in $M_{m+n}(\mathbb{R})$ und wir schreiben Elemente in diesem Raum als Blockmatrizen der Form $\left(\begin{smallmatrix} X_1 & 0 \\ 0 & X_2 \end{smallmatrix}\right)$ mit $X_1 \in M_n(\mathbb{R})$ und $X_2 \in M_m(\mathbb{R})$. Aus der Definition der Matrixmultiplikation folgt sofort, dass $\left(\begin{smallmatrix} X_1 & 0 \\ 0 & X_2 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} Y_1 & 0 \\ 0 & Y_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} X_1Y_1 & 0 \\ 0 & X_2Y_2 \end{smallmatrix}\right)$ gilt. Daraus folgt einerseits sofort ein analoges Resultat für den Kommutator. Andererseits ist natürlich der Schnitt dieses Teilraumes mit $GL(n+m,\mathbb{R})$ eine abgeschlossene Teilmenge und eine Blockmatrix $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$ ist genau dann invertierbar, wenn $A \in GL(n,\mathbb{R})$ und $B \in GL(m,\mathbb{R})$ gilt. Damit sehen wir aber sofort, dass die invertierbaren Blockdiagonalmatrizen eine abgeschlossene Untergruppe in $GL(n+m,\mathbb{R})$ bilden. Diese Gruppe ist in offensichtlicher Weise isomorph zu $GL(n,\mathbb{R}) \times GL(m,\mathbb{R})$ und ihre Lie Algebra ist in der gleichen Weise isomorph zu $\mathfrak{gl}(n,\mathbb{R}) \times \mathfrak{gl}(m,\mathbb{R})$ mit komponentenweiser Klammer. Schließlich folgt aus der Definition des Matrizenexponentials sofort, dass $\exp\left(\begin{smallmatrix} X_1 & 0 \\ 0 & X_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} \exp(X_1) & 0 \\ 0 & \exp(X_2) \end{smallmatrix}\right)$ gilt.

Daraus folgt das allgemeine Resultat leicht: Die Matrizen der Form $\left(\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right)$ mit $A \in G$ und $B \in H$ bilden eine abgeschlossene Untergruppe der Gruppe aller invertierbaren Blockdiagonalmatrizen und damit auch von $GL(n+m,\mathbb{R})$. Für die Lie Algebra erhalten wir den Raum aller Blockmatrizen deren Komponenten in $\mathfrak{g}$ bzw. $\mathfrak{h}$ liegen und damit folgen alle Behauptungen. □

Damit können wir nun eine Beschreibung der Gruppe $SO(4)$ geben, die wir als die spezielle orthogonale Gruppe von $\mathbb{H}$ betrachten können (siehe Abschnitt 2.11). Andererseits können wir das Produkt $SU(2) \times SU(2)$ von zwei Kopien der Matrixgruppe $SU(2)$ bilden. Für $A, B \in SU(2)$ und $X \in \mathbb{H}$ betrachten wir nun AXB^{-1}, was wegen $SU(2) \subset \mathbb{H}$ nach den obigen Resultaten wieder in $\mathbb{H}$ liegt. Fixieren wir A und B und betrachten die Abbildung $\varphi_{(A,B)} : \mathbb{H} \to \mathbb{H}$, die durch $\varphi_{(A,B)}(X) := AXB^{-1}$ gegeben ist. Dann ist diese Funktion offensichtlich linear und invertierbar mit Inverser $\varphi_{(A^{-1},B^{-1})}$. Damit können wir die Abbildung $(A, B) \mapsto \varphi_{(A,B)}$ als Funktion $\varphi : SU(2) \times SU(2) \to GL(\mathbb{H})$ betrachten.

Satz 3.5. *Die Funktion φ hat Werte in $SO(\mathbb{H})$ und $\varphi : SU(2) \times SU(2) \to SO(\mathbb{H})$ ist ein surjektiver glatter Homomorphismus mit Kern $\{(\mathbb{I},\mathbb{I}), (-\mathbb{I},-\mathbb{I})\}$. Insbesondere ist die Lie Algebra $\mathfrak{so}(4)$ isomorph zu $\mathfrak{su}(2) \times \mathfrak{su}(2)$, während $SO(4)$ als Gruppe isomorph zu $(SU(2) \times SU(2))/\mathbb{Z}_2$ ist. Als Raum (und als Mannigfaltigkeit) erhält man $SO(4)$ indem man in einem Produkt $S^3 \times S^3$ jeweils den Punkt (x, y) mit dem Punkt $(-x,-y)$ identifiziert.*

Beweis. Aus der Definition folgt sofort, dass $\varphi_{(A_1,B_1)} \circ \varphi_{(A_2,B_2)}$ eine Matrix $X \in \mathbb{H}$ auf $A_1A_2XB_2^{-1}B_1^{-1} = (A_1A_2)X(B_1B_2)^{-1}$ abbildet, also stimmt das mit $\varphi_{(A_1A_2,B_1B_2)}$ überein. Damit ist φ ein Homomorphismus. Von oben wissen wir, dass $\langle X, X\rangle = \det(X)$ für alle $X \in \mathbb{H}$ gilt. Wegen $\det(A) = \det(B) = 1$ folgt $\det(AXB^{-1}) = \det(X)$, also ist $\|\varphi_{(A,B)}(X)\| = \|X\|$ für alle $X \in \mathbb{H}$ und somit hat φ Werte in $O(\mathbb{H})$. Da $SU(2)$ bogenzusammenhängend ist, ist auch $SU(2) \times SU(2)$ bogenzusammenhängend und wir folgern, dass die Werte von φ sogar in $SO(\mathbb{H})$ liegen müssen.

Aus der Glattheit der Matrizenmultiplikation und der Inversion folgt sofort, dass für fixes $X \in \mathbb{H}$ die Abbildung $(A, B) \mapsto AXB^{-1}$ glatt als Funktion $SU(2) \times SU(2) \to M_2(\mathbb{C})$ ist. Setzt man aber nun für X ein Element einer fixen Orthonormalbasis für $\mathbb{H}$ ein und entwickelt im Bild nach dieser Orthonormalbasis, dann kann man die Koeffizienten als innere Produkte schreiben, also hängen auch sie glatt von (A, B) ab. Diese Koeffizienten bilden aber gerade die Eintragungen der Matrixdarstellung von $\varphi_{(A,B)}$ in $L(\mathbb{H}, \mathbb{H})$ und die Glattheit von φ folgt.

Damit können wir die Ableitung $\varphi' : \mathfrak{su}(2) \times \mathfrak{su}(2) \to \mathfrak{so}(\mathbb{H})$ bilden. Aus unseren Resultaten über die Ableitung der Matrixmultiplikation und der Inversion in Satz 1.2 folgt sofort, das für $Z, W \in \mathfrak{su}(2)$ die Ableitung $\varphi'_{(Z,W)}$ eine Matrix $Y \in \mathbb{H}$ auf $ZY - YW$ abbildet. Wir wissen, dass $\mathfrak{su}(2)$ Dimension 3 hat, also hat $\mathfrak{su}(2) \times \mathfrak{su}(2)$ Dimension 6. Für den Raum $\mathfrak{so}(4)$ der schiefsymmetrischen 4×4–Matrizen ergibt sich als Dimension $\frac{1}{2} \cdot 4 \cdot 3 = 6$, also können wir zeigen, dass φ' ein linearer Isomorphismus ist, indem wir Injektivität von φ' verifizieren. Nun gilt aber $\varphi'_{(Z,W)} = 0$ genau dann, wenn $ZY = YW$ für alle $Y \in \mathbb{H}$ gilt. Setzen wir $Y = \mathbb{I}$, dann folgt $Z = W$, und wir sehen, dass $Z \in \mathfrak{su}(2)$ eine Matrix sein muss, die mit allen $Y \in \mathbb{H}$ kommutiert. Aber im Beweis von Satz 3.4 haben wir schon gesehen, dass diese Eigenschaft für alle $Y \in \mathfrak{su}(2) \subset \mathbb{H}$ schon $Z = 0$ impliziert. Damit ist φ' ein linearer Isomorphismus und nach Satz 2.10 folgt die Surjektivität von φ.

Auch die Beschreibung des Kerns von φ kann man auf den Beweis von Satz 3.4 zurückführen. Ein Paar (A, B) liegt genau dann in diesem Kern, wenn $AXB^{-1} = X$ für alle $X \in \mathbb{H}$ gilt. Setzen wir $X = \mathbb{I}$ ein, dann erhalten wir $AB^{-1} = \mathbb{I}$ und damit $B = A$. Aus dem Beweis von Satz 3.4 wissen wir aber schon, dass $AXA^{-1} = X$ für alle $X \in \mathfrak{su}(2) \subset \mathbb{H}$ nur für $A = \pm\mathbb{I}$ gelten kann. Da $(-\mathbb{I}, -\mathbb{I})$ offensichtlich im Kern von φ liegt, erhalten wir die Beschreibung des Kerns und der Rest folgt dann völlig analog wie im Beweis von Satz 3.4. □

Dieses Resultat ist in mehrfacher Hinsicht sehr bemerkenswert. Abgesehen von der grundsätzlich unerwarteten Verbindung zwischen reellen und komplexen Matrizen ist insbesondere die Tatsache, dass $\mathfrak{so}(4)$ isomorph zu einem Produkt ist, gänzlich unerwartet. Diese Produktstruktur hat algebraische Konsequenzen. Betrachten wir nämlich in $\mathfrak{g} := \mathfrak{su}(2) \times \mathfrak{su}(2)$ die Teilmenge $\mathfrak{h} = \{(X, 0) : X \in \mathfrak{su}(2)\}$, dann ist das nicht nur eine Teilalgebra, sonder erfüllt sogar, dass die Klammer eines Elements von $\mathfrak{h}$ mit einem beliebigen Element von $\mathfrak{g}$ wieder in $\mathfrak{h}$ liegt. Somit ist also $\mathfrak{h} \subset \mathfrak{g}$ ein Ideal, siehe Abschnitt 2.10. Analoges gilt natürlich für die Teilmenge $\{(0, X) : X \in \mathfrak{su}(2)\}$, also gibt es in der 6–dimensionalen Lie Algebra $\mathfrak{so}(4)$ zwei dreidimensionale Ideale.

Im Gegensatz dazu haben wir in Beispiel 2.8 gesehen, dass ein Ideal in $\mathfrak{g} := \mathfrak{sl}(2, \mathbb{R})$, das $\neq \{0\}$ ist, schon mit $\mathfrak{g}$ übereinstimmen muss. (Wir haben dort nur die einfache Tatsache verifiziert, dass der Kern eines Homomorphismus von Lie Algebren automatisch ein Ideal ist und dann mit Idealen gearbeitet.) Die Algebra $\mathfrak{g}$ besitzt also nur die trivialen Ideale $\{0\}$ und $\mathfrak{g}$ und analog zur Gruppentheorie nennt man so eine Lie Algebra *einfach* (wenn zusätzlich $\dim(\mathfrak{g}) > 1$ gilt). Es stellt sich heraus, dass die Lie Algebren

$\mathfrak{sl}(n,\mathbb{R})$, $\mathfrak{sl}(n,\mathbb{C})$ und $\mathfrak{su}(n)$ für $n \geq 2$, sowie auch $\mathfrak{so}(n)$ für $n = 3$ und $n \geq 5$ alle einfach sind. Nur $\mathfrak{so}(4)$ ist ein Produkt von zwei einfachen Lie Algebren und man nennt Produkte von endlich vielen solchen Algebren *halbeinfach*. Es stellt sich heraus, dass man einfache (und damit halbeinfache) Lie Algebren über $\mathbb{R}$ und $\mathbb{C}$ mit Methoden der linearen Algebra vollständig klassifizieren kann und dabei viele interessante Beziehungen zu anderen Teilen der Mathematik auftauchen.

Bemerkung 3.5. Alternativ zu der obigen Beschreibung in Termen der Matrixgruppe $SU(2)$ kann man die Beschreibungen von $SO(3)$ und $SO(4)$ aus den letzten beiden Abschnitt auch ganz in Termen des Schiefkörpers $\mathbb{H}$ der *Hamilton'schen Quaternionen* interpretieren. Dazu verwendet man, dass die Einheitsmatrix $\mathbb{I}$ zusammen mit der Basis von $\mathfrak{su}(2)$ aus Abschnitt 3.4 ein reelle Basis $\{1, i, j, k\}$ von $\mathbb{H}$ bildet, sodass $i^2 = j^2 = -1$ und $ij = -ji = k$ ist (woraus sich alle Produkte der Basiselemente ergeben). Dann schreibt man $q \in \mathbb{H}$ als $a + bi + cj + dk$ mit $a, b, c, d \in \mathbb{R}$ und definiert die Multiplikation durch Bilinearität. Weiters definiert man eine quaternionische Konjugation, indem man $\bar{q} := a - bi - cj - dk$ definiert und verifiziert, dass $(p, q) \mapsto \frac{1}{2}(p\bar{q} + q\bar{p})$ ein inneres Produkt auf $\mathbb{H}$ definiert, sodass $q\bar{q} = |q|^2 \cdot 1$ für die zugehörige Norm gilt. Damit bildet $\{q : |q| = 1\}$ eine Untergruppe von $\mathbb{H}$, die Gruppe der *Einheitsquaternionen*, die oft mit $Sp(1)$ bezeichnet wird. In der Matrizeninterpretation ist $Sp(1) = SU(2)$. Für Satz 3.4 betrachtet man dann den von i, j und k aufgespannten Teilraum der *rein imaginären Quaternionen* und darauf die Wirkung von $p \in Sp(1)$ durch $q \mapsto pqp^{-1}$. Für Satz 3.5 betrachtet man die Wirkung von $(p_1, p_2) \in Sp(1) \times Sp(1)$ auf $\mathbb{H}$, die durch $q \mapsto p_1 q p_2^{-1}$ gegeben ist.

3.6. Darstellungen. Eine wichtige Klasse von glatten Homomorphismen zwischen Matrixgruppen (und allgemeiner zwischen Lie Gruppen) sind die sogenannten Darstellungen. Der Vorteil hier ist, dass man Methoden der linearen Algebra zur Konstruktion und zum Studium von Darstellungen verwenden kann.

Definition 3.6. (1) Eine *Darstellung* einer Matrixgruppe G *auf einem endlichdimensionalen Vektorraum* V ist ein glatter Homomorphismus $\varphi : G \to GL(V)$. Ist V ein komplexer Vektorraum und hat φ komplex lineare Abbildungen als Werte, so spricht man von einer *komplexen Darstellung*.

(2) Eine *Darstellung* einer Lie Algebra $\mathfrak{g}$ auf einem endlichdimensionalen Vektorraum V ist ein Homomorphismus $f : \mathfrak{g} \to \mathfrak{gl}(V)$ von Lie Algebren.

(3) Ein *Morphismus* zwischen Darstellungen φ von G auf V und ψ von G auf W ist eine lineare Abbildung $F : V \to W$ sodass $F \circ \varphi(g) = \psi(g) \circ F$ für alle $g \in G$ gilt. Ein bijektiver Morphismus heißt ein *Isomorphismus*. Analog definiert man Morphismen und Isomorphismen zwischen Darstellungen von Lie Algebren.

Eine Darstellung einer Lie Algebra ist ein viel einfacheres Objekt als eine Darstellung einer Matrixgruppe. Während zunächst nicht klar ist, wie restriktiv die Glattheitsbedingung bei einer Darstellung einer Matrixgruppe ist, hat man es im Fall der Lie Algebra nur mit einer einzelnen lineare Abbildung zu tun, die noch eine zusätzliche Bedingung (nämlich die Verträglichkeit mit den Lie Klammern) erfüllen muss. Offensichtlich ist für jede Darstellung $\varphi : G \to GL(V)$ die Ableitung φ' eine Darstellung von $\mathfrak{g}$ auf V, also kann man etwa Satz 2.10 direkt auf Darstellungen anwenden. Für jede Matrixgruppe hat man eine offensichtliche Darstellung. Ist nämlich G eine abgeschlossene Untergruppe von $GL(n,\mathbb{R})$, dann definiert die Inklusion von G nach $GL(n,\mathbb{R})$ natürlich eine Darstellung von G auf $\mathbb{R}^n$, die oft als die *Standarddarstellung* von G bezeichnet wird. Andererseits haben wir in Satz 2.11 gesehen, dass wir für jede Matrixgruppe G mit Lie

Algebra $\mathfrak{g}$ eine natürliche Darstellung $\mathrm{Ad} : G \to GL(\mathfrak{g})$ erhalten (die sogar Werte in der abgeschlossenen Untergruppe $\mathrm{Aut}(\mathfrak{g})$ hat), die *adjungierte Darstellung*.

Man kann Darstellungen auch in dem aus der Algebra bekannten Bild der Gruppenwirkungen betrachten. Ist $\varphi : G \to GL(V)$ eine Darstellung dann kann man die Abbildung $G \times V \to V$ betrachten, die durch $(g, v) \mapsto \varphi(g)v$ gegeben ist. Ist φ aus dem Zusammenhang klar, dann wird diese Abbildung oft einfach mit $(g, v) \mapsto g \cdot v$ bezeichnet. In diesem Bild erhält man die definierende Eigenschaften einer Gruppenwirkung, nämlich $e \cdot v = v$ für das neutrale Element $e \in G$ und alle $v \in V$ und $gh \cdot v = g \cdot (h \cdot v)$ für alle $g, h \in G$ und $v \in V$. Zusätzlich haben wir natürlich die Information, dass für jedes $g \in G$ die Abbildung $v \mapsto g \cdot v$ linear ist.

Für Darstellungen einer Lie Algebra gibt es eine ähnliche Beschreibung, die sogar noch einfacher ist. Für eine lineare Abbildung $f : \mathfrak{g} \to \mathfrak{gl}(V)$ kann man natürlich die Funktion $(X, v) \mapsto f(X)(v)$ als bilineare Funktion $\mathfrak{g} \times V \to V$ betrachten. Wenn die betroffene Darstellung klar ist, bezeichnet man diese Abbildung ebenfalls mit $(X, v) \mapsto X \cdot v$. Umgekehrt kann man für eine bilineare Abbildung $F : \mathfrak{g} \times V \to V$ und jedes Element $X \in \mathfrak{g}$ eine lineare Abbildung $f(X) : V \to V$ durch $f(X)(v) := F(X, v)$ definieren und erhält so eine lineare Abbildung $f : \mathfrak{g} \to \mathfrak{gl}(V)$. Man sieht dann sofort, dass die Bedingung, dass f ein Homomorphismus von Lie Algebren ist, in der einfacheren Notation äquivalent zu $[X, Y] \cdot v = X \cdot (Y \cdot v) - Y \cdot (X \cdot v)$ ist.

