Technische Universität Berlin
Fakultät II - Institut für Mathematik
Fachgebiet Algorithmische Algebra

Die erwartete Anzahl reeller Nullstellen von invarianten zufälligen Polynomen

Bachelorarbeit

im Studiengang Mathematik

vorgelegt von: Denes Burkhard Stolte

Matrikelnummer: 345410

Erstgutachter: Prof. Dr. Peter Bürgisser

Zweitgutachter: Prof. Dr. Jörg Liesen

Eidesstattliche Erklärung zur Bachelorarbeit

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Die selbstständige und eigenhändige Anfertigung versichert an Eides statt

Berlin, den

____

Denes Burkhard Stolte

Inhaltsverzeichnis

1 Einleitung **3**
1.1 Aufbau 5
1.2 Danksagung 6

2 Vorbereitungen **7**
2.1 Darstellungstheorie von Gruppen 7
2.2 Unitäre/Orthogonale G-Moduln 10
2.3 Lie-Algebren 11
2.4 Darstellungstheorie von Lie-Algebren 14
2.5 Die Lie-Algebra einer algebraischen Gruppe 15

3 Der Raum der homogenen Polynome **17**
3.1 Unitär invariantes Skalarprodukt 18
3.1.1 Irreduzibler $\mathcal{U}(n+1)$-Modul 20
3.1.2 Eindeutiges Skalarprodukt 24
3.2 Orthogonal invariante Skalarprodukte 25
3.2.1 Harmonische Polynome 26
3.2.2 Die Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln . 30
3.2.3 Die Legendre-Harmonischen 32
3.2.4 Irreduzibilität von $\mathcal{Y}_d^n(\mathbb{R})$ 34

4 Invariante normalverteilte Vektoren und Matrizen **38**
4.1 Invariante zufällige Vektoren 38
4.2 Invariante zufällige symmetrische Matrizen 39
4.3 Die erwartete Determinante einer standard-normalverteilten Matrix 39

5 Invariante zufällige Polynome **41**
5.1 Der Parameter eines invarianten zufälligen zentrierten Polynoms 42
5.2 Klassifikation invarianter zufälliger zentrierter Polynome . . 46
5.2.1 Klassifikation unitär invarianter zufälliger Polynome . 47
5.2.2 Klassifikation indefiniter Skalarprodukte 48
5.2.3 Klassifikation invarianter zentrierter Normalverteilungen 50

6 Die erwartete Anzahl reeller Nullstellen von Systemen invarianter zufälliger Polynome **55**
6.1 Haupttheorem 56
6.2 Die Rice-Formel 57
6.3 Beweis des Haupttheorems 58
6.4 Das Ergebnis von Shub & Smale 62

1 Einleitung

Die Frage nach der Lösung algebraischer Gleichungen existiert schon seit über 4000 Jahren. Die alten Babylonier waren zum Beispiel schon 2000 v. Chr. in der Lage, quadratische Gleichungen der Form $x^2 + q = px$ zu lösen (wenn auch nicht exakt), welche äquivalent zum Gleichungssystem $xy = q$ und $x + y = p$ sind. Betrachten wir ein allgemeines Polynom f vom Grad d mit Koeffizienten aus $\mathbb{R}$. Wohlbekannt ist dann unter dem ***Fundamentalsatz der Algebra***, dass das Polynom f über dem komplexen Zahlenkörper $\mathbb{C}$ genau d Nullstellen besitzt, wenn wir die Vielfachheiten mitzählen. Auch wenn wir zwar keine Aussage über das exakte Aussehen dieser Nullstellen haben, so wissen wir dennoch, dass jedes Polynom in einer Variablen vom Grad d, mit Koeffizienten in $\mathbb{R}$ oder $\mathbb{C}$, genau d-viele komplexe Nullstellen besitzt. Doch wie viele davon sind ***reell***?