Im Fall von Gruppen ist die Charakterisierung wegen der Glattheitsbedingung etwas komplizierter:

Lemma 3.6. *Sei G eine Matrixgruppe, V ein endlichdimensionaler Vektorraum und betrachten wir eine Wirkung von G auf V, die wir als $(g, v) \mapsto g \cdot v$ schreiben. Nehmen wir weiters an, dass für jedes fixe $g \in G$ die Abbildung $v \mapsto g \cdot v$ linear und für jedes fixe $v \in V$ die Abbildung $g \mapsto g \cdot v$ eine glatte Funktion $G \to V$ ist. Dann definiert $\varphi(g)(v) := g \cdot v$ einen glatten Homomorphismus $\varphi : G \to GL(V)$.*

Beweis. Für jedes $g \in G$ ist $\varphi(g)$ eine lineare Abbildung $V \to V$ und aus den Wirkungseigenschaften folgt sofort, dass $\varphi(g) \circ \varphi(g^{-1}) = \varphi(e) = \mathrm{id}_V$ gilt, also ist $\varphi(g) \in GL(V)$. Aus der Algebra ist bekannt, dass die Wirkungseigenschaften zeigen, dass φ einen Homomorphismus $G \to GL(V)$ definiert, also bleibt nur noch die Glattheit zu zeigen. Dazu sei $\{v_1, \dots, v_n\}$ eine Basis für V und für jedes $i = 1, \dots, n$ sei $\lambda_i : V \to \mathbb{K}$ die eindeutige lineare Abbildung, die $\lambda_i(v_i) = 1$ und $\lambda_i(v_j) = 0$ für $j \neq i$ erfüllt. Ist $A(g) = (a_{ij}(g))$ die Matrixdarstellung von $\varphi(g)$ bezüglich der Basis $\{v_i\}$, dann gilt nach Definition $a_{ij}(g) = \lambda_i(g \cdot v_j)$ für alle i und j. Nun ist aber nach Voraussetzung $g \mapsto g \cdot v_j$ eine glatte Funktion $G \to V$ und weil λ_j linear ist, ist auch $g \mapsto a_{ij}(g)$ glatt. Wie wir aus Abschnitt 2.11 wissen, bedeutet das genau, dass φ glatt als Funktion $G \to GL(V)$ ist und damit eine Darstellung definiert. □

Beispiel 3.6. Wie schon erwähnt können viele Konstruktionen der linearen Algebra in natürlicher Weise auf Darstellungen angewandt werden, wir besprechen hier nur einige Beispiele. Dabei ist zu beachten, dass die Konstruktionen für Matrixgruppen und für Lie Algebren nicht gleich sind, sondern durch eine Ableitung miteinander verbunden sind. Sei G eine Matrixgruppe mit Lie Algebra $\mathfrak{g}$ und seien Darstellungen von G auf Vektorräumen V und W gegeben.

(1) Die einfachste dieser Konstruktionen ist die *direkte Summe* von Darstellungen. Wir betrachten $V \oplus W = V \times W$ mit der Operation $g \cdot (v, w) := (g \cdot v, g \cdot w)$, wobei auf der rechten Seite die gegebenen Wirkungen auf V und W verwendet werden. Nach Konstruktion sind die Voraussetzungen von Lemma 3.6 erfüllt, also erhalten wir so eine Darstellung von G auf $V \oplus W$. Die Konstruktion für Darstellungen von Lie Algebren

ist in diesem Fall völlig analog, man definiert einfach $X \cdot (v, w) = (X \cdot v, X \cdot W)$ und verifiziert sofort, dass die Darstellungseigenschaft erfüllt ist.

Analog kann man für eine Darstellung von G auf V und eine Darstellung von H auf W eine Darstellung von $G \times H$ auf $V \oplus W$ definieren, indem man einfach $(g, h) \cdot (v, w) := (g \cdot v, h \cdot w)$ setzt (und analog für Lie Algebren). Tatsächlich haben wir diese Konstruktion schon in Abschnitt 3.5 verwendet um $GL(n, \mathbb{R}) \times GL(m, \mathbb{R})$ zu einer Matrixgruppe zu machen.

(2) Betrachten wir den Raum $L(V, W)$ aller lineare Abbildungen von V nach W und definieren eine Abbildung $G \times L(V, W) \to L(V, W)$ durch $(g \cdot f)(v) := g \cdot (f(g^{-1} \cdot v))$, wobei auf der rechten Seite der erste Punkt die Wirkung auf W und der zweite Punkt die Wirkung auf V bezeichnet. Zunächst ist für fixes $g \in G$ und $f \in L(V, W)$ die so definierte Funktion $g \cdot f$ eine Komposition von drei linearen Abbildungen und damit selbst linear, also erhalten wir tatsächlich eine Funktion nach $L(V, W)$. Aus den Eigenschaften der Wirkungen auf V und W folgt sofort, dass dies eine Gruppenwirkung definiert, also $e \cdot f = f$ und $gh \cdot f = g \cdot (h \cdot f)$ gilt. Für $f_1, f_2 \in L(V, W)$ und $t \in \mathbb{K}$ betrachten wir $g \cdot (f_1 + tf_2)$. Das bildet $v \in V$ auf $g \cdot (f_1(g^{-1} \cdot v) + tf_2(g^{-1} \cdot v))$ ab und weil die Wirkung auf W linear ist schließen wir leicht, dass $g \cdot (f_1 + tf_2) = (g \cdot f_1) + t(g \cdot f_2)$ gilt, also ist die Wirkung linear in der zweiten Variablen.

Um die Glattheit der Wirkung zu verifizieren, wählen wir Basen $\{v_1, \dots, v_n\}$ für V und $\{w_1, \dots, w_m\}$ für W. Damit erhalten wir eine Basis für $L(V, W)$ und die Entwicklung von $f \in L(V, W)$ in dieser Basis entspricht gerade dem Übergang zur Matrixdarstellung von f bezüglich der beiden Basen. Stellen wir die Wirkungen von G auf V und W ebenfalls durch Matrizen dar, dann liefern sie glatte Funktionen $A : G \to GL(n, \mathbb{R})$ und $B : G \to GL(m, \mathbb{R})$. Ist $C \in M_{m,n}(\mathbb{R})$ die Matrixdarstellung zu f, dann erhält man nach Konstruktion die Matrix zu $g \cdot f$ als $B(g)CA(g)^{-1}$. Aus der Glattheit der Matrixmultiplikation und –inversion folgt damit, das die Matrix von $g \cdot f$ glatt von g abhängt und nach Lemma 3.6 haben wir eine Darstellung von G auf $L(V, W)$ konstruiert. Nach Definition ist $f \in L(V, W)$ genau dann ein Morphismus von Darstellungen, wenn $g \cdot f = f$ für alle $g \in G$ gilt.

Auch hier kann man analog aus einer Darstellung von G auf V und einer Darstellung von H auf W eine Darstellung von $G \times H$ auf $L(V, W)$ konstruieren. Schließlich kann man $W = \mathbb{R}$ wählen und darauf die triviale G–Darstellung $g \cdot t = t$ für alle $g \in G$ und $t \in \mathbb{R}$ betrachten. Dann ist $L(V, \mathbb{R}) = V^*$ der Dualraum zu V und man erhält die *duale Darstellung*, die explizit durch $(g \cdot \alpha)(v) = \alpha(g^{-1} \cdot v)$ für $\alpha \in V^*$, $g \in G$ und $v \in V$ definiert ist.

(3) Um zu verstehen, wie Darstellungen einer Lie Algebra auf V und W eine Darstellung auf $L(V, W)$ liefern, betrachtet man die Konstruktion für Gruppen aus (2) und differenziert. Wählen wir Basen für V und W, wie oben, dann ist die Wirkung aus (2) durch $g \cdot C = B(g)CA(g)^{-1}$ gegeben. Damit rechnet man analog zu Abschnitt 2.11 leicht nach, dass die Ableitung der Wirkung durch $X \cdot C = (DB(\mathbb{I})(X))C - C(DA(\mathbb{I})(X))$ gegeben ist. Übersetzt in die Sprache der linearen Abbildungen bedeutet das, dass wir für Darstellungen einer Lie Algebra auf V und W die zugehörige Darstellung auf $L(V, W)$ durch $(X \cdot f)(v) := X \cdot (f(v)) - f(X \cdot v)$ definieren sollten. Eine einfache Rechnung (siehe Übungen) zeigt dann, dass man so tatsächlich allgemein eine Darstellung von $\mathfrak{g}$ auf $L(V, W)$ erhält. Insbesondere liefert das die duale Darstellung auf V^*, die durch $(X \cdot \alpha)(v) = -\alpha(X \cdot v)$ für $\alpha \in V^*$, $X \in \mathfrak{g}$ und $v \in V$ gegeben ist.

(4) Die Konstruktionen aus (2) und (3) verallgemeinern sich einfach auf bi– und multilineare Abbildungen, wir beschränken uns hier auf einige Spezialfälle. Für Darstellungen von G auf V und W betrachten wir den Raum aller bilinearen Abbildungen

$V \times W \to \mathbb{R}$, der üblicherweise mit $V^* \otimes W^*$ bezeichnet wird. Darauf definiert man $(g \cdot \alpha)(v, w) := \alpha(g^{-1} \cdot v, g^{-1} \cdot w)$ und rechnet leicht, dass dies wieder eine Gruppenwirkung durch lineare Abbildungen definiert. Fixieren wir Basen $\{v_i\}$ und $\{w_j\}$ für V und W, dann erhalten wir eine entsprechende Basis $\{\alpha_{ij}\}$ für $V^* \otimes W^*$ wobei die Abbildungen α_{ij} durch $\alpha_{ij}(v_k, w_\ell) = \delta_{ik}\delta_{j\ell}$ charakterisiert ist. Fixieren wir i und j, dann wissen wir, dass $g \mapsto g^{-1} \cdot v_i$ und $g \mapsto g^{-1} \cdot w_j$ glatte Funktionen $G \to V$ und $G \to W$ definieren. Jede Funktion $\alpha \in V^* \otimes W^*$ ist bilinear und damit glatt als Abbildung $V \times W \to \mathbb{R}$, also ist auch $g \mapsto (g \cdot \alpha)(v_i, w_j)$ eine glatte Funktion $G \to \mathbb{R}$. Nach Konstruktion ist aber $(g \cdot \alpha)(v_i, w_j) \in \mathbb{R}$ genau der Koeffizient von α_{ij} in der Darstellung von $g \cdot \alpha$ als Linearkombination dieser Basiselemente. Damit folgt aber, dass $g \mapsto g \cdot \alpha$ für jedes $\alpha \in V^* \otimes W^*$ eine glatte Funktion $G \to V^* \otimes W^*$ definiert. Nach Lemma 3.6 erhalten wir damit eine Darstellung von G auf $V^* \otimes W^*$.

Das geht analog für den Raum der k–linearen Abbildungen $V^k \to \mathbb{R}$, der üblicherweise mit $\otimes^k V^*$ bezeichnet wird. Hier ist die Wirkung für $g \in G$ und beliebige Elemente $v_1, \dots, v_k \in V$ durch $(g \cdot \alpha)(v_1, \dots, v_k) = \alpha(g^{-1} \cdot v_1, \dots, g^{-1} \cdot v_k)$ gegeben. Aus der Definition folgt sofort, dass für eine symmetrische Abbildung α (also eine, die bei Permutation der Argumenten unverändert bleibt) auch $g \cdot \alpha$ symmetrisch ist. Damit liefert die gleiche Definition auch eine Darstellung auf dem Raum $S^k V^*$ der k–linearen, symmetrischen Funktionen $V^k \to \mathbb{R}$. Analog ist für eine alternierende k–lineare Abbildung (also eine, die bei Permutation der Elemente mit $\sigma \in \mathfrak{S}_k$ mit dem Signum $\operatorname{sgn}(\sigma)$ multipliziert wird) auch $g \cdot \alpha$ alternierend. Somit erhalten wir auch eine Darstellung auf dem Raum $\Lambda^k V^*$ der alternierenden k–linearen Abbildungen $V^k \to \mathbb{R}$.

Das funktioniert auch mit Darstellungen von Lie Algebren, wobei man aber wieder eine differenzierte Wirkung verwenden muss. Dabei ergibt sich auf $V^* \otimes W^*$ als Wirkung $(X \cdot \alpha)(v, w) := -\alpha(X \cdot v, w) - \alpha(v, X \cdot w)$. Analog erhält man (siehe Übungen) auf $\otimes^k V^*$, $S^k V^*$ und $\Lambda^k V^*$ die Vorschrift

$$(X \cdot \alpha)(v_1, \dots, v_k) = -\sum_{i=1}^k \alpha(v_1, \dots, v_{i-1}, X \cdot v_i, v_{i+1}, \dots, v_k). \tag{3.2}$$

3.7. Invariante Teilräume – Irreduzibilität. Hat man den Begriff der Darstellung verdaut, dann ergeben sich die folgenden Begriffe in natürlicher Weise:

Definition 3.7. Betrachten wir eine Darstellung einer Matrixgruppe G auf einem endlichdimensionalen Vektorraum V, die wir mit $(g, v) \mapsto g \cdot v$ bezeichnen.

(1) Ein *invarianter Teilraum* in V ist ein Teilraum $W \subset V$, sodass für alle $g \in G$ und $w \in W$ auch $g \cdot w \in W$ gilt.

(2) Die Darstellung V heißt *irreduzibel*, wenn $\{0\}$ und V die einzigen invarianten Teilräume von V sind.

(3) Die Darstellung V heißt *zerlegbar*, wenn es invariante Teilräume $W_1, W_2 \subset V$ gibt, die beide Dimension > 0 haben, sodass $V = W_1 \oplus W_2$ gilt. Gibt es keine solchen Teilräume, dann nennt man V *unzerlegbar*.

Diese Definitionen übertragen sich in offensichtlicher Weise auf Darstellungen von Lie Algebren.

Tatsächlich haben wir schon (allgemeine) Beispiele für invariante Teilräume gesehen, nämlich die Teilräume $S^k V^*$ und $\Lambda^k V^*$ in $\otimes^k V^*$ in Beispiel 3.6 (4).

Bevor wir die Zusammenhänge zwischen diesen Konzepten besprechen, beweisen wir ein erstes Resultat.

Lemma 3.7. *Sei $\varphi : G \to GL(V)$ eine Darstellung einer Matrixgruppe G auf einem Vektorraum V und sei $W \subset V$ ein invarianter Teilraum. Dann liefert φ in natürlicher Weise Darstellungen φ_W von G auf W und $\underline{\varphi}$ von G auf V/W. Analoges gilt für Darstellungen von Lie Algebren.*

Beweis. Nach Definition gilt für jedes Element $g \in G$, dass der zugehörige lineare Isomorphismus $\varphi(g) : V \to V$ erfüllt, dass $\varphi(g)(w) \in W$ für alle $w \in W$ gilt. Damit kann man einerseits $\varphi_W(g) := \varphi(g)|_W : W \to W$ bilden, was wieder ein linearer Isomorphismus ist. Andererseits erhält man eine induzierte lineare Abbildung $\underline{\varphi}(g) : V/W \to V/W$, die durch $\underline{\varphi}(g)(v + W) := \varphi(g)(v) + W$ gegeben ist. Damit erhalten wir Funktionen $\varphi_W : G \to GL(W)$ und $\underline{\varphi} : G \to GL(V/W)$ und aus der Konstruktion verifiziert man sofort, dass dies Gruppenhomomorphismen sind.

Um die Glattheit zu verifizieren, wählen wir eine Basis $\{w_1, \dots, w_k\}$ für W und ergänzen sie durch Vektoren $\{v_1, \dots, v_{n-k}\}$ zu einer Basis von V. Dann ist aus der linearen Algebra bekannt, dass die Nebenklassen $v_1 + W, \dots, v_{n-k} + W$ eine Basis für den Quotientenraum V/W bilden. Bezeichnen wir für $g \in G$ mit $A(g) = (a_{ij}(g))$ die Matrix zu $\varphi(g)$ bezüglich dieser Basis, dann bedeutet Invarianz des Teilraumes $W \subset V$ genau, dass $a_{ij}(g) = 0$ für alle g gilt, soferne $i \le k$ und $j > k$ ist. In Blockform mit Blöcken der Größe k und $n - k$ bedeute das genau, dass $A(g) = \begin{pmatrix} B(g) & D(g) \\ 0 & C(g) \end{pmatrix}$ für passende Matrizen $B(g)$, $C(g)$ und $D(g)$ ist. Nun ist aber nach Konstruktion $B(g)$ genau die Matrixdarstellung von $\varphi_W(g)$ bezüglich der Basis $\{w_1, \dots, w_k\}$ von W und $C(g)$ die Matrixdarstellung von $\underline{\varphi}(g)$ bezüglich der Basis $\{v_1 + W, \dots, v_{n-k} + W\}$ von V/W. □

Ist $\varphi : G \to GL(V)$ eine zerlegbare Darstellung und sind $W_1, W_2 \subset V$ die entsprechenden Teilräume, dann können wir die Identifikation $V = W_1 \oplus W_2$ als Isomorphismus zwischen der Darstellung φ und der direkten Summe von φ_{W_1} und φ_{W_2} im Sinne von Beispiel 3.6 (1) betrachten. Umgekehrt sind die beiden Summanden in einer direkten Summe natürlich invariante Teilräume, also ist eine Darstellung genau dann unzerlegbar, wenn sie nicht als Summe von zwei Darstellungen mit echt kleinerer Dimension geschrieben werden kann. Mehr oder weniger nach Definition kann man jede Darstellung als direkte Summe von unzerlegbaren Darstellungen schreiben, also genügt es, unzerlegbare Darstellungen zu verstehen.

Irreduzible Darstellungen sind nach Definition unzerlegbar, aber im Allgemeinen muss eine unzerlegbare Darstellung nicht irreduzibel sein. Ein Beispiel liefern die Standarddarstellungen der Gruppen $B(n, \mathbb{R})$ aus Beispiel 2.4 (3) (siehe Übungen). Es gibt jedoch wichtige Klassen von Matrixgruppen (und Lie Algebren) für die unzerlegbare Darstellungen automatisch irreduzibel sein müssen ("vollständige Reduzibilität"). Insbesondere gilt das für kompakte und für halbeinfache Matrixgruppen, die in vielen Anwendungen eine zentrale Rolle spielen. Für dieser Gruppen ist man primär daran interessiert, irreduzible Darstellungen zu verstehen.

Ein Schlüsselresultat ist, dass man im Fall von zusammenhängenden Gruppen invariante Teilräume auf dem Niveau der Lie Algebra studieren kann, was viele Fragen wesentlich vereinfacht:

Satz 3.7. *Sei G eine zusammenhängende Matrixgruppe mit Lie Algebra $\mathfrak{g}$ und sei $\varphi : G \to GL(V)$ eine Darstellung mit zugehöriger Lie Algebren Darstellung $\varphi' : \mathfrak{g} \to \mathfrak{gl}(V)$. Dann ist ein Teilraum $W \subset V$ genau dann G–invariant, wenn er $\mathfrak{g}$–invariant ist. Insbesondere ist V genau dann unzerlegbar bzw. irreduzibel als G–Darstellung, wenn die entsprechende Eigenschaft für die zugehörige $\mathfrak{g}$–Darstellung erfüllt ist.*

Beweis. Nach Definition gilt $\varphi'(X) = \frac{d}{dt}|_{t=0}\varphi(\exp(tX))$ für alle $X \in \mathfrak{g}$. Damit folgt aber für $v \in V$ sofort, dass $X \cdot v = \varphi'(X)(v) = \frac{d}{dt}|_{t=0}\varphi(\exp(tX))(v) = \frac{d}{dt}|_{t=0}\exp(tX) \cdot v$ gilt. Ist $W \subset V$ ein G–invarianter Teilraum, dann ist $\exp(tX) \cdot w \in W$ für alle $w \in W$ und $t \in \mathbb{R}$. Damit folgt $X \cdot w \in W$ nach Lemma 2.3, also ist W auch $\mathfrak{g}$–invariant.

Nehmen wir umgekehrt an, dass $W \subset V$ ein $\mathfrak{g}$–invarianter Teilraum ist. Dann gilt $\varphi'(X)(w) \in W$ für alle $w \in W$. Induktiv erhalten wir $(\varphi'(X))^k(w) \in W$ für alle $k \in \mathbb{N}$. Nun gilt nach Satz 2.8 $\varphi(\exp(X)) = \exp(\varphi'(X))$ für alle $X \in \mathfrak{g}$. Die Exponentialabbildung auf der rechten Seite ist aber hier das Matrizenexponential $\exp : \mathfrak{gl}(V) \to GL(V)$. Damit kann man aber $\varphi(\exp(X))(w)$ als $\sum_{k=0}^{\infty} \frac{1}{k!}(\varphi'(X))^k(w)$ schreiben. Von oben wissen wir, dass alle endlichen Partialsummen in W liegen und da Teilräume von endlichdimensionalen Vektorräumen immer abgeschlossen sind, liegt auch die unendliche Summe in W. Damit folgt aber $\varphi(\exp(X))(w) \in W$ für alle $w \in W$ und $X \in \mathfrak{g}$. Betrachten wir nun $H := \{h \in G : \forall w \in W : \varphi(h)(w) \in W\} \subset G$, dann gilt offensichtlich $e \in H$ und aus $h_1, h_2 \in H$ folgt $h_1h_2 \in H$. Für $h \in H$ ist $\varphi(h)|_W : W \to W$ injektiv also ein linearer Isomorphismus. Damit kann man ein gegebenes Element $w \in W$ als $\varphi(h)(w')$ für $w' \in W$ schreiben und damit gilt $\varphi(h^{-1})(w) = w' \in W$, also $h^{-1} \in H$. Somit ist $H \subset G$ eine Untergruppe, die alle Elemente der Form $\exp(X)$ enthält. Nach Satz 2.9 folgt $H = G$ also ist der Teilraum W auch G–invariant.

Die Aussagen über Unzerlegbarkeit und Irreduzibilität folgen offensichtlich aus der Aussage über invariante Teilräume. □

Wir geben nun ein Beispiel für die Verwendung der Lie Algebra, das auch zeigt, dass man für interessante Matrixgruppen viele irreduzible Darstellungen findet.

Proposition 3.7. *Sei $G := SL(2, \mathbb{R})$, $\mathbb{R}^2$ die Standarddarstellung von G und für $k \geq 1$ betrachten wir die Darstellung $S^k\mathbb{R}^{2*}$ der symmetrischen multilinearen Abbildungen aus Beispiel (4) von 3.6. Dann gilt* $\dim(S^k\mathbb{R}^{2*}) = k + 1$ *und jede der Darstellungen $S^k\mathbb{R}^{2*}$ ist irreduzibel.*

Beweis. Sei $V := \mathbb{R}^2$ und sei $\{e_1, e_2\}$ die Standardbasis. Dann definieren wir für $i = 0, \dots, k$ eine k–lineare Abbildung $\alpha_i : V^k \to \mathbb{R}$ wie folgt. Betrachten wir k Vektoren $v_1, \dots, v_k \in V$, schreiben $v_j = \binom{x_j}{y_j}$ und definieren $\alpha_i(v_1, \dots, v_k) = \sum_J \prod_{j\in J} x_j \prod_{\ell\notin J} y_\ell$, wobei die Summe über alle i–elementigen Teilmengen von $\{1, \dots, k\}$ geht. Insbesondere ist also $\alpha_0(v_1, \dots, v_k) = y_1 \dots y_k$,

$$\alpha_1(v_1, \dots, v_k) = x_1y_2 \dots y_k + y_1x_2y_3 \dots y_k + \dots + y_1 \dots y_{k-1}x_k,$$

und so weiter. Aus dieser Definition folgt sofort, dass jedes α_i eine k–lineare, symmetrische Abbildung definiert. Setzt man in α_i nur Kopien der Basiselemente e_1 und e_2 ein, dann erhält man 1 falls genau i Kopien von e_1 und $k - i$ Kopien von e_2 eingesetzt werden und 0 sonst. Daraus folgt offensichtlich, dass die Abbildungen $\alpha_0, \dots, \alpha_k$ linear unabhängig in S^kV^* sind.

Betrachten wir nun eine beliebige k–lineare Abbildung $\alpha : V^k \to \mathbb{R}$. Dann können wir Basiselemente in α einsetzen und wir betrachten die reellen Zahlen $a_0, \dots, a_k$, die durch $a_i := \alpha(e_1, \dots, e_1, e_2, \dots, e_2)$ definiert sind, wobei wir i mal e_1 und $k-i$ mal e_2 einsetzen. Für Vektoren $v_1, \dots, v_k$ wie oben können wir nun $v_i = x_ie_1 + y_ie_2$ in $\alpha(v_1, \dots, v_k)$ einsetzen und nach Multilinearität expandieren. Das gibt eine große Summe wobei jeder Summand aus einem reellen Faktor mal einem Ausdruck besteht, in dem an jeder Stelle entweder e_1 oder e_2 in α eingesetzt wird. Bezeichnet man mit $J \subset \{1, \dots, k\}$ die Menge jener Eintragungen von α, in denen e_1 steht, dann ist der Vorfaktor genau $\prod_{j\in J} x_j \prod_{\ell\notin J} y_\ell$, während man die verbleibenden Eintragungen wegen der Symmetrie von α umordnen kann und damit gerade a_i erhält. Damit folgt aber $\alpha = \sum_i a_i\alpha_i$, also bilden die α_i eine Basis für den Raum S^kV^*, der somit Dimension $k+1$ hat.

Um die Irreduzibilität zu beweisen, können wir nach Satz 3.7 mit der zugehörigen Darstellung der Lie Algebra $\mathfrak{sl}(2,\mathbb{R})$ arbeiten. Dazu betrachten wir die Basis $\{E,H,F\}$ aus Beispiel 2.8, also $E = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $H = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ und $F = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ und berechnen die Wirkung dieser Elemente auf die Basiselemente α_i. Nach Definition ist $He_1 = e_1$ und $He_2 = -e_2$, daraus folgt sofort, dass $H\cdot\alpha_i$ ein Vielfaches von α_i sein muss, und es genügt $(H\cdot\alpha_i)(e_1,\dots,e_1,e_2,\dots,e_2)$ mit i Kopien von e_1 und $k-i$ Kopien von e_2 zu berechnen. Expandieren wir das nach Formel (3.2) aus Abschnitt 3.6, dann erhalten wir i mal -1 und $k-i$ mal $+1$, also insgesamt $H\cdot\alpha_i = (k-2i)\alpha_i$. Damit sind alle α_i Eigenvektoren für die Wirkung von H und die auftretenden Eigenwerte sind alle verschieden, nämlich k, $k-2$, $\dots$, $-k+2$, $-k$.

Andererseits ist $Ee_1 = 0$ und $Ee_2 = e_1$, woraus wir sofort sehen, dass $E\cdot\alpha_0 = 0$ gilt, während für $i>0$ die Funktion $E\cdot\alpha_i$ ein Vielfaches von α_{i-1} sein muss. Um den Faktor zu bestimmen können wir für $i>0$ wieder $(E\cdot\alpha_i)(e_1,\dots,e_1,e_2,\dots,e_2)$ mit $i-1$ Kopien von e_1 und $k-i+1$ Kopien von e_2 berechnen und erhalten $E\cdot\alpha_i = -(k-i+1)\alpha_{i-1}$. Insbesondere ist der Faktor immer ungleich 0. Analog ist $Fe_1 = e_2$ und $Fe_2 = 0$ und es folgt $F\cdot\alpha_k = 0$ und für $i<k$ muss $F\cdot\alpha_i$ ein Vielfaches von α_{i+1} sein. Für den Faktor erhalten wir analog wie oben $F\cdot\alpha_i = (i+1)\alpha_{i+1}$ und auch dieser Faktor ist immer ungleich Null.

Damit können wir nun die Irreduzibilität direkt beweisen. Sei nämlich $W \subset S^kV^*$ ein Teilraum $\neq \{0\}$, der $\mathfrak{sl}(2,\mathbb{R})$–invariant ist. Dann finden wir eine Element $0 \neq w = \sum_{i=0}^n w_i\alpha_i \in W$, also finden wir einen Index i, sodass $w_i \neq 0$ gilt. Nun benutzen wir, dass für die diagonalisierbare lineare Abbildung $\varphi(h)$ die Projektion auf einen Eigenraum als Polynom in $\varphi(h)$ geschrieben werden kann. Explizit betrachten wir die lineare Abbildung $\pi_i := \prod_{j\neq i}(\varphi(h) - (k-2j)\,\mathrm{id}) : S^kV^* \to S^kV^*$, wobei wir die Reihenfolge der Komposition beliebig wählen dürfen. Damit folgt für $j \neq i$ sofort $\pi_i(\alpha_j) = 0$ während $\pi_i(\alpha_i) = \prod_{j\neq i}((k-2i)-(k-2j))\alpha_i = (2^{k-1}\prod_{j\neq i}(j-i))\alpha_i$ und damit ein nichttriviales Vielfaches von α_i ist. Aus der Invarianz von W folgt sofort, dass $\pi_i(w) \in W$ gilt und nach Konstruktion ist das ein nichttriviales Vielfaches von α_i. Damit folgt $\alpha_i \in W$ und wegen der Invarianz liegen auch $E\cdot\alpha_i$ und $F\cdot\alpha_i$ und damit α_{i-1} und α_{i+1} in W und iterativ erhält man $W = V$. □

Dieses Resultat zeigt viele allgemeine Züge der Darstellungstheorie von halbeinfachen Lie Algebren auf, wobei man dort üblicherweise wegen der besseren Diagonalisierbarkeitsresultate über $\mathbb{C}$ arbeitet. Tatsächlich bilden ja E, F und H auch eine (komplexe) Basis für $\mathfrak{sl}(2,\mathbb{C})$ und die Relationen $[H,E] = 2E$, $[H,F] = -2F$ und $[E,F] = H$, die wir schon aus Beispiel 2.8 kennen, gelten für beide Lie Algebren. Diese bedingen viele der schönen Eigenschaften, die wir im obigen Beweis beobachtet haben. Betrachten wir eine beliebige Darstellung von $\mathfrak{sl}(2,\mathbb{K})$ auf V und sei $v \in V$ ein Eigenvektor für die Wirkung von H, also $H\cdot v = \lambda v$ für $\lambda \in \mathbb{K}$. Dann betrachten wir $E\cdot v$ und aus der Definition einer Darstellung folgt, dass

$$H\cdot(E\cdot v) = E\cdot(H\cdot v) + [H,E]\cdot v = \lambda(E\cdot v) + 2E\cdot v = (\lambda+2)E\cdot v$$

gelten muss. Also ist $E\cdot v$ ein Eigenvektor zum Eigenwert $\lambda+2$ und analog ist $F\cdot v$ ein Eigenvektor zum Eigenwert $\lambda-2$. Schließlich kann man noch Informationen über die Gesamtheit der möglichen Eigenwerte der Wirkung von H bekommen. Wegen $H = [E,F]$ gilt $\varphi(H) = \varphi(E)\circ\varphi(F) - \varphi(F)\circ\varphi(E)$, wobei $\varphi : \mathfrak{sl}(2,\mathbb{K}) \to \mathfrak{gl}(V)$ die Darstellung bezeichnet. Daraus folgt aber $\mathrm{tr}(\varphi(H)) = 0$, also muss für diagonalisierbares $\varphi(H)$ die Summe aller Eigenwerte 0 sein. Über $\mathbb{C}$ gibt es allgemeine Resultate, die sicherstellen, dass $\varphi(H)$ immer diagonalisierbar ist und man kann beweisen, dass jede endlichdimensionale, irreduzible Darstellung von $\mathfrak{sl}(2,\mathbb{C})$ isomorph zu einer der Darstellungen $S^k\mathbb{C}^{2*}$ ist, für die Irreduzibilität wie in Proposition 3.7 folgt.

3.8. Bewegungen und affine Erweiterungen. Zum Abschluss dieses Kapitels beschäftigen wir uns mit einer Klasse von Beispielen von Matrixgruppen die eine Verbindung zwischen der klassischen Geometrie und der Differentialgeometrie über die Theorie der Matrixgruppen darstellt. Der Startpunkt für diese Überlegungen ist die *affine Geometrie*, in der man Geraden, Ebenen und ihre höherdimensionalen Analoga ohne Begriffe von Abstand und Winkel studiert. Ein typisches Resultat der affinen Geometrie ist der Strahlensatz, der keine Aussagen über Längen von Vektoren trifft, sondern nur über Verhältnissen von Längen paralleler Vektoren, die in der affinen Geometrie Sinn machen.

Den richtigen Rahmen dafür bildet ein affiner Raum, den man erhält, indem man in einem Vektorraum (der in diesem Zusammenhang als der "modellierende Vektorraum" bezeichnet wird) "den Ursprung vergisst". Damit kann man einerseits zu zwei Punkten im affinen Raum den Verbindungsvektor als Element des modellierenden Vektorraumes betrachten, andererseits kann man an einen Punkt im affinen Raum einen Vektor aus dem modellierenden Vektorraum "anhängen" und so zu einem anderen Punkt des affinen Raumes gelangen. Die allgemeine Definition eines affinen Raumes erhält man, indem man eine abstrakte Version der offensichtlichen Eigenschaften der beiden entstehenden Operationen verlangt, wir werden diese allgemeine Version aber nicht benötigen.

Für unsere Zwecke wird der Begriff einer *affinen Bewegung* von zentraler Bedeutung sein. Im abstrakten Setting definiert man affine Bewegungen als bijektive Funktionen zwischen affinen Räumen, die affine Geraden auf affine Geraden abbilden. Dann zeigt man, dass diese Bewegungen als Kompositionen von Translationen und linearen Abbildungen geschrieben werden können. Bekannter als die affine Geometrie ist die *euklidische Geometrie* für die der Abstand von zwei Punkten der fundamentale Begriff ist. Um die dazu passenden *euklidischen Räume* zu definieren, betrachtet man einen affinen Raum zusammen mit einem positiv definiten inneren Produkt auf dem modellierenden Vektorraum und definiert den Abstand zwischen zwei Punkten als die Länge des Verbindungsvektors. *Euklidische Bewegungen* sind dann als bijektive Funktionen zwischen euklidischen Räumen definiert, die Distanzen zwischen Punkten bewahren und man zeigt wieder (siehe Satz 9.5 in [**LinAlg**]), dass solche Abbildungen als Komposition von Translationen und orthogonalen linearen Abbildungen geschrieben werden können.

Die Verbindung zu Matrixgruppen ist relativ leicht herzustellen. Wir betrachten $\mathbb{R}^{n+1}$ und darin die affine Hyperebene $E := \{(x,1) : x \in \mathbb{R}^n\}$, die wir als Modell für einen n–dimensionalen affinen Raum verwenden. Dann definieren wir

$$\mathrm{Aff}(n) := \{C \in GL(n+1,\mathbb{R}) : \forall y \in E : Cy \in E\}.$$

Wir können leicht zeigen, dass dies genau die Gruppe der affinen Bewegungen ist:

Proposition 3.8. *(1)* $\mathrm{Aff}(n)$ *ist eine abgeschlossene Untergruppe von* $GL(n+1,\mathbb{R})$ *und damit eine Matrixgruppe.*

(2) Die Gruppe $\mathrm{Aff}(n)$ *besteht genau aus den Matrizen der Blockform* $\left(\begin{smallmatrix} A & b \\ 0 & 1 \end{smallmatrix}\right)$ *mit* $A \in GL(n,\mathbb{R})$ *und* $b \in \mathbb{R}^n$. *Diese Matrix bildet* $\binom{x}{1}$ *auf* $\binom{Ax+b}{1}$ *ab, also wird* $\mathrm{Aff}(n)$ *von Translationen und linearen Abbildungen erzeugt.*

(3) Die Lie Algebra $\mathfrak{aff}(n)$ *von* $\mathrm{Aff}(n)$ *besteht genau aus jenen Matrizen in* $M_{n+1}(\mathbb{R})$*, deren letzte Zeile nur aus Nullen besteht.*

Beweis. Wir behaupten zunächst, dass für jede Matrix $C \in \mathrm{Aff}(n)$ und jeden Vektor $y \in \mathbb{R}^{n+1}$ mit letzter Koordinate $t = y_{n+1}$ auch die letzte Koordinate von Cy gleich t sein muss. Für $t = 1$ folgt das aus der Definition und damit folgt es für $t \neq 0$ aus der Linearität von C. Nun können wir aber den Punkt $y = \binom{x}{0}$ für $x \in \mathbb{R}^n$ als Limes der Folge $y_k := \binom{x}{1/k}$ schreiben. Weil C stetig ist, ist Cy der Limes der Folge Cy_k und wir wissen schon dass die letzte Koordinate von Cy_k gerade $\frac{1}{k}$ ist, und damit folgt die Behauptung.