Betrachten wir zum Beispiel das Polynom $h = aX^2 + bX + c$ vom Grad 2 in einer Variablen mit reellen Koeffizienten. Dann lernt man schon in der Schule, dass die Nullstellen vom Polynom h durch

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

gegeben sind. Schon bei diesem einfachen Beispiel beobachten wir, dass das Polynom h, abhängig vom Wert der sogenannten Diskriminante $D \coloneqq b^2 - 4ac$, unterschiedlich viele reelle Nullstellen besitzt. Ist die Diskriminante größer als Null, so besitzt h genau zwei verschiedene reelle Nullstellen. Gilt $D = 0$, so besitzt das Polynom h genau eine reelle Nullstelle mit Vielfachheit 2. Ist die Diskriminante kleiner als Null, so besitzt h sogar gar keine reellen Nullstellen, dafür aber 2 zueinander konjugierte komplexe Nullstellen. Die Frage nach der Anzahl reeller Nullstellen von Polynomen ist also eine deutlich schwierigere und soll zum zentralen Thema dieser Arbeit werden. Sie führt uns zu zufälligen Polynomen und deren erwarteter Anzahl reeller Nullstellen.

Sei $f = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0$ ein Polynom vom Grad d in einer Variablen mit Koeffizienten aus $\mathbb{R}$. Wählen wir die Koeffizienten a_i normalverteilt, so wird f zu einem zufälligen Polynom.

Aussagen über die erwartete Anzahl reeller Nullstellen von zufälligen Polynomen einer Variablen sind schon etwas länger bekannt. Bezeichnen wir mit $N^f(\mathbb{R})$ die Nullstellen von f, die in $\mathbb{R}$ liegen und wählen wir die Koeffizienten a_i unabhängig und standard-normalverteilt, so zeigte Kac [14], dass die erwartete Anzahl reeller Nullstellen $\mathbb{E}\big[N^f(\mathbb{R})\big]$ von f asymptotisch $\frac{2}{\pi}\ln(d)$ ist, für $d \to \infty$.

Eine zentrale Annahme, die es ermöglichte, die Frage nach der Anzahl reeller Nullstellen auf Systeme von Polynomen zu erweitern, machten Shub & Smale [13], nämlich die Invarianz der Verteilung unter der Gruppenaktion der orthogonalen Gruppe. Diese Annahme ist von einem geometrischen Standpunkt aus gesehen sehr natürlich und wird auch Gegenstand dieser Arbeit sein. Shub & Smale wählten die Koeffizienten a_i von f zentriert, unabhängig und normalverteilt mit Varianz $\binom{d}{i}$ und erhielten $\sqrt{d}$ für die erwartete Anzahl reeller Nullstellen von f. Man beobachtet also, dass die erwartete Anzahl reeller Nullstellen abhängig von der gewählten Wahrscheinlichkeitsverteilung ist.

Um die orthogonale Invarianz besser zu verstehen, betrachten wir das äquivalente "homogenisierte" Problem im eindimensionalen reellen projektiven Raum $\mathbb{P}^1$. D.h. wir homogenisieren das Polynom f und betrachten nun dessen Nullstellen im $\mathbb{P}^1$. Das homogenisierte Polynom ist dann ein homogenes Polynom vom Grad d in zwei Variablen, mit einer Normalverteilung, die nun invariant unter der Aktion der orthogonalen Gruppe $\mathcal{O}(2)$ ist.