(1) Aus der Definition folgt sofort, dass $\mathbb{I}$ in $\mathrm{Aff}(n)$ liegt und für $C, D \in \mathrm{Aff}(n)$ auch $CD \in \mathrm{Aff}(n)$ gilt. Für $C \in \mathrm{Aff}(n)$ und $x \in E$ liegt $CC^{-1}x = x$ in E also liegt nach der der Behauptung auch $C^{-1}x$ in E. Damit ist $C^{-1} \in \mathrm{Aff}(n)$ und somit $\mathrm{Aff}(n) \subset GL(n, \mathbb{R})$ eine Untergruppe. Sei $(C_k)_{k\in\mathbb{N}}$ eine Folge von Elementen von $\mathrm{Aff}(n)$, die gegen eine invertierbare Matrix C konvergiert. Dann konvergiert für jedes $y \in E$ die Folge $(C_k y)_{k\in\mathbb{N}}$ gegen Cy und nach Konstruktion liegen die $C_k y$ alle in E. Nun ist E aber klarerweise abgeschlossen in $\mathbb{R}^{n+1}$, also folgt $Cy \in E$ und damit $C \in \mathrm{Aff}(n)$. Damit ist $\mathrm{Aff}(n)$ abgeschlossen in $GL(n + 1, \mathbb{R})$ und somit eine Matrixgruppe.

(2) Aus der Behauptung wissen wir, dass für $C \in \mathrm{Aff}(n)$ und $i = 1, \dots, n$ die letzte Koordinate von Ce_i immer Null sein muss, während Ce_{n+1} letzte Koordinate 1 haben muss. Damit erhalten wir die behauptete Blockform und die Determinante einer Matrix mit dieser Blockform stimmt mit $\det(A)$ überein, also muss $A \in GL(n, \mathbb{R})$ gelten. Eine Matrix dieser Form bildet aber offensichtlich $\binom{x}{1}$ auf $\binom{Ax+b}{1}$ ab und liegt damit in $\mathrm{Aff}(n)$ und damit ist der Rest von (2) offensichtlich.

(3) Ist $c : I \to M_{n+1}(\mathbb{R})$ eine glatte Kurve, die $c(0) = \mathbb{I}$ erfüllt und Werte in $\mathrm{Aff}(n)$ hat, dann ist die letzte Zeile von $c(t)$ konstant, also besteht die letzte Zeile von $c'(0)$ nur aus Nullen. Betrachten wir umgekehrt eine Matrix, deren letzte Zeile nur aus Nullen besteht, dann gilt das gleiche für alle Potenzen dieser Matrix. Mit Hilfe von (2) folgt daraus aber sofort, dass $\exp(tX)$ für alle t in $\mathrm{Aff}(n)$ liegt, was den Beweis vervollständigt. □

Beschreibt man eine Matrix wie in Teil (2) als Paar (A, b) mit $A \in GL(n, \mathbb{R})$ und $b \in \mathbb{R}^n$, dann ist die Multiplikation durch $(A_1, b_1)(A_2, b_2) = (A_1A_2, b_1 + A_1b_2)$ gegeben. Ähnlich kann man Elemente in $\mathfrak{aff}(n)$ als Paare (X, c) mit $X \in \mathfrak{gl}(n, \mathbb{R})$ und $c \in \mathbb{R}^n$ schreiben, indem man die Nullzeile unten weglässt und dann die letzte Spalte abtrennt. In diesem Bild bekommt die Lie Klammer die Form

$$[(X_1, c_1), (X_2, c_2)] = ([X_1, X_2], X_1c_2 - X_2c_1).$$

Eine allgemeine Version davon bilden die Konstruktionen von *semi-direkten Produkten.* Für eine allgemeine Matrixgruppe G und eine Darstellung V von G kann man analog wie oben eine Multiplikation definieren, die $G \times V$ zu einer Gruppe macht. Mit etwas Mühe kann man auch zeigen, dass dies eine Matrixgruppe ist (im Setting von Lie Gruppen ist das leichter). Für einer Lie Algebra $\mathfrak{g}$ und eine Darstellung V von $\mathfrak{g}$ kann man analog $\mathfrak{g} \times V$ zu einer Lie Algebra machen. Insbesondere ist die Lie Algebra des semi-direkten Produkts von G und V genau das semi-direkte Produkt der Lie Algebra $\mathfrak{g}$ von G mit V.

Der Übergang zu euklidischen Bewegungen ist nun relativ einfach. Man definiert $\mathrm{Euc}(n)$ einfach als die Menge aller jener Matrizen $\left(\begin{smallmatrix} A & b \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Aff}(n)$, für die $A \in O(n)$ (oder – je nach Konvention – $A \in SO(n)$) gilt. Aus der Beschreibung der Multiplikation von oben folgt sofort, dass diese Matrizen eine Abgeschlossene Untergruppe von $\mathrm{Aff}(n)$ bilden, also ist auch $\mathrm{Euc}(n)$ eine Matrixgruppe. Im Bild der Paare von oben ist die Lie Algebra dieser Gruppe $\mathfrak{euc}(n) = \{(X, c) \in \mathfrak{aff}(n) : X^t = -X\}$. Die geometrische

Interpretation ergibt sich dann wie folgt. Für eine affine Bewegung $f(x) = Ax + b$ sieht man sofort, dass die Ableitung durch $Df(x)(v) = Av$ gegeben ist. Konzeptuell gesehen kann man die affine Ebene $\{z_{n+1} = 1\}$ als Teilmannigfaltigkeit von $\mathbb{R}^{n+1}$ betrachten und den Tangentialraum in jedem Punkt mit dem modellierenden Vektorraum $\mathbb{R}^n$ identifizieren. Der Übergang zu euklidischen Bewegungen erfolgt dann wie schon erwähnt dadurch, dass man diesen $\mathbb{R}^n$ mit dem üblichen inneren Produkt ausstattet und so ein inneres Produkt auf jedem Tangentialraum erhält. Die euklidischen Bewegungen sind dann dadurch charakterisiert, dass ihre Ableitung in jedem Punkt orthogonal bezüglich der inneren Produkte auf den beteiligten Tangentialräumen ist. In diesem Sinne wird die euklidische Geometrie zum Modellfall für die sogenannte Riemann'sche Geometrie, siehe Abschnitte 1.1 und 1.2 von [**Riem**].

Natürlich liefert nun jede abgeschlossene Untergruppe $G \subset GL(n, \mathbb{R})$ in ähnlicher Weise eine abgeschlossene Untergruppe von $\mathrm{Aff}(n)$, die man als die *affine Erweiterung* von G bezeichnet. Natürlich ist diese Gruppe genau das semi-direkte Produkt von G mit der Standarddarstellung $\mathbb{R}^n$ die man aus der Inklusion $G \to GL(n, \mathbb{R})$ erhält. Somit ist die Lie Algebra der affinen Erweiterung genau das semi-direkte Produkt der Lie Algebra $\mathfrak{g}$ von G mit $\mathbb{R}^n$. Analog zum Fall der euklidischen Gruppe kann man diese Konstruktion so verstehen, dass die Gruppe G eine "geometrische Strukturen" auf dem affinen Raum definiert. Dieser Standpunkt ist fundamental für den entsprechenden Teil der Differentialgeometrie.

Ein Beispiel einer affinen Erweiterung, das in der Physik große Bedeutung hat ist die *Poincaré Gruppe*. Dazu betrachtet man einen affinen Raum der Dimension 4 und stattet den modellierenden Vektorraum $\mathbb{R}^4$ mit einer nicht-degenerierten symmetrischen Bilinearform der Signatur $(3, 1)$ (oder, je nach Konvention, der Signatur $(1, 3)$) aus. Bezeichnet man die Koordinaten mit $(t, x) = (t, x_1, x_2, x_3)$ dann kann man die Bilinearform durch

$$\langle (t, x), (s, y) \rangle := -ts + \textstyle\sum_{i=1}^{3} x_i y_i$$

definieren. Die orthogonale Gruppe dieser Bilinearform wird mit $O(3, 1)$ bezeichnet, in der Physik heißt sie die *Lorentzgruppe* und ihre Elemente nennt man Lorentztransformationen. Die affine Erweiterung der Lorentzgruppe nennt man die Poincaré Gruppe.

Der entsprechende affine Raum heißt in der Physik der *Minkowski Raum* (obwohl in der Physik meist nicht zwischen dem affinen Raum und dem unterliegenden Vektorraum unterschieden wird) und wird als *Raumzeit* also gemeinsames Modell für Raum und Zeit betrachtet. Die wesentliche Rolle der Bilinearform dabei ist, dass man bei Vektoren $v \in \mathbb{R}^4$ zwischen den Fällen $\langle v, v \rangle > 0$, $\langle v, v \rangle = 0$ und $\langle v, v \rangle < 0$ unterscheiden kann, die als "raumartig", "lichtartig" und "zeitartig" bezeichnet werden. Dementsprechend kann man für zwei Punkte im Minkowski Raum sehen, in welche dieser Klassen der Verbindungsvektor zwischen den Punkten fällt.

Physikalisch gesehen entsprechen die Geraden in lichtartigen Richtungen genau den raumzeitlichen Bahnen von Lichtstrahlen (was bedeutet, dass wir die Zeit als Länge messen, indem wir die Lichtgeschwindigkeit gleich 1 setzen). Raumzeitliche Punkte mit raumartigem Verbindungsvektor können einander nicht beeinflussen, weil eine Bewegung von einem Punkt zum anderen nur mit Überlichtgeschwindigkeit möglich wäre, und so weiter. In diesem Sinne ist die spezielle Relativitätstheorie der Physik genau das Analogon der euklidischen Geometrie, wobei aber der euklidische Raum durch den Minkowski Raum ersetzt wird. Tatsächlich kann man die Grundeffekte der speziellen Relativitätstheorie wie Zeitdialatation und Längenkontraktion relativ leicht aus der Struktur einiger grundlegender Lorentztransformationen ablesen.

KAPITEL 4

Exkurs: Homogene Räume und abstrakte Mannigfaltigkeiten

Zum Abschluss der Vorlesung werden wir einen kurzen Blick auf weiterführende Aspekte der Theorie der Matrixgruppen und der Analysis auf Mannigfaltigkeiten werfen. Zur Motivation betrachten wir die Frage von Quotienten von Matrixgruppen, die man sowohl im Sinne von Quotientengruppen als auch allgemeiner im Sinne von Räumen von Nebenklassen verstehen kann. Dabei stellt sich leicht heraus, dass nur im Fall einer einer abgeschlossenen (normalen) Untergruppe Hoffnung auf einen "schönen" Quotienten besteht. In diesem Fall kann man auch relativ leicht sehen, dass der Quotient analytisch immer schön aussehen sollte, nur ist gar nicht klar, wie man ihn als Teilmenge eine endlichdimensionalen Vektorraumes realisieren sollte. Die einfachste Lösung für dieses Problem ist der Übergang zu einem allgemeineren Mannigfaltigkeitsbegriff.

4.1. Quotienten von Matrixgruppen. Sei G eine Matrixgruppe und $H \subset G$ eine normale Untergruppe. Dann können wir G/H als die Menge der linken Nebenklassen $gH = \{gh : h \in H\}$ definieren und darauf eine Multiplikation durch $(g_1H)(g_2H) := (g_1g_2)H$ definieren. Aus der Algebra ist bekannt, dass diese Multiplikation die Menge G/H zu einer Gruppe und die natürliche Abbildung $\pi : G \to G/H$, $\pi(g) := gH$ zu einem Gruppenhomomorphismus macht. Nun stellt sich natürlich die Frage, ob man G/H so zu einer Matrixgruppe machen kann, dass der Homomorphismus π glatt ist.

In manchen Fällen wissen wir schon, dass das möglich ist. Nach Satz 2.10 kann man für einen surjektiven glatten Homomorphismus $\varphi : G \to H$ zwischen Matrixgruppen den Quotienten $G/\operatorname{Ker}(\varphi)$ mit H identifizieren. Explizit haben wir das etwa in den Abschnitten 3.4 und 3.5 verwendet um $SO(3)$ mit $SU(2)/\{\pm\mathbb{I}\}$ und $SO(4)$ mit $(SU(2) \times SU(2))/\mathbb{Z}_2$ zu identifizieren. Schon das erste dieser Beispiele zeigt uns sowohl das Potential solcher Konstruktionen, als auch Schwierigkeiten, die wir erwarten müssen. Als Raum können wir $SU(2)$ mit der Einheitssphäre in $\mathbb{C}^2 = \mathbb{R}^4$ identifizieren und aus Satz 3.4 wissen wir, dass man daraus $SO(3)$ erhält, in dem man jeden Punkt x mit dem antipodalen Punkt $-x$ identifiziert. Diese Punkte x und $-x$ sind aber genau die beiden Schnittpunkte eines 1-dimensionalen reellen Teilraumes von $\mathbb{R}^4$ mit S^3. Damit kann man den Quotienten $SO(3)$ als den Raum aller eindimensionalen Teilräume von $\mathbb{R}^4$, also den projektiven Raum $\mathbb{R}P^3$ betrachten. Also treten interessante Räume als Quotienten von Matrixgruppen auf.

Tatsächlich ist hier die Tatsache, dass $\mathbb{R}P^3$ zu einer Gruppe gemacht werden kann, nicht so bedeutsam. Man kann allgemeiner sehen, dass viele interessante Mengen mit Mengen von Nebenklassen von Gruppen nach einer Untergruppe identifiziert werden können. Einen konzeptuellen Zugang dazu liefern die aus der Algebra bekannten Gruppenwirkungen. Für eine Menge X und eine Gruppe G betrachten wir eine *Linkswirkung* von G auf X, also eine Funktion $G \times X \to X$, die wir als $(g, x) \mapsto g \cdot x$ schreiben. Dabei verlangen wir, dass $e \cdot x = x$ und $gh \cdot x = g \cdot (h \cdot x)$ für das neutrale Element $e \in G$ und beliebige $g, h \in G$ und $x \in X$ gelten. Daraus ergeben sich die folgenden Definitionen in natürlicher Weise:

Definition 4.1. Betrachten wir eine Linkswirkung $(g,x) \mapsto g \cdot x$ einer Gruppe G auf einer Menge X und sei $x_0 \in X$ ein fixer Punkt.

(1) Der *Orbit* oder die *Bahn von* x_0 ist die Teilmenge $G \cdot x_0 := \{g \cdot x_0 : g \in G\} \subset X$.

(2) Die *Standgruppe* oder die *Isotropieuntergruppe* von x_0 ist die Teilmenge $G_{x_0} := \{g \in G : g \cdot x_0 = x_0\} \subset G$.

Die folgenden Tatsachen sind vermutlich aus der Algebra bekannt:

Lemma 4.1. *Für jede Linkswirkung $G \times X \to X$ und jeden Punkt $x_0 \in X$ gilt:*

(1) Die Standgruppe G_{x_0} ist eine Untergruppe von G.

(2) Die Funktion $g \mapsto g \cdot x_0$ induziert eine Bijektion zwischen dem Raum G/G_{x_0} der linken Nebenklassen und dem Orbit $G \cdot x_0$.

Beweis. (1) folgt sofort aus der Definition einer Wirkung.

(2) Betrachten wir die Funktion $p : G \to X$, die durch $p(g) := g \cdot x_0$ gegeben ist. Sind $g_1, g_2 \in G$ so, dass $g_1 \cdot x_0 = g_2 \cdot x_0$, dann wirken wir auf beide Seiten dieser Gleichung mit g_1^{-1}. Auf der linken Seite erhalten wir $e \cdot x_0 = x_0$, auf der rechten Seite $g_1^{-1} g_2 \cdot x_0$, also folgt $g_1^{-1} g_2 \in G_{x_0}$ und damit $g_1 G_{x_0} = g_2 G_{x_0}$. Ist umgekehrt $g_1 G_{x_0} = g_2 G_{x_0}$, dann gibt es ein Element $h \in G_{x_0}$, sodass $g_2 = g_1 h$ gilt. Dann ist aber $g_2 \cdot x_0 = g_1 \cdot (h \cdot x_0) = g_1 \cdot x_0$. Insgesamt erhalten wir, dass $p(g_1) = p(g_2)$ äquivalent zu $g_1 G_{x_0} = g_2 G_{x_0}$ ist und damit steigt p zu einer injektiven Funktion $G/G_{x_0} \to X$ ab. Aber nach Definition gilt $p(G) = G \cdot x_0$, also folgt die Behauptung. □

Das liefert schon mit einfachen Matrixgruppen sehr interessante Beispiele, bei denen a priori absolut nicht klar ist, ob sie "schöne" Räume bilden.

Beispiel 4.1. Für $k < n \in \mathbb{R}^n$ betrachten wir die Menge $Gr(k, \mathbb{R}^n)$ aller k–dimensionalen Teilräume von $\mathbb{R}^n$. Ist $V \subset \mathbb{R}^n$ ein k-dimensionaler Teilraum und $f : \mathbb{R}^n \to \mathbb{R}^n$ ein linearer Isomorphismus, dann ist natürlich $f(V) \subset \mathbb{R}^n$ ebenfalls ein k-dimensionaler Teilraum. Damit folgt sofort, dass $(A, V) \mapsto A(V)$ eine Linkswirkung von $G := GL(n, \mathbb{R})$ auf $Gr(k, \mathbb{R}^n)$ definiert. Betrachten wir den k-dimensionalen Teilraum $\mathbb{R}^k \subset \mathbb{R}^n$, dann folgt aus bekannten Resultaten der linearen Algebra sofort, dass die Bahn von $\mathbb{R}^k$ die ganze Menge $Gr(k, \mathbb{R}^n)$ ist: Für einen gegebenen Teilraum $V \in Gr(k, \mathbb{R}^n)$ muss man nur einfach eine Basis $v_1, \dots, v_k$ für V wählen und durch Vektoren $v_{k+1}, \dots, v_n$ zu einer Basis von $\mathbb{R}^n$ erweitern. Die Matrix A mit den Spaltenvektoren $v_1, \dots, v_n$ ist dann invertierbar und bildet die Basis $\{e_1, \dots, e_k\}$ für $\mathbb{R}^k$ auf die Basis $\{v_1, \dots, v_k\}$ von V ab, also gilt $A(\mathbb{R}^k) = V$. Damit können wir die Menge $Gr(k, \mathbb{R}^n)$ mit der Menge $G/G_{\mathbb{R}^k}$ von linken Nebenklassen identifizieren.

Man kann diese Wirkung natürlich auf jede Untergruppe von $GL(n, \mathbb{R})$ einschränken. Betrachten wir etwa $SL(n, \mathbb{R})$, $O(n)$ oder $SO(n)$, dann kann unter Benutzung von passenden Basen wie oben zeigen, dass die Bahn von $\mathbb{R}^k \subset \mathbb{R}^n$ wiederum ganz $Gr(k, \mathbb{R}^n)$ ist, die daher mit einer Menge von Nebenklassen der entsprechenden Gruppe identifiziert werden kann.

Für das Beispiel des projektiven Raumes $\mathbb{R}P^3$ ist aus der Konstruktion als $S^3/\sim$ klar, dass $\mathbb{R}P^3$ lokal schön aussieht. Betrachtet man etwa eine offene Hemisphäre in S^3 dann enthält diese keine antipodalen Punkte und wird damit bijektiv auf eine Teilmengen von $\mathbb{R}P^3$ abgebildet. Um $\mathbb{R}P^3$ zu einer Teilmannigfaltigkeit zu machen, müsste man diesen Raum aber erst einmal als Teilmenge eines $\mathbb{R}^N$ realisieren und es ist nicht klar, woher so eine Realisierung kommen könnte. Die Matrixgruppe $SO(3) \subset \mathbb{R}^9$ liefert so eine Realisierung und nachdem eine positiv orientierte Orthonormalbasis von $\mathbb{R}^3$ durch ihre ersten beiden Vektoren eindeutig bestimmt ist, kann man sich auf die ersten beiden

Spalten beschränken, was eine Realisierung in $\mathbb{R}^6$ liefert. Ein Zusammenhang zu der Realisierung von $SU(2)$ in $\mathbb{R}^4$ ist aber nicht erkenntlich. Tatsächlich kann man zeigen, dass $\mathbb{R}P^3$ als Teilmenge von $\mathbb{R}^5$ realisiert werden kann, nicht aber als Teilmenge von $\mathbb{R}^n$ für $n \leq 4$.

Im allgemeinen Fall einer Matrixgruppe G und einer Untergruppe $H \subset G$ kann man leicht einen ersten Schritt zu einer Beschreibung der Menge G/H von Nebenklassen machen. Nach Definition ist G/H der Quotient von G nach einer Äquivalenzrelation (nämlich $g_1 \sim g_2$ genau dann, wenn $g_1^{-1}g_2 \in H$). Als Matrixgruppe trägt aber G eine Topologie und diese liefert in natürlicher Weise eine Topologie auf G/H, die *Quotiententopologie*. Man definiert eine Teilmenge $U \subset G/H$ als offen (bzw. abgeschlossen) genau dann, wenn $\pi^{-1}(U) \subset G$ offen (bzw. abgeschlossen) ist, womit insbesondere die Funktion $\pi : G \to G/H$ stetig ist. Diese Topologie hat die schöne Eigenschaft, dass für jeden topologischen Raum X eine Funktion $f : G/H \to X$ genau dann stetig ist, wenn $f \circ \pi : G \to X$ stetig ist. Daher wird sie auch als *finale Topologie bezüglich* π bezeichnet. Eine Minimalvoraussetzung für einen schönen Quotienten ist, dass diese Topologie Hausdorff ist. Insbesondere müssen dann einpunktige Teilmengen von G/H abgeschlossen sein, was impliziert, dass $H = \pi^{-1}(eH)$ abgeschlossen in G sein muss. Damit ist aber dann H selbst eine Matrixgruppe und die Lie Algebra $\mathfrak{h}$ von H ist ein Teilraum der Lie Algebra $\mathfrak{g}$ von G. Damit kann man hoffen, den Raum G/H lokal durch ein Komplement von $\mathfrak{h}$ in $\mathfrak{g}$ zu beschreiben (und wir werden sehen, dass das funktioniert). Eine Realisierung von G/H als Menge von Matrizen (oder als Teilmenge eines $\mathbb{R}^N$) ist aber nicht in Sicht.

4.2. Von Teilmannigfaltigkeiten zu abstrakten Mannigfaltigkeiten. Man kann die oben angesprochenen Probleme vermeiden, indem man den Begriff der Teilmannigfaltigkeit verallgemeinert, was natürlich auch in vielen anderen Bereichen nützlich ist. Dazu können wir zunächst bemerken, dass wir in der Definition der Teilmannigfaltigkeit eigentlich wesentlich mehr gefordert hatten, als tatsächlich benötigt wird. In der Definition einer k–dimensionalen Teilmannigfaltigkeit $M \subset \mathbb{R}^n$ hatten wir ja verlangt dass es für jeden Punkt $x \in M$ offene Teilmengen $U, V \subset \mathbb{R}^n$ mit $x \in U$ und einen Diffeomorphismus $\Phi : U \to V$ gibt, der sich zu einer Bijektion zwischen $M \cap U$ und $\mathbb{R}^k \cap V$ einschränkt. Zerlegt man den Zielraum $\mathbb{R}^n$ als $\mathbb{R}^k \times \mathbb{R}^{n-k}$, dann kann man Φ in Komponenten (Φ_1, Φ_2) zerlegen. Die Komponente $\Phi_2 : U \to \mathbb{R}^{n-k}$ liefert eine lokale Realisierung von M als "reguläre Nullstellenmenge", die ausreicht um M zu einer Teilmannigfaltigkeit zu machen (siehe Übungen). Analog kann man die Inverse $\Psi := \Phi^{-1}$ auf die beiden Faktoren im Produkt einschränken und wir werden zeigen, dass man die Teilmannigfaltigkeitsstruktur aus der Einschränkung von Ψ auf $\mathbb{R}^k$ zurückgewinnen kann.

Definition 4.2. Sei $M \subset \mathbb{R}^n$ eine Teilmenge mit der von $\mathbb{R}^n$ induzierten Topologie. Man sagt, dass *M lokal reguläre k-dimensionale Parametrisierungen besitzt*, wenn folgendes gilt: Für jeden Punkt $x \in M$ gibt es offene Teilmengen $U \subset \mathbb{R}^n$ mit $x \in U$ und $V \subset \mathbb{R}^k$ und eine glatte Funktion $\psi : V \to U$, die einen Homöomorphismus von V auf $M \cap U$ definiert, sodass für jedes $y \in V$ die Ableitung $D\psi(y) : \mathbb{R}^k \to \mathbb{R}^n$ injektiv ist.

Proposition 4.2. *(1) Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine k-dimensionale Teilmannigfaltigkeit, wenn sie lokal reguläre k–dimensionale Parametrisierungen besitzt. Für so eine lokale Parametrisierung $\psi : V \to U$ und $y \in V$ gilt $T_{\psi(y)}M = D\psi(y)(\mathbb{R}^k) \subset \mathbb{R}^n$.*

(2) Seien $M \subset \mathbb{R}^n$ und $N \subset \mathbb{R}^m$ Teilmannigfaltigkeiten der Dimension k und ℓ und sei $f : M \to N$ eine Funktion. Dann ist f genau dann glatt, wenn f stetig ist und für jedes $x \in M$ lokale reguläre Parametrisierungen $\psi_1 : V_1 \to U_1$ für M mit $x \in U_1$ und $\psi_2 : V_2 \to U_2$ für N mit $f(x) \in U_2$ gibt, sodass $\psi_2^{-1} \circ f \circ \psi_1$ eine glatte Funktion im Sinne der Analysis zwischen den offenen Teilmengen $\psi_1^{-1}(f^{-1}(U_2) \cap U_1) \subset \mathbb{R}^k$ und $\psi_2(U_2) \subset \mathbb{R}^\ell$ definiert.

Beweis. (1) Ist M eine k–dimensionale Teilmannigfaltigkeit, dann finden wir für $x \in M$ offene Teilmengen $U, \tilde{V} \subset \mathbb{R}^n$ mit $x \in U$ und einen Diffeomorphismus $\Phi : U \to \tilde{V}$ der sich zu einer Bijektion zwischen $M \cap U$ und $V := \mathbb{R}^k \cap \tilde{V}$ einschränkt. Dann ist $\psi := \Phi^{-1}|_V : V \to M \cap U$ natürlich glatt als Einschränkung der glatten Funktion Φ^{-1} und damit insbesondere stetig und bijektiv nach Konstruktion. Nach Konstruktion ist $\psi^{-1} = \Phi|_{M\cap U}$, was als Einschränkung der stetigen Funktion Φ ebenfalls stetig ist, also ist ψ ein Homöomorphismus. Schließlich ist für $y \in V$ natürlich $D\psi(y) = D\Phi^{-1}(y)|_{\mathbb{R}^k}$ und weil $D\Phi^{-1}(y)$ ein linearer Isomorphismus ist, ist $D\psi(y)$ injektiv, also ist ψ eine lokale reguläre Parametrisierung für M. Wir bemerken auch, dass das Bild $D\psi(y)(\mathbb{R}^k)$ in diesem Fall nach dem Beweis von Satz 2.6 mit $T_{\psi(y)}M$ übereinstimmt.

Nehmen wir also umgekehrt an, dass $M \subset \mathbb{R}^n$ lokal reguläre Parametrisierungen besitzt, betrachten $x \in M$ und eine Parametrisierung $\psi : V \to M \cap U$ mit $x \in U$. Damit ist $x = \psi(y_0)$ für einen eindeutigen Punkt $y_0 \in V$ und nach Voraussetzung ist $E := D\psi(y_0)(\mathbb{R}^k)$ ein k-dimensionaler Teilraum von $\mathbb{R}^n$. Wir wählen einen komplementären Teilraum $F \subset \mathbb{R}^n$ und identifizieren ihn mit $\mathbb{R}^{n-k}$, womit wir auch $\mathbb{R}^k \times F$ mit $\mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n$ identifizieren können. Darin ist $V \times F$ eine offene Teilmenge und wir betrachten die Funktion $\Psi : V \times F \to \mathbb{R}^n$, die durch $\Psi(y, v) := \psi(y) + v$ definiert ist. Insbesondere gilt $\Psi(y_0, 0) = x$ und $D\Psi(y_0, 0)(w, v) = D\psi(y_0)(w) + v$ und der erste Summand liegt in E und der zweite in F. Damit folgt aber aus $D\Psi(y_0, 0)(w, v) = 0$ sofort $D\psi(y_0)(w) = 0$ und $v = 0$, also auch $w = 0$. Also ist $D\Psi(y_0, 0)$ injektiv und damit ein linearer Isomorphismus.

Nach dem inversen Funktionensatz (Satz 1.1) finden wir eine offene Umgebung von $(y_0, 0)$ auf der sich Ψ zu einem Diffeomorphismus einschränkt und wir dürfen annehmen, dass diese von der Form $V_1 \times V_2$ für offene Umgebungen V_1 von y_0 in V und V_2 von 0 in F ist. Da $\psi : V \to M \cap U$ ein Homöomorphismus ist, ist $\psi(V_1)$ offen in $M \cap U$, also finden wir eine offene Teilmenge $W \subset \mathbb{R}^n$, sodass $W \cap M = \psi(V_1)$ gilt. Nun ist auch $\tilde{U} := W \cap \Psi(V_1 \times V_2)$ offen und damit auch $\tilde{V} := \Psi^{-1}(\tilde{U})$ und $\Psi : \tilde{V} \to \tilde{U}$ ist ein Diffeomorphismus. Nach Konstruktion muss eine Punkt $z \in \tilde{U} \cap M$ schon in $W \cap M = \psi(V_1)$ liegen, also gibt es einen Punkt $y \in V_1$ mit $z = \psi(y) = \Psi(y, 0)$. Damit gilt aber $(y, 0) \in \tilde{U}$ und $z = \Psi(y, 0)$ also $\Psi^{-1}(z) \in \mathbb{R}^k \subset \mathbb{R}^n$ und damit erfüllt $\Phi = \Psi^{-1} : \tilde{U} \to \tilde{V}$ alle Bedingungen von Definition 2.5. Da $\psi = \Psi|_{\mathbb{R}^k}$ ist, haben wir oben schon gesehen, dass $T_xM = D\psi(y_0)(\mathbb{R}^k)$ gilt.

(2) Nach Definition ist für eine lokale Parametrisierung $\psi_1 : V_1 \to U_1 \subset M$ die Funktion ψ glatt als Funktion $V_1 \to M$. Andererseits sehen wir aus dem Beweis von Teil (1) dass wir für $\psi_2 : V_2 \to U_2 \subset N$ die inverse Funktion $\psi_2^{-1} : U_2 \to V_2$ als Einschränkung einer lokalen Trivialisierung $\Psi : \tilde{U}_2 \to \tilde{V}_2$ auf $U_2 = \tilde{U}_2 \cap N$ schreiben können. Ist $f : M \to N$ glatt, dann kann man es lokal als Einschränkung $\tilde{f}|_M$ schreiben, wobei $\tilde{f}$ glatt im Sinne der Analysis und somit stetig ist. Daher ist auch f stetig, also ist $f^{-1}(U_2)$ offen in M, also $f^{-1}(U_2) \cap U_1$ offen in U_1. Da ψ_1 stetig ist, ist auch $\psi_1^{-1}(f^{-1}(U_2) \cap U_1)$ offen. Natürlich ist $\psi_2^{-1} \circ f \circ \psi_1$ auf dieser Menge definiert und bildet sie nach $\psi_2(U_2)$ ab. Aus dieser Beschreibung, der Glattheit von f und Satz 2.6 folgt sofort die Glattheit von $\psi_2^{-1} \circ f \circ \psi_1$ im Sinne der Analysis.

Die umgekehrte Richtung ist ähnlich einfach: Nehmen wir an, dass f stetig ist und wir für $x \in M$ lokale Parametrisierungen ψ_1 und ψ_2 gefunden haben, sodass $\psi_2^{-1} \circ f \circ \psi_1 : \psi_1^{-1}(U_1 \cap f^{-1}(U_2)) \to V_2$ glatt im Sinne der Analysis ist. Nach Definition ist auch $\psi_2 : V_2 \to U_2$ glatt in diesem Sinn, also ist $f \circ \psi_1 = \psi_2 \circ \psi_2^{-1} \circ f \circ \psi_1$ glatt als Funktion von der offenen Teilmenge $\psi_1^{-1}(U_1 \cap f^{-1}(U_2)) \subset \mathbb{R}^k$ nach $\mathbb{R}^m$. Nach dem Beweis von (1) finden wir offene Teilmengen $\tilde{U}_1 \in \mathbb{R}^n$ mit $x \in \tilde{U}_1$ und $\tilde{V}_1 \in \mathbb{R}^k \times \mathbb{R}^{n-k}$ und einen Diffeomorphismus $\Psi_1 : \tilde{V} \to \tilde{U}$, sodass $\Psi_1(y, 0) = \psi_1(y)$ gilt. Da $f \circ \psi_1$ glatt im Sinne der Analysis ist, definiert auch $F(y, v) = f(\psi_1(y))$ eine glatte Funktion auf einer offenen Umgebung von $\psi_1^{-1}(x)$ und somit ist $\tilde{f} = F \circ \Psi_1^{-1}$ eine glatte Funktion, die auf einer offenen Umgebung von x in $\mathbb{R}^n$ definiert ist und die auf dem Schnitt mit M nach Konstruktion mit $f \circ \psi_1 \circ \psi_1^{-1} = f$ übereinstimmt. Da der Punkt x beliebig war ist f glatt im Sinn von Definition 2.5. □

Die Bedingung, dass eine lokale Parametrisierung ein Homöomorphismus sein muss, kommt vielleicht etwas überraschend, es ist aber leicht einzusehen, dass sie nötig ist. Betrachten wir etwa die Teilmenge $M \subset \mathbb{R}^2$, die aus der x–Achse und der positiven y–Achse (also $\{(0, t) : t > 0\}$) besteht. Diese ist natürlich keine Teilmannigfaltigkeit, weil es keine lokale Trivialisierung um $0 \in M$ geben kann. Andererseits definieren natürlich $t \mapsto (t, 0)$ und $t \mapsto (0, t)$ glatte Funktionen $\psi_1 : \mathbb{R} \to \mathbb{R}^2$ und $\psi_2 : (0, \infty] \to \mathbb{R}^2$ mit injektiven Ableitungen. Gemeinsam erfüllen diese beiden Funktionen alle Bedingungen von Definition 2.5, abgesehen davon, dass ψ_1 kein Homöomorphismus auf den Schnitt von M mit einer offenen Teilmenge von $\mathbb{R}^2$ sein kann. (Insbesondere ist $\psi_1(\mathbb{R}) \subset M$ nicht offen in der Teilraumtopologie.)