Wir betrachten jetzt allgemeiner homogene Polynome vom Grad d in den $n+1$ Variablen $X_0, \ldots, X_n$ mit reellen Koeffizienten. Der Raum dieser Polynome, den wir mit $\mathcal{H}_d^n(\mathbb{R})$ bezeichnen, bildet einen reellen Vektorraum der Dimension $\binom{d+n}{d}$. Ein Polynom f aus $\mathcal{H}_d^n(\mathbb{R})$ können wir dann in der Form $f = \sum_{|\alpha|=d} f_\alpha X_0^{\alpha_0} \cdots X_n^{\alpha_n}$ darstellen. Sind die Koeffizienten f_α unabhängig zentriert und normalverteilt mit Varianz $\binom{d}{\alpha} := \frac{d!}{\alpha_0! \cdots \alpha_n!}$, so nennen wir ein solches zufälliges Polynom ***Kostlan-verteilt***. Shub & Smale [13] zeigten, dass für ein System $f_1 = 0, \cdots, f_n = 0$, von Kostlan-verteilten Polynomen $f_i \in \mathcal{H}_{d_i}^n(\mathbb{R})$ die erwartete Anzahl reeller Nullstellen gleich $\sqrt{d_1 \cdots d_n}$ ist. Dieses Resultat ähnelt sehr dem ***Satz von Bézout***. Dieser besagt, dass die Anzahl komplexer Nullstellen des Polynom-Systems $f_1 = 0, \cdots, f_n = 0$ von homogenen Polynomen im Allgemeinen gleich $d_1 \cdots d_n$ ist. Azaïs & Wschebor [11] haben später einen alternativen Beweis für das Ergebnis von Shub & Smale mithilfe der sogenannten ***Rice-Formel*** gefunden.

Ziel dieser Arbeit ist es, das Ergebnis von Shub & Smale mit der Beweismethode von Azaïs & Wschebor (Rice-Formel) zu verknüpfen, um das Ergebnis von Shub & Smale für allgemeine normalverteilte zentrierte Polynome zu verallgemeinern, ohne dabei die elegante Form des Resultats von Shub & Smale zu verlieren. Mit anderen Worten möchten wir die Anzahl reeller Nullstellen für ein System $f_1 = 0, \ldots, f_n = 0$ von unabhängigen zufälligen zentrierten Polynomen $f_i \in \mathcal{H}_{d_i}^n(\mathbb{R})$ mit normalverteilten Koeffizienten berechnen, deren Verteilung invariant unter der Aktion der orthogonalen Gruppe $\mathcal{O}(n+1)$ ist. Eine Schlüsselfunktion bildet dabei der ***Parameter*** $\delta(f)$ eines invarianten zufälligen zentrierten Polynoms f aus $\mathcal{H}_d^n(\mathbb{R})$ (Definition 5.6). Der Parameter eines Polynoms wurde erstmals von Podkorytov [15] betrachtet. Er ordnet dem Polynom f einen Wert zu, dessen Quadratwurzel $\sqrt{\delta(f)}$ wir als den erwarteten Grad des Polynoms f verstehen können. Podkorytov [15] ist sehr knapp gehalten und viele Beweise über den Parameter werden ausgelassen. In Bürgisser [9] wird der Parameter eines invarianten zufälligen zentrierten Polynoms wieder aufgegriffen und unter anderem werden die Resultate über den Parameter aus [15] bewiesen. Für Kostlan-verteilte Polynome $f \in \mathcal{H}_d^n(\mathbb{R})$ gilt $\delta(f) = d$.

Ein besonderer Fokus dieser Arbeit liegt dabei auf dem Verständnis invarianter Normalverteilungen und deren Einfluss auf den Parameter.

Als Hauptergebnis dieser Arbeit erhalten wir das folgende Theorem (vgl. Theorem 6.2).

Theorem (Hauptheorem). *Sei $f = (f_1, \cdots, f_n) = 0$ ein System von invarianten unabhängigen zentrierten normalverteilten Polynomen f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$. Dann erhalten wir $\sqrt{\delta(f_1)\cdots\delta(f_n)}$ für die erwartete Anzahl reeller Nullstellen des Systems $f = 0$.*

Wir beobachten die direkte Ähnlichkeit zum Ergebnis von Shub & Smale und für Kostlan-verteilte Polynome erhalten wir genau $\sqrt{d_1 \cdots d_n}$ für die erwartete Anzahl reeller Nullstellen.