Die Beschreibung glatter Funktionen aus Proposition 4.2 bildet nicht nur die Überleitung zum allgemeinen Konzept von Mannigfaltigkeiten, sie impliziert auch leicht, dass der inverse Funktionensatz für Teilmannigfaltigkeiten von $\mathbb{R}^n$ gilt:

Korollar 4.2 (Inverser Funktionensatz für Teilmannigfaltigkeiten). *Seien $M \subset \mathbb{R}^n$ und $N \subset \mathbb{R}^m$ Teilmannigfaltigkeiten und sei $f : M \to N$ eine glatte Funktion. Ist $x \in M$ ein Punkt, sodass $T_x f : T_x M \to T_{f(x)} N$ ein linearer Isomorphismus ist (was nur möglich ist, wenn M und N die gleiche Dimension haben), dann gibt es eine offene Umgebung U von x in M, sodass $f(U) \subset N$ offen ist, $f|_U : U \to f(U)$ bijektiv ist und die Inverse dazu ebenfalls glatt ist.*

Beweis. Wählen wir lokale Parametrisierungen $\psi_1 : V_1 \to U_1$ für M mit $x \in U_1$ und $\psi_2 : V_2 \to U_2$ für N mit $f(x) \in U_2$. Indem wir U_1 durch $U_1 \cap f^{-1}(U_2)$ und V_1 durch das Urbild dieser offenen Menge unter ψ_1 ersetzen, dürfen wir annehmen, dass $U_1 \subset f^{-1}(U_2)$ gilt. Dann folgt aus dem Beweis von Proposition 4.2, dass $g := \psi_2^{-1} \circ f \circ \psi_1$ glatt als Funktion $V_1 \to V_2$ im Sinne der Analysis ist. Wie im Beweis der Proposition sehen wir auch, dass die glatte Funktion $\psi_2 \circ g : V_1 \to \mathbb{R}^m$ mit $f \circ \psi_1$ übereinstimmt. Setzen wir $y = \psi_1^{-1}(x)$, dann gilt $\psi_2(g(y)) = f(x)$ und $D(\psi_2 \circ g)(y) = D\psi_2(f(x)) \circ Dg(y)$. Schreiben wir andererseits f als Einschränkung einer glatten Funktion $\tilde{f}$ wie in Definition 2.5, dann können wir $D(f \circ \psi_1)(y) = D\tilde{f}(x) \circ D\psi_1(y)$ schreiben. Aus Proposition 4.2 wissen wir, dass $D\psi_1(y)$ ein linearer Isomorphismus $\mathbb{R}^k \to T_x M$ ist und auf diesem Teilraum stimmt $D\tilde{f}(x)$ mit $T_x f$ überein. Nach Voraussetzung ist $T_x f$ ein linearer Isomorphismus, also ist $D(f \circ \psi_1)(y) = D(\psi_2 \circ g)(y)$ ein linearer Isomorphismus von $\mathbb{R}^k$ auf den Teilraum $T_{f(x)} N \subset \mathbb{R}^m$. Da aber $D\psi_2(f(x)) : \mathbb{R}^k \to T_{f(x)} N$ ein linearer Isomorphismus ist, folgern wir, dass $Dg(y) : \mathbb{R}^k \to \mathbb{R}^k$ ein linearer Isomorphismus ist.

Damit können wir aber den gewöhnlichen inversen Funktionensatz (Satz 1.1) auf die Funktion $g : V_1 \to V_2$ im Punkt y anwenden. Das liefert eine offene Teilmenge

$W_1 \subset V_1$, sodass $W_2 := g(W_1) \subset V_2$ offen ist und sich g zu einem Diffeomorphismus $W_1 \to W_2$ einschränkt. Nun setzen wir $U := \psi_1(W_1)$ und $V := \psi_2(g(W_1))$. Weil die ψ_i Homöomorphismen sind, sind das offene Teilmengen in M bzw. N und natürlich sind $\psi_1 : W_1 \to U$ und $\psi_2 : W_2 \to V$ bijektiv. Aus der Konstruktion folgt sofort, dass sich die Funktion $f = \psi_2 \circ g \circ \psi_1^{-1}$ zu einer Bijektion $U \to f(U) = V$ einschränkt. Außerdem folgt auch sofort, dass $\psi_1^{-1} \circ (f|_U)^{-1} \circ \psi_2$ invers zu $g|_{W_1}$ und damit glatt ist. Nach Proposition 4.2 folgt daraus aber, dass auch $(f|_U)^{-1}$ eine glatte Funktion ist. □

4.3. Abstrakte Mannigfaltigkeiten. Die Überlegungen aus Abschnitt 4.2 führen nun in natürlicher Weise zu der Idee eine abstrakte Version von lokalen Parametrisierungen zu verwenden um auf allgemeineren Mengen sinnvoll von glatten Funktionen sprechen zu können. Wir sehen auch, dass es günstiger sein könnte, mit einem topologischen Raum M statt einfach mit einer Menge zu beginnen. Für so einen Raum kann man dann die Existenz von lokalen Parametrisierungen zumindest als Homöomorphismen von offenen Teilmengen eines $\mathbb{R}^n$ verlangen, wobei es sich allerdings eingebürgert hat, die inversen von lokalen Parametrisierungen zu betrachten. Leider macht es keinen Sinn zu verlangen, dass die Parametrisierung selbst glatt sind, weil dieser Begriff für Funktionen nach M (noch) keinen Sinn macht. Trotzdem kann man versuchen, Glattheit von auf M definierten Funktionen durch Glattheit von Kompositionen mit lokalen Parametrisierungen zu definieren. Damit das einen sinnvollen Begriff ergibt, müssen aber die verschiedenen Parametrisierungen eine Verträglichkeitsbedingung erfüllen. Schließlich möchte man noch vermeiden, dass die gewählten Parametrisierungen eine ausgezeichnete Rolle spielen, was die Definition leider etwas sperrig macht:

Definition 4.3. Sei M ein topologischer Raum.

(1) Eine *Karte auf* M ist eine offene Teilmenge $U \subset M$ zusammen mit einem Homöomorphismus u von U auf eine offene Teilmenge $u(U) \subset \mathbb{R}^k$.

(2) Zwei Karten (U_1, u_1) und (U_2, u_2) auf M heißen *verträglich* wenn $U_{12} := U_1 \cap U_2$ entweder leer ist, oder die *Kartenwechselabbildung* $u_2 \circ u_1^{-1} : u_1(U_{12}) \to u_2(U_{12})$ glatt im Sinne der Analysis ist. (Das macht Sinn, weil U_{12} offen in U_1 und U_2 ist und daher $u_1(U_{12})$ und $u_2(U_{12})$ offen in $\mathbb{R}^k$ sind.)

(3) Ein *Atlas* für M ist eine Familie $\{(U_i, u_i) : i \in I\}$ von paarweise verträglichen Karten auf M, sodass die Mengen U_i eine offene Überdeckung von M bilden, also $M = \cup_{i \in I} U_i$ gilt.

(4) Zwei Atlanten $\mathcal{A}_1$ und $\mathcal{A}_2$ für M heißen *äquivalent*, wenn jede Karte von $\mathcal{A}_1$ mit jeder Karte von $\mathcal{A}_2$ verträglich ist.

(5) Eine *glatte Mannigfaltigkeit* ist ein topologischer Raum M, der Hausdorff ist und das zweite Abzählbarkeitsaxiom AA2 erfüllt, zusammen mit einer Äquivalenzklasse von Atlanten auf M. (Die topologischen Voraussetzungen “separabel” und “metrisierbar” sind eher unproblematisch und werden für uns im weiteren keine Rolle spielen.)

Nach Definition sind die Kartenwechselabbildungen Diffeomorphismen zwischen offenen Teilmengen, also müssen zwei Karten mit nichtleerem Schnitt immer in Räume gleicher Dimension abbilden. Damit ist diese Dimension auf Zusammenhangskomponenten von M konstant und meist nimmt man an, dass sie auch für alle Komponenten gleich ist, wodurch eine Mannigfaltigkeit eine wohldefinierte Dimension hat.

Der Vorteil von Karten gegenüber lokalen Parametrisierungen ist, dass man für eine Karte $u : U \to \mathbb{R}^k$ (mit $U \subset M$ offen), die Funktion $u : U \to \mathbb{R}^k$ in Komponenten $(u_1, \dots, u_k)$ zerlegen kann. Diese Komponenten definieren *lokalen Koordinaten* auf $U \subset M$, mit denen man oft sehr handlich rechnen kann.

Man kann einen (geeigneten) topologischen Raum M zu einer Mannigfaltigkeit machen, indem man einen Atlas auf M konstruiert und (formal) dessen Äquivalenzklasse betrachtet. Die umständliche Definition ist nur notwendig, damit der gewählten Atlas keine besondere Rolle spielt. Die übliche Sprechweise ist dann, dass für eine Mannigfaltigkeit M "eine Karte auf M" gerade "eine Karte aus einem der Atlanten der gegebenen Äquivalenzklasse" bedeutet. Das klingt aber wesentlich komplizierter, als es in der Praxis ist.

Beispiel 4.3. Wir werden in Kürze substantielle Beispiele besprechen, daher beschränken wir uns hier auf eher triviale Beispiele, die aber die Denkweise illustrieren.

(1) Ist $U \subset \mathbb{R}^n$ eine offene Teilmenge, dann können wir (U, id_U) als Karte auf U betrachten und diese Karte bildet einen Atlas für U. Damit sind offene Teilmengen in $\mathbb{R}^n$ automatisch glatte Mannigfaltigkeiten.

(2) Sei allgemeiner M eine glatte Mannigfaltigkeit und $W \subset M$ eine offene Teilmenge. Ist $\mathcal{A} = \{(U_i, u_i) : i \in I\}$ ein Atlas für M, dann ist für jedes $i \in I$ der Durchschnitt $U_i \cap W$ offen in M (und damit auch in W und in U_i). Schränken wir uns auf jene i ein, für die der Durchschnitt nichtleer ist, dann ist $u_i(U_i \cap W)$ offen in $\mathbb{R}^k$ und wir können $(U_i \cap W, u_i|_{U_i \cap W})$ als Karte auf M betrachten. Damit erhält man Kartenwechselabbildungen, die einfach Einschränkungen der Kartenwechselabbildungen zwischen den ursprünglichen Karten und damit glatt sind. Also erhalten wir einen Atlas auf W, der W zu einer glatten Mannigfaltigkeit macht. Beginnt man mit einem zu $\mathcal{A}$ äquivalenten Atlas von M, dann überlegt man sofort, dass man einen äquivalenten Atlas für W erhält. Damit hängt die Struktur auf W nicht von irgendwelchen Wahlen ab, sondern nur von der Mannigfaltigkeitsstruktur auf M.

(3) Sei $M \subset \mathbb{R}^n$ eine k-dimensionale Teilmannigfaltigkeit, die wir mit der Teilraumtopologie ausstatten (die automatisch Hausdorff ist und AA2 erfüllt). Ist $x \in M$ und $\Phi : \tilde{U} \to \tilde{V}$ eine lokale Trivialisierung wie in Definition 2.5 mit $x \in \tilde{U}$, dann ist $U := \tilde{U} \cap M$ eine offene Teilmenge, die x enthält. Setzt man $V := \tilde{V} \cap \mathbb{R}^k$, dann folgt aus dem Beweis von Proposition 4.2 sofort, dass $\Phi|_U : U \to V$ ein Homöomorphismus ist und damit eine Karte auf M definiert. Liefert $\tilde{\Phi}$ eine weitere Karte dieser Art, dann ist die resultierende Kartenwechselabbildung einfach eine Einschränkung der glatten Funktion $\tilde{\Phi} \circ \Phi^{-1}$ und damit ebenfalls glatt. Damit ist M in natürlicher Weise eine glatte Mannigfaltigkeit.

4.4. Glatte Funktionen zwischen Mannigfaltigkeiten. Da glatte Mannigfaltigkeiten topologische Räume sind, können wir für zwei Mannigfaltigkeiten M und N eine stetige Funktion $f : M \to N$ betrachten. Sei $x \in M$ ein Punkt und betrachten wir Karten (U, u) für M mit $x \in U$ und (V, v) für N mit $f(x) \in V$. Dann ist $f^{-1}(V) \subset M$ offen (und enthält x), also ist auch $u(f^{-1}(V) \cap U)$ eine offene Teilmenge in $\mathbb{R}^k$. Damit ist aber $v \circ f \circ u^{-1} : u(f^{-1}(V) \cap U) \to v(V)$ eine stetige Funktion zwischen zwei offenen Teilmengen von $\mathbb{R}^k$. Damit bietet sich folgende Definition an:

Definition 4.4. (1) Seien M und N glatte Mannigfaltigkeiten und $f : M \to N$ eine Funktion. Dann ist f *glatt*, wenn f stetig ist und es zu jedem Punkt $x \in M$ Karten (U, u) für M und (V, v) für N mit $x \in U$ und $f(x) \in V$ gibt, sodass die Funktion $v \circ f \circ u^{-1} : u(f^{-1}(V) \cap U) \to v(V)$ glatt im Sinne der Analysis ist.

(2) Ein *Diffeomorphismus* zwischen zwei Mannigfaltigkeiten M und N ist eine bijektive glatte Funktion $f : M \to N$, sodass auch die inverse Funktion $f^{-1} : N \to M$ glatt ist. Mann nennt M und N *diffeomorph*, wenn es einen Diffeomorphismus zwischen M und N gibt.

Für Karten (U,u) und (V,v) wie in Teil (1) der Definition heisst $v \circ f \circ u^{-1}$ die *lokale Koordinatendarstellung von* f. Für $y \in U$ beschreibt sie ja gerade die lokalen Koordinaten $v_j(f(y))$ als Funktionen der lokalen Koordination $u_i(y)$. Die Glattheit von stetigen Funktionen ist also gerade über die Glattheit von Darstellungen in lokalen Koordinaten definiert.

PROPOSITION 4.4. *Seien M und N glatte Mannigfaltigkeiten und $f : M \to N$ eine glatte Funktion.*

(1) Für beliebige Karten (U,u) auf M und $(V.v)$ auf N, sodass $f^{-1}(V) \cap U$ nichtleer ist, ist die Funktion $v \circ f \circ u^{-1} : u(f^{-1}(V) \cap U) \to v(V)$ glatt im Sinne der Analysis.

(2) Die Karten auf M sind genau die Diffeomorphismen zwischen offenen Teilmengen von M und offenen Teilmengen von $\mathbb{R}^k$.

(3) Ist P eine weitere glatte Mannigfaltigkeiten und $g : N \to P$ eine glatte Funktion, dann ist auch die Komposition $g \circ f : M \to P$ glatt. Insbesondere ist Diffeomorphie eine Äquivalenzrelation.

BEWEIS. (1) Es genügt zu zeigen, dass $v \circ f \circ u^{-1}$ lokal um jeden Punkt $z \in u(f^{-1}(V) \cap U)$ glatt ist. Nach Definition finden wir zu $x := u(z)$ Karten $(\tilde{U}, \tilde{u})$ für M mit $x \in \tilde{U}$ und $(\tilde{V}, \tilde{v})$ für N mit $f(x) \in \tilde{V}$, sodass $\tilde{v} \circ f \circ \tilde{u}^{-1} : \tilde{u}(\tilde{U} \cap f^{-1}(\tilde{V})) \to \tilde{v}(\tilde{V})$ glatt ist. Nun ist $\hat{U} := U \cap \tilde{U} \cap f^{-1}(V \cap \tilde{V})$ eine offene Umgebung von x in M, also ist $u(\hat{U})$ eine offene Umgebung von z in $u(f^{-1}(V) \cap U)$. Nach Konstruktion sind die Kartenwechselabbildungen $\tilde{u} \circ u^{-1} : u(\hat{U}) \to \tilde{u}(\hat{U})$ und $v \circ \tilde{v}^{-1} : \tilde{v}(V \cap \tilde{V}) \to v(V \cap \tilde{V})$ jeweils glatt im Sinne der Analysis. Nun ist aber $\tilde{v} \circ f \circ \tilde{u}^{-1}$ glatt auf $\tilde{u}(\hat{U})$ und bildet diese Teilmenge nach $\tilde{v}(V \cap \tilde{V})$ ab. Als Komposition von drei glatten Funktionen ist damit auch

$$(v \circ \tilde{v}^{-1}) \circ (\tilde{v} \circ f \circ \tilde{u}^{-1}) \circ (\tilde{u} \circ u^{-1}) = v \circ f \circ u^{-1}$$

glatt im Sinne der Analysis als Funktion $u(\hat{U}) \to v(V)$, also lokal um z.

(2) Ist (U,u) eine Karte auf M, dann sind $U \subset M$ und $u(U) \subset \mathbb{R}^k$ offene Teilmengen und damit Mannigfaltigkeiten. Nach Definition ist u ein Homöomorphismus, also sind $u : U \to u(U)$ und $u^{-1} : u(U) \to U$ stetig. Betrachten wir $u : U \to u(U)$, dann können wir für jeden Punkt $x \in U$ die Karten (U,u) für U und $(u(U), \mathrm{id}_{u(U)})$ für $u(U)$ verwenden und natürlich ist $\mathrm{id}_{u(U)} \circ u \circ u^{-1} = \mathrm{id}_{u(U)}$ glatt im Sinne der Analysis, also ist u glatt. Analog ist die Funktion $u^{-1} : u(U) \to U$ glatt, weil $u \circ u^{-1} \circ \mathrm{id}_{u(U)} = \mathrm{id}_{u(U)}$ glatt ist, also ist $u : U \to u(U)$ ein Diffeomorphismus.

Sei umgekehrt $W \in M$ offen und $f : W \to f(W)$ ein Diffeomorphismus auf eine offene Teilmenge $f(W) \subset \mathbb{R}^k$. Dann sind f und f^{-1} als glatte Funktionen stetig und damit ist f ein Homomorphismus. Für eine Karte (U,u) auf M mit $W \cap U \neq \emptyset$ ist nach (1) die Funktion $\mathrm{id}_{f(W)} \circ f \circ u^{-1} : u(W \cap U) \to f(W \cap U)$ glatt im Sinne der Analysis. Analog folgt aus Glattheit von f^{-1}, dass $u \circ f^{-1} : f(W \cap U) \to u(U)$ glatt im Sinne der Analysis ist. Damit sind die Kartenwechsel glatt und wir können (W,f) zu einem beliebigen Atlas auf M hinzufügen und erhalten wieder einen Atlas. Damit ist (W,f) eine Karte auf M.

(3) Für $x \in M$ wählen wir beliebige Karten (U,u) für M mit $x \in U$, (V,v) für N mit $f(x) \in V$ und (W,w) für P mit $g(f(x)) \in W$. Dann ist $\hat{U} := U \cap f^{-1}(V \cap g^{-1}(W))$ eine offene Umgebung von x in M und für $\hat{u} : u|_{\hat{U}}$ ist $(\hat{U}, \hat{u})$ eine Karte auf M mit $x \in \hat{U}$. Nun ist aber $f(\hat{U}) \subset V \cap g^{-1}(W)$, also $(g \circ f)(\hat{U}) \subset W$. Damit gilt aber $f \circ \hat{u}^{-1} = v^{-1} \circ v \circ f \circ \hat{u}^{-1}$ und damit

$$w \circ g \circ f \circ \hat{u}^{-1} = (w \circ g \circ v^{-1}) \circ (v \circ f \circ \hat{u}^{-1}) : \hat{u}(\hat{U}) \to w(W).$$

Das ist glatt im Sinne der Analysis als Komposition zweier Funktionen, die nach (1) glatt im Sinne der Analysis sind, also ist $g \circ f$ eine glatte Funktion. Damit ist auch die Komposition von zwei Diffeomorphismen ein Diffeomorphismus, also ist Diffeomorphie eine Äquivalenzrelation. □

Zusammen mit Beispiel 4.3 (3) und Proposition 4.2 zeigt dieses Resultat auch dass wir tatsächlich Verallgemeinerungen der bisher studierten Begriffe erhalten. Betrachten wir Teilmannigfaltigkeiten $M \subset \mathbb{R}^n$ und $N \subset \mathbb{R}^m$ wie in Beispiel 4.3 (3) als glatte Mannigfaltigkeiten, dann ist eine Funktion $f : M \to N$ genau dann glatt im Sinn von Definition 2.5 wenn sie glatt im Sinn von Definition 4.4 ist. Das umfasst insbesondere den Fall offener Teilmengen von $\mathbb{R}^n$.

4.5. Homogene Räume als Mannigfaltigkeiten. Betrachten wir nun eine Matrixgruppe G, eine abgeschlossene Untergruppe $H \subset G$ (die damit selbst eine Matrixgruppe ist), den Raum G/H der linken Nebenklassen und die kanonische Abbildung $p : G \to G/H$. Wir statten G/H mit der Quotiententopologie aus, definieren also eine Teilmenge $U \subset G/H$ als offen genau dann, wenn $p^{-1}(U) \subset G$ offen ist. Damit ist p stetig und wir möchten G/H so zu einer Mannigfaltigkeit machen, dass die Funktion p glatt ist.

Satz 4.5. *Sei $G \subset GL(n,\mathbb{R})$ eine Matrixgruppe, $H \subset G$ eine abgeschlossene Untergruppe. Dann kann der Raum G/H der linken Nebenklassen auf natürliche Art zu einer Mannigfaltigkeit der Dimension* $\dim(G) - \dim(H)$ *gemacht werden. Die kanonische Abbildung $p : G \to G/H$ ist glatt und für eine beliebige Mannigfaltigkeit M ist eine Funktion $F : G/H \to M$ genau dann glatt, wenn $p \circ F : G \to M$ glatt ist.*

Beweis. Da $H \subset G$ abgeschlossen ist, ist H selbst eine Matrixgruppe. Beschreibt man die Lie Algebren durch $\mathfrak{h} = \{X \in M_n(\mathbb{R}) : \forall t \in \mathbb{R} : \exp(tX) \in H\}$ und analog für $\mathfrak{g}$, dann ist offensichtlich $\mathfrak{h} \subset \mathfrak{g}$. Damit ist $\mathfrak{h}$ ein linearer Teilraum von $\mathfrak{g}$ und wir wählen einen komplementären Teilraum $\mathfrak{k} \subset \mathfrak{g}$. Wählen wir schließlich noch ein Komplement $\mathfrak{l}$ zu $\mathfrak{g}$ in $M_n(\mathbb{R})$, dann ist $\mathfrak{k} \oplus \mathfrak{l}$ ein Komplement zu $\mathfrak{h}$ in $M_n(\mathbb{R})$. Nun können wir die Argumente aus dem Beweis von Satz 2.6 sowohl auf G als auch auf H anwenden.

Zunächst finden wir offene Nullumgebungen V in $\mathfrak{h}$ und $\tilde{W}$ in $\mathfrak{k} \oplus \mathfrak{l}$, sodass $(X,Y) \mapsto \exp(X)\exp(Y)$ ein Diffeomorphismus von $\tilde{W} \times V$ auf eine offene Umgebung $\tilde{U}$ von $\mathbb{I}$ in $GL(n,\mathbb{R})$ ist, deren Schnitt mit H nur aus den Elementen der Form $\exp(Y)$ für $Y \in V$ besteht. Indem wir $\tilde{W}$ wenn nötig Verkleinern, dürfen wir annehmen, dass es die Form $W_1 \times W_2$ für Nullumgebungen W_1 in $\mathfrak{k}$ und W_2 in $\mathfrak{l}$ hat. Durch weiteres Verkleinern können wir auch erreichen, dass $\varphi(X,Y) := \exp(X)\exp(Y)$ einen Diffeomorphismus von $W_1 \times V$ auf eine offene Umgebung U' von $\mathbb{I}$ in G definiert, wobei wir annehmen dürfen, dass $U' \subset \tilde{U} \cap G$ gilt. Schließlich definiert $(X_1, X_2) \mapsto \exp(X_1)^{-1}\exp(X_2)$ eine stetige Funktion $\mathfrak{k} \times \mathfrak{k} \to G$. Damit gibt es eine offene Nullumgebung $\hat{W}$ in $\mathfrak{k}$, sodass $\exp(X_1)^{-1}\exp(X_2) \in U'$ für alle $X_1, X_2 \in \hat{W}$ gilt und wir definieren $W := W_1 \cap \hat{W}$. Offensichtlich ist das eine offene Nullumgebung in $\mathfrak{k}$.

Nun betrachten wir die (offensichtlich glatte) Funktion $f : \mathfrak{k} \times H \to G$, die durch $f(X,h) := \exp(X)h$ definiert ist und behaupten, dass sie einen Diffeomorphismus von $W \times H$ auf eine offene Umgebung U von $\mathbb{I}$ in G definiert. Für $X_1, X_2 \in W$ und $h_1, h_2 \in H$ folgt aus $\exp(X_1)h_1 = \exp(X_2)h_2$ natürlich $\exp(X_1)^{-1}\exp(X_2) = h_1 h_2^{-1} \in H \cap U' \subset H \cap \tilde{U}$. Damit muss es aber eine Element $Y \in V$ geben, sodass $\exp(X_1)^{-1}\exp(X_2) = \exp(Y)$ gilt. Das bedeutet aber gerade, dass $\varphi(X_1, Y) = \varphi(X_2, 0)$ gilt und weil φ injektiv ist, folgt $X_1 = X_2$ und $Y = 0$. Damit folgt aber auch $h_1 = h_2$ und wir sehen, dass $f : W \times H \to G$ injektiv ist.

Nach Konstruktion gilt $f(X, \exp(Y)) = \varphi(X, Y)$ für $Y \in \mathfrak{h}$ hinreichend nahe bei 0. Damit können wir f auf einer Umgebung von $W \times \{\mathbb{I}\}$ als $f = \varphi \circ (\mathrm{id}, \exp^{-1})$ schrieben. Daraus folgt sofort, dass für jedes $X \in W$ die Tangentialabbildung $T_{(X,e)}f$ invertierbar und nach Korollar 4.2 ist f lokal um diesen Punkt ein Diffeomorphismus. Natürlich gilt $f \circ (\mathrm{id}, \rho^h) = \rho^h \circ f$ und da ρ^h sowohl auf H als auch auf G ein Diffeomorphismus ist, ist auch $T_{(X,h)}f$ ein linearer Isomorphismus und f auch lokal um diesen Punkt eine Diffeomorphismus. Damit enthält aber $U := f(W \times H) \subset G$ für jeden Punkt (X, h) eine offene Umgebung von $f(X, h)$ und ist damit selbst offen und da $f : W \times H \to U$ bijektiv ist und lokal glatte Inverse besitzt, ist f ein Diffeomorphismus.

Betrachten wir die Teilmenge $p(U) \in G/H$. Nach Konstruktion gilt für $g \in U$ und $h \in H$ offensichtlich $gh \in U$ also folgt $p^{-1}(p(U)) = U$, also ist $p(U)$ offen in G/H. Nun definieren wir $\psi : W \to p(U)$ durch $\psi(X) := p(\exp(X))$. Dann ist ψ nach Konstruktion surjektiv und für $X_1, X_2 \in W$ folgt aus $\psi(X_1) = \psi(X_2)$ natürlich, dass $\exp(X_1)H = \exp(X_2)H$ gelten muss. Da aber $f(X, h) = \exp(X)h$ injektiv auf $W \times H$ ist, folgt daraus natürlich, dass $X_1 = X_2$ gilt, also ist ψ bijektiv. Offensichtlich ist $\psi = p \circ \exp$ stetig für die Quotiententopologie. Ist $W' \subset W$ offen, dann ist $p^{-1}(\psi(W')) = \{\exp(X)h : X \in W', h \in H\} = f(W' \times H)$ und weil f ein Diffeomorphismus und daher ein Homöomorphismus ist, ist diese Teilmenge offen in G. Damit ist aber auch $\psi(W')$ offen in G/H und damit in $p(U)$, also ist $\psi : W \to p(U)$ ein Homöomorphismus.

Für $\tilde{g}H \in G/H$ ist $g^{-1}\tilde{g}H$ eine wohldefinierte Nebenklasse in G/H, also können wir eine Funktion $\ell_g : G/H \to G/H$ durch $\ell_g(\tilde{g}H) := g\tilde{g}H$ definieren. Nach Konstruktion gilt $\ell_g \circ p = p \circ \lambda_g$ für die Linkstranslation λ_g. Damit ist $\ell_g \circ p$ stetig, also ist auch $\ell_g : G/H \to G/H$ stetig für die Quotiententopologie. Nun ist aber natürlich $\ell_{g^{-1}}$ invers zu ℓ_g und ebenfalls stetig also ist jedes ℓ_g ein Homöomorphismus. Für $g \in G$ betrachten wir die offene Teilmenge $U_g := \{g \exp(X)H : X \in W\} = \ell_g(p(U)) \subset G/H$ und definieren $u_g := \psi^{-1} \circ \ell_{g^{-1}} : U_g \to W$. Dann ist jedes u_g ein Homöomorphismus und wir behaupten, $\{(U_g, u_g) : g \in G\}$ ein Atlas für G/H ist. Offensichtlich ist $\cup_{g \in G} U_g = G/H$, also müssen wir nur noch die Verträglichkeit der Karten überprüfen. Nehmen wir also an, dass $U_g \cap U_{g'} \neq \emptyset$ für $g, g' \in G$ gilt und betrachten $u_{g'} \circ u_g^{-1} : W \to W$. Für $X \in W$ ist $u_g^{-1}(X) = g \exp(X)H$ und

$$u_{g'}(g \exp(X)H) = \psi^{-1}((g')^{-1} g \exp(X)H)$$

und das ist die erste Komponente von $f^{-1}((g')^{-1} g \exp(X)) \in W \times H$. Damit ist aber $u_{g'} \circ u_g^{-1} = pr_1 \circ f^{-1} \circ \lambda_{(g')^{-1}g} \circ \exp$ und das ist glatt als Komposition glatter Funktionen. Damit haben wir G/H zu einer glatten Mannigfaltigkeit der Dimension $\dim(\mathfrak{k}) = \dim(\mathfrak{g}) - \dim(\mathfrak{h})$ gemacht.

Von oben wissen wir, dass es eine Nullumgebung V in $\mathfrak{h}$ gibt, sodass $\varphi(X, Y) = \exp(X)\exp(Y)$ einen Diffeomorphismus von $W \times V$ auf eine offene Umgebung von $\mathbb{I}$ in G definiert. Damit definiert für jedes $g \in G$ die Funktion $(X, Y) \mapsto g \exp(X) \exp(Y)$ einen Diffeomorphismus von $W \times V$ auf eine offene Umgebung von $g \in G$, dessen Inverse wir nach Proposition 4.4 wir als Karte $z_g : Z_g \to W \times V$ auf G verwenden können. Nun ist aber

$$u_g \circ p \circ z_g^{-1}(X, Y) = (u_g \circ p)(g \exp(X) \exp(Y)) = u_g(g(\exp(X)H)) = X$$

und weil das für jedes $g \in G$ funktioniert ist p glatt. Für eine glatte Funktion F von G/H in eine beliebige Mannigfaltigkeit M ist damit auch $F \circ p : G \to M$ glatt. Andererseits kann man jedes Element in U_g eindeutig als $g \exp(X)H$ für ein $X \in W$ schreiben. Damit können wir eine Funktion $\sigma_g : U_g \to G/H$ durch $\sigma_g(g \exp(X)H) := g \exp(X)$ und offensichtlich gilt $z_g \circ \sigma_g \circ u_g^{-1}(X) = X$, also ist σ_g glatt. Nach Konstruktion erfüllt das $p \circ \sigma_g = \mathrm{id}_{U_g}$. Sei nun $F : G/H \to M$ eine Funktion in eine Mannigfaltigkeit, sodass $F \circ p : G \to M$ glatt ist. Dann folgt aus der Stetigkeit von $F \circ p$ die Stetigkeit von F und weil man $F|_{U_g}$ als $(F \circ p) \circ \sigma_g$ schreiben kann, ist diese Einschränkung glatt, woraus sofort folgt, dass F glatt ist. □

Daraus können wir sofort einige weitere Folgerungen ziehen: Im Beweis haben wir für jedes $g \in G$ die Funktionen $\ell_g : G/H \to G/H$ betrachtet, die durch $\ell_g(\tilde{g}H) = g\tilde{g}H$ gegeben ist. Wir habe auch schon bemerkt, dass $\ell_g \circ p = p \circ \lambda_g$ für die Linkstranslation $\lambda_g : G \to G$ gilt. Da λ_g glatt ist, ist auch $p \circ \lambda_g$ glatt, also ist nach dem letzten Teil des Satzes $\ell_g : G/H \to G/H$ glatt. Schließlich haben wir aber auch schon gesehen, dass $\ell_{g^{-1}}$ invers zu ℓ_g ist, also ist jedes ℓ_g ein Diffeomorphismus. Andererseits gilt nach Konstruktion $\ell_{\mathbb{I}} = \mathrm{id}_{G/H}$ und $\ell_{g_1 g_2} = \ell_{g_1} \circ \ell g_2$ also haben wir eine Wirkung der Gruppe G auf der Mannigfaltigkeit G/H definiert, in der die Wirkung jedes $g \in G$ ein Diffeomorphismus ist. Das ist ein typisches Beispiel für eine *glatte Wirkung* einer Matrixgruppe auf einer Mannigfaltigkeit. Insbesondere finden für beliebige Punkte $x = g_1 H \in G/H$ und $y = g_2 H \in G/H$ einen Diffeomorphismus von G/H (nämlich $\ell_{g_2 g_1^{-1}}$) der x auf y abbildet. Das bedeutet, dass G/H "überall gleich aussieht", weshalb man diese Räume auch als *homogene Räume* bezeichnet.

4.6. Beispiele. (1) Betrachten wir die Einheitssphäre $S^{n-1} \subset \mathbb{R}^n$, also $\{v \in \mathbb{R}^n : \|v\| = 1\}$. Für eine Matrix $A \in O(n)$ und $v \in S^{n-1}$ ist $\|Av\| = \|v\|$, also $Av \in S^{n-1}$. Man rechnet sofort nach, dass $(A, v) \mapsto Av$ eine Wirkung der Gruppe $O(n)$ auf S^{n-1} definiert. Für einen beliebigen Einheitsvektor $v \in S^{n-1}$ gibt es natürlich eine Orthonormalbasis von $\mathbb{R}^n$, deren erstes Element v ist. Schreibt man die Basisvektoren als Spalten in eine Matrix A, dann ist $A \in O(n)$ und $Ae_1 = v$, also ist die Bahn $O(n) \cdot e_1$ von e_1 die ganze Sphäre S^{n-1}. Die Standgruppe H von e_1 in $O(n)$ besteht offensichtlich aus allen Matrizen in $O(n)$, deren erste Spalte e_1 ist. Da aber die restlichen Spalten orthogonal auf die erste Spalte stehen, muss die Matrix Blockform $\left(\begin{smallmatrix} 1 & 0 \\ 0 & B \end{smallmatrix}\right)$ haben. Eine Matrix dieser Form ist aber genau dann orthogonal, wenn $B \in O(n-1)$ gilt, also ist insbesondere $H \cong O(n-1)$. Die Abbildung $A \mapsto Ae_1$ induziert somit eine bijektive Funktion $O(n)/O(n-1) \to S^{n-1}$. Die Komposition dieser Funktion mit $p : O(n) \to O(n)/O(n-1)$ ist aber gerade $A \mapsto Ae_1$ und damit offensichtlich glatt, also ist auch unsere Funktion $O(n)/O(n-1) \to S^{n-1}$ glatt. Man kann leicht explizit verifizieren, dass diese Funktion ein Diffeomorphismus ist, also stimmt die "übliche" Mannigfaltigkeitsstruktur auf S^{n-1} mit der auf $O(n)/O(n-1)$ überein. Analog geht das mit $SO(n)$.

Analog kann man die Einheitssphäre $S^{2n-1} \subset \mathbb{C}^n$ betrachten, auf der neben $SO(2n)$ und $O(2n)$ auch die Untergruppen $SU(n)$ und $U(n)$ mit nur einer Bahn wirken. Damit bekommt man verschiedene Darstellungen von S^{2n-1} als homogener Raum, zum Beispiel $S^{2n-1} \cong SU(n)/SU(n-1)$ die wieder die übliche Mannigfaltigkeitsstruktur auf S^{2n-1} liefern. Man kann diese Darstellungen so interpretieren, dass man auf den Sphären verschiedene geometrische Strukturen finden kann, sodass die Wirkungen der Gruppen jeweils genau die Diffeomorphismen der Sphäre liefern, die mit der entsprechenden Struktur verträglich sind.

Diese Beispiele bilden auch den Hintergrund für unser Studium der Zusammenhangseigenschaften der klassischen Gruppen in Abschnitt 3.3. Der wesentliche Schritt dort war zu zeigen, dass die Gruppen $SO(n)$ und $SU(n)$ für zusammenhängend sind und die Tatsache, dass die Sphären S^{n-1} zusammenhängend sind war die Basis für ein Induktionsargument. Tatsächlich kann man allgemein zeigen, dass für eine abgeschlossene Untergruppe H in einer Matrixgruppe G der homogene Raum G/H genau dann zusammenhängend ist, wenn jede Zusammenhangskomponente von G die Untergruppe H

schneidet, siehe Lemma 1.17 in [**LieGrp**]. Daraus folgt insbesondere, dass G höchstens so viele Zusammenhangskomponenten besitzt wie H.