1.1 Aufbau

Der Aufbau der Arbeit gestaltet sich folgendermaßen. In Kapitel 2 geben wir zunächst eine kurze Einführung in die Darstellungstheorie von Gruppen und Lie-Algebren. Ein wichtiges Ergebnis ist das Lemma von Schur (Lemma 2.5), welches (bis auf Multiplikation mit einem positiven Skalar) uns eineutige Skalarprodukte auf irreduziblen G-Moduln liefert. Außerdem erhalten wir eine Aussage darüber, unter welchen Bedingungen G-Moduln irreduzibel sind (Satz 2.34). Diese Ergebnisse nutzen wir dann in Kapitel 3, um zu zeigen, dass der komplexe Vektorraum $\mathcal{H}^n_d(\mathbb{C})$ der homogenen Polynome vom Grad d in $n+1$ Variablen bezüglich der natürlichen Aktion der unitären Gruppe $\mathcal{U}(n+1)$ ein irreduzibler $\mathcal{U}(n+1)$-Modul ist. Mit anderen Worten bedeutet dies: Der Vektorraum $\mathcal{H}^n_d(\mathbb{C})$ besitzt keinen nicht-trivialen linearen Unterraum, der $\mathcal{U}(n+1)$-invariant ist. Mit dem Lemma von Schur erhalten wir dann, dass $\mathcal{O}(n+1)$-invariante Skalarprodukte auf $\mathcal{H}^n_d(\mathbb{C})$ (bis auf Multiplikation mit einem positiven Skalar) eindeutig sind.

Viel interessanter gestaltet sich der reelle Vektoraum $\mathcal{H}^n_d(\mathbb{R})$ der homogenen Polynome vom Grad d in $n+1$ Variablen. Für diesen finden wir heraus, dass $\mathcal{H}^n_d(\mathbb{R})$ in mehrere irreduzible $\mathcal{O}(n+1)$-Moduln zerfällt, auf denen jeweils ein (bis auf Multiplikation mit einem positiven Skalar) eindeutiges Skalarprodukt existiert. Damit erhalten wir dann, dass $\mathcal{O}(n+1)$-invariante Skalarprodukte auf $\mathcal{H}^n_d(\mathbb{R})$ durch positive Gewichte parametrisiert sind. Die Resultate aus Kapitel 3 greifen wir in Kapitel 5 wieder auf, um $\mathcal{O}(n+1)$-invariante zentrierte Normalverteilungen zu klassifizieren. Dazu sei bemerkt, dass (positiv definite) $\mathcal{O}(n+1)$-invariante Skalarprodukte auf $\mathcal{H}^n_d(\mathbb{R})$ zu invarianten zentrierten Normalverteilungen auf $\mathcal{H}^n_d(\mathbb{R})$ korrespondieren.

In Kapitel 4 erarbeiten wir Eigenschaften invarianter normalverteilter Vektoren und Matrizen. Diese Eigenschaften werden wir für den Beweis des Haupttheorems wieder gebrauchen, da ein Vektor $f = (f_1, \dots, f_n)$ von invarianten zufälligen Polynomen f_i einem invarianten zufälligen Vektor entspricht.

Den Parameter eines Polynoms und dessen Eigenschaften betrachten wir in Abschnitt 5.1. Insbesondere berechnen wir den Parameter in Abhängigkeit der Gewichte, die $\mathcal{O}(n+1)$-invariante Skalarprodukte parametrisieren (und damit auch invariante zentrierte Normalverteilungen).

Zu guter Letzt formulieren wir das Haupttheorem in Kapitel 6. Außerdem beweisen wir das Ergebnis von Shub & Smale mithilfe der in dieser Arbeit vorgestellten Methoden.

1.2 Danksagung

Ich möchte meinem Betreuer Prof. Dr. Peter Bürgisser für seine Hilfe bei der Erstellung meiner Arbeit danken. Lehrreich und auch beeindruckend finde ich die Herangehensweise von Prof. Bürgisser an manche Probleme, sowie seine Erfahrung, an bestimmten Stellen die richtigen (und wichtigen) Fragen zu stellen. Vor allem möchte ich mich aber auch dafür bedanken, dass Prof. Bürgisser immer versucht hat, mir auch Sachverhalte über den Horizont meiner Arbeit aufzuzeigen und zu erklären.

Auch möchte ich mich bei meinen Eltern bedanken, die mir trotz ihrer geographischen Entfernung immer nah waren und mich unterstützt haben.

2 Vorbereitungen

Wir möchten hier eine elementare Einführung in die Darstellungstheorie von Gruppen und Lie-Algebren geben. Die Darstellungstheorie ermöglicht es, abstrakte algebraische Strukturen in der Sprache der linearen Abbildungen von Vektorräumen zu untersuchen. Wir werden sehen, dass Gruppen und Lie-Algebren eng miteinander verbunden sind. Wir folgen hier weitestgehend Teleman [1], sowie Humphreys [3] und [4].

2.1 Darstellungstheorie von Gruppen

Haben wir ganz allgemein eine Gruppe G und eine Menge X, so sagen wir G ***agiert*** auf X, falls es eine Abbildung $\rho : G \times X \to X$ gibt, die mit der Gruppenoperation auf G kompatibel ist, d.h.

$$\rho(g_1, \rho(g_2, x)) = \rho(g_1 g_2, x) \text{ und } \rho(1_G, x) = x.$$

Hier bezeichnet 1_G das neutrale Element der Gruppe G. Wir schreiben $g.x$ oder einfach gx und meinen $\rho(g, x)$. Die Abbildung ρ nennen wir dann eine ***Gruppenaktion*** von G auf X.

In der Geometrie hat man geometrische Objekte X und möchte wissen, welches die Symmetriegruppen dieser Objekte sind, das heißt welche Abbildungen gewisse geometrische Eigenschaften erhalten. In der Darstellungstheorie hat man eine Gruppe G und möchte wissen, auf welchen Objekten X die Gruppe G agiert und man versucht diese (bis auf Isomorphie) zu klassifizieren ([1]). Wir beschränken uns nun auf den Fall, wo X ein Vektorraum über einem Körper ist.

Definition 2.1. Sei V ein Vektorraum über einem Körper K und sei G eine Gruppe. Dann nennen wir die Abbildung

$$\rho : G \to GL(V)$$

eine ***(lineare) Darstellung*** von G auf V, falls ρ ein Gruppenhomomorphismus ist. Die Menge $GL(V)$ bezeichnet die Menge aller Automorphismen von V.

Beispiel 2.2. Sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum. Wählen wir eine Basis, so können wir die linearen Abbildungen $V \to V$ durch Matrizen darstellen. Die Menge $GL(V)$ enspricht dann der Menge der invertierbaren $n \times n$-Matrizen mit Einträgen aus $\mathbb{C}$. Diese bildet zusammen mit der Matrixmultiplikation eine Gruppe, die wir speziell mit $GL(n, \mathbb{C})$ bezeichnen. Für $G = GL(n, \mathbb{C})$ ist dann $GL(n, \mathbb{C}) \to GL(n, \mathbb{C})$, $A \mapsto A$, eine Darstellung von $GL(n, \mathbb{C})$ auf V.

Das Setting, in dem wir im Weiteren arbeiten möchten, ist das der ***Moduln***. Haben wir eine Darstellung $\rho : G \to GL(V)$ einer Gruppe G auf einem Vektorraum V über K, so ist $\rho(g)$ aus $GL(V)$ für alle g aus G und wir können eine Gruppenaktion $G \times V \to V$, $(g, v) \mapsto g.v$, durch $g.v \coloneqq \rho(g)(v)$

definieren. Mit anderen Worten bedeutet dies: Wir können die Darstellung als eine Gruppenaktion von G auf V auffassen, die die Vektorraumeigenschaften erhält. Diese Eigenschaft werden wir verwenden, um G-Moduln zu definieren.