(2) Betrachten wir $G := GL(n,\mathbb{R})$ und $k < n$ und die Wirkung $A \cdot V := A(V)$ von G auf der Menge $Gr(k,\mathbb{R}^n)$ der k–dimensionalen Teilräume von $\mathbb{R}^n$ aus Beispiel 4.1. Sei $H := G_{\mathbb{R}^k} \subset G$ die Standgruppe des Teilraumes $\mathbb{R}^k \subset \mathbb{R}^n$. Nach Konstruktion bedeutet $A \in H$ genau, dass die ersten k Spaltenvektoren von A den Teilraum $\mathbb{R}^k$ aufspannen. Da die Spaltenvektoren einer invertierbaren Matrix aber immer linear unabhängig sind, ist das äquivalent dazu, dass die ersten k Spaltenvektoren in $\mathbb{R}^k \subset \mathbb{R}^n$ liegen. Für $A = (a_{ij})$ bedeutet das aber genau, dass $a_{ij} = 0$ für $i > k$ und $j \leq k$ gelten muss. Damit ist aber offensichtlich, dass H abgeschlossen in G ist und damit ist G/H nach Satz 4.5 eine glatte Mannigfaltigkeit. Aus Abschnitt 4.1 wissen wir, dass $AH \mapsto A(\mathbb{R}^k)$ eine Bijektion zwischen G/H und $Gr(k,\mathbb{R}^n)$ definiert, also haben wir so $Gr(k,\mathbb{R}^n)$ zu einer glatten Mannigfaltigkeit gemacht ("Grassmann Mannigfaltigkeiten"). Aus dem letzten Teil von Satz 4.5 folgt auch, dass man glatte Funktionen von $Gr(k,\mathbb{R}^n)$ in eine beliebige Mannigfaltigkeit M gut beschreiben kann. Für eine Funktion $f : Gr(k,\mathbb{R}^n) \to M$ erhält man $p \circ f : G \to M$ ja gerade, indem man eine Matrix A auf den Wert von f auf dem von den ersten k Spaltenvektoren von A aufgespannten Teilraum abbildet und Glattheit von f ist äquivalent zu Glattheit dieser Funktion.

Diese Konstruktion kann man leicht variieren. Beginnen wir mit der Einschränkung der Wirkung auf $O(n)$ und setzen K wieder die Standgruppe von $\mathbb{R}^k$. Für $A \in K$ müssen dann die ersten k Spaltenvektoren eine Orthonormalbasis für $\mathbb{R}^k$ bilden, während die restlichen Spaltenvektoren orthogonal auf die ersten k stehen müssen. Damit folgt aber sofort, dass $K \subset O(n)$ genau aus jenen Matrizen in G besteht, die blockdiagonal mit Blöcken der Größe k und $n-k$ sind (vergleiche mit Abschnitt 3.5). Nun ist eine Blockdiagonalmatrix offensichtlich genau dann orthogonal, wenn die beiden Blöcke auf der Diagonale orthogonal sind, also ist $K \cong O(k) \times O(n-k)$. Damit können wir $Gr(k,\mathbb{R}^n)$ auch mit dem homogenen Raum $O(n)/(O(k) \times O(n-k))$ identifizieren. Da $p : O(n) \to O(n)/(O(k) \times O(n-k))$ stetig und surjektiv ist folgt aus der Kompaktheit von $O(n)$ in diesem Bild sofort, dass wir $Gr(k,\mathbb{R}^n)$ sogar zu einer kompakten Mannigfaltigkeit gemacht haben.

Die Inklusion $i : O(n) \to GL(n,\mathbb{R})$ hat die Eigenschaft, dass $O(n) \cap H = K$ gilt. Damit induziert aber $AK \mapsto AH$ eine wohldefinierte Funktion $j : O(n)/K \to G/H$, die $j \circ p = p \circ i$ erfüllt und daher glatt ist. Man verifiziert leicht, dass diese Funktion einen Diffeomorphismus definiert, also liefern beide Konstruktionen die gleiche Mannigfaltigkeitsstruktur auf $Gr(k,\mathbb{R}^n)$. Eine natürliche Interpretation dieser Situation ist, dass man eine zusätzliche geometrische Struktur auf $Gr(k,\mathbb{R}^n)$ wählen kann (in diesem Fall eine Riemann–Metrik) und die Untergruppe $O(n) \subset GL(n,\mathbb{R})$ genau aus jenen Matrizen besteht, deren Wirkung auf $Gr(k,\mathbb{R}^n)$ mit dieser zusätzlichen Struktur verträglich sind. (Man kann auch schon die Wirkungen von Elementen von $GL(n,\mathbb{R})$ als jene Diffeomorphismen von $Gr(k,\mathbb{R}^n)$ charakterisieren, die mit einer "gröberen" geometrischen Struktur verträglich sind.)

(3) Die Ideen aus (2) kann man auf viele Fragen der linearen Algebra anwenden. Aus Abschnitt 3.6 kennen wir die Darstellung von $G := GL(n,\mathbb{R})$ auf dem Raum $S^2\mathbb{R}^{n*}$ der symmetrischen Bilinearformen auf $\mathbb{R}^n$, die durch $(A \cdot \beta)(v,w) = \beta(A^{-1}v, A^{-1}w)$ definiert ist. Das standard innere Produkt $\beta_0 := \langle\ ,\ \rangle$ ist so eine Bilinearform und seine Standgruppe ist nach Definition $O(n) \subset GL(n,\mathbb{R})$. Damit sehen wir aber, dass die Abbildung $A \mapsto A \cdot \beta_0$ eine Bijektion zwischen dem homogenen Raum $GL(n,\mathbb{R})/O(n)$ und der Bahn $G \cdot \beta_0$ induziert. Wir behaupten, dass diese Bahn genau aus allen positiv

definiten symmetrischen Bilinearformen auf $\mathbb{R}^n$ besteht, also ist $GL(n,\mathbb{R})/O(n)$ der Raum aller inneren Produkte auf $\mathbb{R}^n$, der somit eine glatte Mannigfaltigkeit ist.

Einerseits gilt für $v \in \mathbb{R}^n$ mit $v \neq 0$ natürlich $(A \cdot \beta_0)(v,v) = \langle A^{-1}v, A^{-1}v\rangle > 0$, also besteht die Bahn nur aus positiv definiten symmetrischen Bilinearformen. Ist andererseits $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ symmetrisch und positiv definit, dann ist aus der linearen Algebra bekannt, dass man eine Orthonormalbasis für β findet, also Vektoren $v_1, \dots, v_n \in \mathbb{R}^n$, sodass $\beta(v_i, v_j) = \delta_{ij}$ gilt. Sei nun A die Matrix mit den Spaltenvektoren $v_1, \dots, v_n$, dann gilt $Ae_i = v_i$ und somit $A^{-1}v_i = e_i$ für alle $i = 1, \dots, n$. Damit folgt aber sofort, dass $(A \cdot \beta_0)(v_i, v_j) = \delta_{ij} = \beta(v_i, v_j)$ für alle i und j gilt und weil die v_i eine Basis bilden folgt daraus $A \cdot \beta_0 = \beta$, was den Beweis der Behauptung vervollständigt.

In diesem Beispiel kann man die Mannigfaltigkeitsstruktur auch einfacher finden: Für eine Bilinearform β auf $\mathbb{R}^n$ gibt es eine eindeutig bestimmte Matrix $A_\beta \in M_n(\mathbb{R})$ sodass $\beta(v, w) = \langle v, Aw\rangle$ gilt. Die Form β ist genau dann symmetrisch wenn A symmetrisch ist in dem Sinne, dass $A^t = A$ ist und die symmetrischen Matrizen bilden einen Teilraum von $M_n(\mathbb{R})$. Aus der linearen Algebra ist bekannt, wie man positiv definite symmetrische Bilinearformen in diesem Bild charakterisieren kann. Das sogenannte Hauptminorenkriterium, sagt dass dies genau dann gilt, wenn für $A = (a_{ij})$ jede der Teilmatrizen $(a_{ij})_{1\leq i,j\leq k}$ positive Determinante hat (siehe Satz 9.12 in [**LinAlg**]). Aus dieser Beschreibung ist aber leicht zu sehen, dass die positiv definiten symmetrischen Matrizen eine offene Teilmenge im Raum aller symmetrischen Matrizen bilden. Damit erhalten wir eine Mannigfaltigkeitsstruktur die mit der Struktur von $GL(n,\mathbb{R})/O(n)$ übereinstimmt.

4.7. Ausblick. Zum Abschluss dieser Einführung in die Theorie der Matrixgruppen und der Analysis auf Mannigfaltigkeiten besprechen wir kurz "wie es in dieser Richtung weitergeht". Insbesondere wollen wir sehen, wie man glatte Funktionen zwischen Mannigfaltigkeiten tatsächlich differenzieren kann. Da die Ableitung einer glatten Funktion eine lineare Approximation an die Funktion in einem Punkt sein soll, benötigt man zunächst passende Vektorräume auf denen die Ableitung wirken kann. Für eine Mannigfaltigkeit M und einen Punkt $x \in M$ muss man also wieder den *Tangentialraum* T_xM an M in x definieren, der für eine k–dimensionale Mannigfaltigkeit ein k–dimensionaler Vektorraum sein soll. Da man im Gegensatz zu Teilmannigfaltigkeiten keinen umgebenden $\mathbb{R}^n$ zur Verfügung hat, ist das etwas aufwändig.

Das Konzept einer glatten Kurve in einer Mannigfaltigkeit ist kein Problem und für eine glatte Kurve $c : I \to M$ mit $c(0) = x$ und eine Karte (u, U) für M mit $x \in U$ ist $u \circ c$ eine lokal um 0 definierte glatte Kurve in der offenen Teilmenge $u(U) \subset \mathbb{R}^k$. Man zeigt dann, dass für Kurven c_1 und c_2 mit $c_1(0) = c_2(0) = x$ die Frage ob $(u \circ c_1)'(0)$ und $(u \circ c_2)'(0)$ gleich sind unabhängig von der Wahl der Karte (u, U) ist. Dann definiert man zwei solche Kurven als äquivalent, wenn $(u\circ c_1)'(0) = (u\circ c_2)'(0)$ gilt und bezeichnet die Menge der Äquivalenzklassen mit T_xM. Natürlich soll die Klasse von c genau den "Tangentialvektor" $c'(0)$ beschreiben. Für jede fixe Karte (u, U) kann man T_xM mit $\mathbb{R}^k$ identifizieren, indem man die Klasse von c auf $(u\circ c)'(0)$ abbildet und damit T_xM zu einem Vektorraum machen. Für verschiedene Karten sind die entsprechenden Identifikationen über die Ableitung der Kartenwechselabbildung in einem passenden Punkt verbunden, also ist die erhaltene Vektorraumstruktur auf T_xM unabhängig von der Wahl der Karte. Für eine glatte Funktion $f : M \to N$ und den Punkt $x \in M$ kann man dann $T_xf : T_xM \to T_{f(x)}N$ definieren, indem man die Klasse einer Kurve c durch x auf die Klasse von $f\circ c$ durch $f(x)$ abbildet. Das entspricht genau der Idee aus Abschnitt 2.6 dass die Ableitung von f den Tangentialvektor $c'(0)$ auf $(f \circ c)'(0)$ abbilden muss, damit die Kettenregel gelten kann. Man rechnet leicht nach, dass das wohldefiniert und linear ist. Um mit den Tangentialräumen einfacher umgehen zu können ist es handlicher, eine kartenunabhängige Beschreibung von T_xM zu verwenden, aber darauf wollen wir hier nicht eingehen.

Um die Ableitung von f als Funktion beschreiben zu können, muss man wieder die verschiedenen Tangentialräume zusammenfassen, was auch schwieriger ist, weil sie nicht als Teilräume eines gemeinsamen umgebenden Vektorraumes definiert sind. Die Idee ist aber wieder, die Menge $TM := \{(x,v) : x \in M, v \in T_xM\}$ zu betrachten und diese zu einer Mannigfaltigkeit zu machen. Dazu benutzt man, dass wir ja oben für eine Karte (u,U) für M und jedes $y \in U$ eine Identifikation von T_yM mit $\mathbb{R}^k$ erhalten haben, die man einfach als $T_yu : T_yM \to \mathbb{R}^k$ betrachten kann. Setzt man nun $TU := \{(x,v) : x \in U, v \in T_xM\} \subset TM$ dann kann man eine Bijektion $Tu : TU \to u(U) \times \mathbb{R}^k$ durch $Tu(x,v) := (u(x), T_xu \cdot v)$ definieren. Nun wählt man einen Atlas $\{(U_i,u_i) : i \in I\}$ für M und definiert eine Topologie auf TM indem man die Abbildungen Tu_i für alle i zu Homöomorphismen erklärt. Dann kann man jedes (TU_i, Tu_i) als Karte auf TM betrachten. Die Kartenwechselabbildungen zwischen solchen Karten sind durch die Kartenwechsel der ursprünglichen Karten und ihre Ableitungen gegeben und damit glatt, also erhält man so einen Atlas für TM, das damit zu einer glatten Mannigfaltigkeit wird. Schließlich zeigt man (relativ einfach), dass diese Struktur nicht von der Wahl des Atlas abhängt, also ist TM in natürlicher Weise eine Mannigfaltigkeit. Für eine glatte Funktion $f : M \to N$ erhält man dann eine glatte Funktion $Tf : TM \to TN$ indem man $Tf(x,v) := (f(x), T_xf(v))$ setzt und es gilt die Kettenregel $T(g \circ f) = Tg \circ Tf$. Damit hat man die wesentlichen Zutaten um die Analysis auf Mannigfaltigkeiten zu entwickeln.

Hat man den Begriff der abstrakten Mannigfaltigkeit an der Hand, dann kann man den Begriff der Matrixgruppe auf naheliegende Weise Verallgemeinern. Dazu muss man nur noch zeigen, dass man für Mannigfaltigkeiten M und N den Produktraum $M \times N$ (mit der Produkttopologie) kanonisch zu einer Mannigfaltigkeit machen kann, was ganz einfach ist. Dann kann man eine *Lie Gruppe* als eine glatte Mannigfaltigkeit G definieren, die mit einer Gruppenstruktur ausgestattet ist, sodass die Multiplikation und die Inversion glatt als Funktionen $\mu : G \times G \to G$ bzw. $\nu : G \to G$ sind. (Man kann auch zeigen, dass die Glattheit von μ schon die Glattheit von ν impliziert, aber das ist nicht so wichtig.) Der Übergang zu Lie Gruppen ist aber primär aus konzeptuellen Gründen handlich, vor allem um verschiedene Konstruktionen durchführen zu können, die aus der Klasse der Matrixgruppen herausführen können. Die bei weitem wichtigsten Beispiele für Lie Gruppen sind aber Matrixgruppen und es gibt auch allgemeine Resultate, etwa dass jede kompakte Lie Gruppe isomorph (als Lie Gruppe) zu einer Matrixgruppe ist.

Für eine Lie Gruppe G definiert man die Lie Algebra $\mathfrak{g}$ als den Tangentialraum T_eG an G im neutralen Element $e \in G$. Mit analytischen Methoden (Vektorfelder und die Lie Klammer von Vektorfeldern) definiert man eine Lie Klammer auf $\mathfrak{g}$. Damit kann man dann wieder glatte Homomorphismen differenzieren und die Ableitungen sind wieder Lie Algebra Homomorphismen. Analytische Methoden (Flüsse von Vektorfeldern) werden auch benötigt um eine allgemeine Version der Exponentialabbildung $\exp : \mathfrak{g} \to G$ zu erhalten, die man dann auch wieder in Termen von 1–Parameter–Untergruppen interpretieren kann. Die adjungierte Darstellung $\mathrm{Ad} : G \to GL(\mathfrak{g})$ erhält man, indem man für jedes Element g die Konjugationsabbildung $h \mapsto ghg^{-1}$ (die ein glatter Homomorphismus ist) differenziert um eine Lie Algebren Homomorphismus $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ zu erhalten. Praktisch alle Resultate dieser Vorlesung haben ein Gegenstück im Setting der Lie Gruppen mit parallelen Definitionen und oft ähnlichen Beweisen. Insbesondere sind abgeschlossene Untergruppen in Lie Gruppen automatisch Teilmannigfaltigkeiten (in passendem Sinn analog zu Definition 2.5) und damit selbst Lie Gruppen. Für eine abgeschlossene Untergruppe H einer Lie Gruppe G kann man wieder den Raum G/H der linken Nebenklassen in natürlicher Weise zu einer Mannigfaltigkeit machen.

Auch Darstellung kann man für allgemeine Lie Gruppen definieren und Darstellungstheorie ist ein weites Feld, insbesondere wenn man auch Darstellungen auf unendlichdimensionalen Räumen zulässt, was Beziehungen zur Funktionalanalysis ins Spiel bringt. In allen Bereichen (Klassifikation und Strukturtheorie, Darstellungstheorie, etc.) spielt der Zusammenhang zwischen einer Lie Gruppe und ihrer Lie Algebra eine zentrale Rolle. Auf dem Niveau der Lie Algebren sind üblicherweise stärkere algebraische Methoden verfügbar, währen auf dem Niveau der Gruppen die analytischen und geometrischen Methoden dominieren.

Literaturverzeichnis

[Kühnel] W. Kühnel, "Matrizen und Lie-Gruppen" Eine geometrische Einführung, Springer Viehweg 2011.

[LinAlg] A. Čap, *Lineare Algebra und Geometrie*, Skriptum zum 3-semestrigen Vorlesungszyklus[1], Version WS14/15-15/16.

[LieGrp] A. Čap, *Lie Groups*, Skriptum[1] (Englisch), version fall term 2020/21.

[Riem] A. Čap, *Riemannian Geometry*, Skriptum[1] (Englisch), version fall term 2021/22.

[1] avail able online at https://www.mat.univie.ac.at/∼cap/lectnotes.html