Definition 2.3. Sei V ein Vektorraum über einen Körper K und G sei eine Gruppe. Dann nennen wir V einen ***G-Modul***, falls es eine Abbildung $G \times V \to V$, $(g, v) \mapsto g.v$, mit folgenden Eigenschaften gibt:

(i) $(g_1 g_2).v = g_1.(g_2.v)$

(ii) $1_G.v = v$

(iii) $g.(\lambda v_1 + v_2) = \lambda g.v_1 + g.v_2$

für alle g, g_1, g_2 aus G, v, v_1, v_2 aus V und λ aus K. Ist V ein G-Modul und U ein Untervektorraum von V, so nennen wir U einen ***G-Untermodul***, falls

$$g.U := \{g.u \mid u \in U\} \subseteq U$$

für alle g aus G gilt. Wir sagen dann auch, dass der Untervektorraum U ***G-invariant*** ist. Außerdem nennen wir den G-Modul V ***irreduzibel***, wenn $\{0\}$ und V die einzigen G-Untermoduln von V sind. Andernfalls nennen wir den G-Modul V ***reduzibel***. Ist V ein endlichdimensionaler Vektorraum, so nennen wir auch den G-Modul V ***endlichdimensional***.

Bemerkung 2.4. Sei V ein Vektorraum über einem Körper K und G eine Gruppe. Angenommen, wir haben eine Darstellung $\rho : G \to GL(V)$ der Gruppe G auf V. Dann gilt $\rho(g) \in GL(V)$ für alle $g \in G$. Definieren wir die Abbildung $G \times V \to V$, $(g, v) \mapsto g.v$ für $g.v := \rho(g)(v)$, so wird der Vektorraum V zu einen G-Modul. Sei umgekehrt der Vektorraum V ein G-Modul. Dann definiert $\rho : G \to GL(V)$, $\rho(g) := \varphi_g$ mit $\varphi_g(v) := g.v$ für $v \in V$ und $g \in G$ eine Darstellung der Gruppe G auf V. Daher sind Modul und Darstellung austauschbare Begriffe.

Seien V und W zwei G-Moduln, V und W Vektorräume über einem Körper K und $\varphi : V \to W$ eine Abbildung. Dann nennen wir φ einen ***G-Modul-Homomorphismus***, falls φ K-linear ist und $\varphi(g.v) = g.\varphi(v)$ für alle v aus V und g aus G gilt. Die Menge der G-Modul-Homomorphismen von V nach W bezeichnen wir mit $\operatorname{Hom}_G(V, W)$. Ist $\varphi : V \to W$ ein bijektiver G-Modul-Homomorphismus, so nennen wir φ einen ***G-Modul-Isomorphismus***. In diesem Fall sagen wir, dass die G-Moduln V und W ***isomorph*** zueinander sind und schreiben $V \cong W$.

Wir kommen zur ersten Aussage über irreduzible G-Moduln: Haben wir einen irreduziblen G-Modul V über $\mathbb{C}$, so besteht die Menge der G-Modul-Endomorphismen $\operatorname{End}_G(V) := \operatorname{Hom}_G(V, V)$ aus Vielfachen der Identität.

Lemma 2.5 (Lemma von Schur)**.** *Sei V ein endlichdimensionaler, irreduzibler G-Modul über $\mathbb{C}$. Dann ist jeder G-Modul-Homomorphismus $\varphi : V \to V$ skalar, d.h. $\varphi = \lambda \operatorname{id}_V$ für ein $\lambda \in \mathbb{C}$.*

Beweis. Da der zugrunde liegende Körper $\mathbb{C}$ ist, besitzt φ einen Eigenwert $\lambda \in \mathbb{C}$ mit $\varphi(v) = \lambda\, v$ für ein v aus V verschieden von Null.

Behauptung: Der Eigenraum $E_\lambda := \{v \in V \mid \varphi(v) = \lambda v\}$ ist G-invariant und verschieden vom Nullraum.
Beweis: Zunächst ist $E_\lambda \neq \{0\}$, da φ einen Eigenvektor $v \neq 0$ zum Eigenwert λ enthält. Sei v aus E_λ, dann gilt

$$\varphi(g.v) = g.\varphi(v) = g.(\lambda v) = \lambda\, g.v$$

für alle g aus G. Also ist $g.v \in E_\lambda$ für alle $g \in G$. ■

Wegen $E_\lambda \neq \{0\}$ und der Irreduzibilität folgt $E_\lambda = V$ und damit $\varphi = \lambda\,\mathrm{id}$. □

Als nächstes möchten wir zeigen, dass G-Modul-Homomorphismen zwischen irreduziblen G-Moduln über $\mathbb{C}$ entweder bijektiv sind oder Null.

Korollar 2.6. *Seien V und W zwei endlichdimensionale G-Moduln über $\mathbb{C}$. Sind V und W irreduzibel, dann gilt:*

$$\dim \mathrm{Hom}_G(V, W) = \begin{cases} 1 & , \textit{ falls } V \cong W. \\ 0 & , \textit{ sonst.} \end{cases}$$

Beweis. Sei $\varphi \in \mathrm{Hom}_G(V, W)$ verschieden von Null. Dann ist $\ker \varphi$ nicht ganz V und G-invariant. Die Irreduzibilität von V impliziert $\ker \varphi = \{0\}$, d.h. φ ist injektiv. Analog ist das Bild $\mathrm{im}\, \varphi$ verschieden von $\{0\}$ und G-invariant. Die Irreduzibilität von W impliziert, dass das Bild $\mathrm{im}\, \varphi$ gleich W ist, d.h. φ ist surjektiv. Damit ist φ bijektiv. Wir erhalten: Es gibt genau dann ein $\varphi \in \mathrm{Hom}_G(V, W)$ verschieden von Null, wenn V und W isomorph sind.

Seien also V und W isomorph. Wir wählen ein φ aus $\mathrm{Hom}_G(V, W)$ verschieden von Null. Ist auch $\tilde{\varphi} \in \mathrm{Hom}_G(V, W)$, so ist $\tilde{\varphi} \circ \varphi^{-1} \in \mathrm{Hom}_G(V, V)$. Mit dem Lemma von Schur (Lemma 2.5) folgt $\tilde{\varphi} \circ \varphi^{-1} = \lambda\, \mathrm{id}_V$ für ein $\lambda \in \mathbb{C}$ und wir erhalten

$$\tilde{\varphi} = \lambda\, \varphi. \tag{1}$$

D.h. es gilt $\mathrm{Hom}_G(V, W) = \mathbb{C}\, \varphi$. □

Das Lemma von Schur (Lemma 2.5) gilt auch in ähnlicher Weise für irreduzible reelle G-Moduln.

Bemerkung 2.7. Sei V ein endlichdimensionaler G-Modul über $\mathbb{R}$ und $\varphi : V \to V$ ein G-Modul-Homomorphismus. Angenommen, φ besitze einen reellen Eigenwert λ, dann gilt $\varphi = \lambda\, \mathrm{id}$.

Wir haben eine allgemeine Definition für G-Moduln angegeben. Jetzt möchten wir diese Definition etwas einschränken, indem wir speziell Vektorräume über $\mathbb{C}$ oder $\mathbb{R}$ betrachten.

2.2 Unitäre/Orthogonale G-Moduln

Im Folgenden sei V ein Vektorraum über $\mathbb{C}$ ($\mathbb{R}$) und G eine Gruppe.

Definition 2.8. Sei V ein G-Modul und V sei mit einem Skalarprodukt $\langle \cdot,\cdot \rangle : V \times V \to \mathbb{C}$ ($\mathbb{R}$) ausgestattet. Dann nennen wir den G-Modul V ***unitär*** (***orthogonal***), falls das Skalarprodukt G-invariant ist,
d.h. falls $\langle g.v_1, g.v_2 \rangle = \langle v_1, v_2 \rangle$ für alle g aus G und alle v_1, v_2 aus V.

Für endlichdimensionale komplexe bzw. reelle Vektorräume V mit Skalarprodukt wissen wir: Ist U ein Untervektorraum von V, so ist auch das Vektorraumkomplement $U^\perp$ ein Untervektorraum von V. Desweiteren ist V die direkte Summe von U und $U^\perp$. Wir werden nun zeigen, dass etwas Vergleichbares auch für endlichdimensionale unitäre bzw. orthogonale G-Moduln gilt.

Proposition 2.9. *Sei V ein endlichdimensionaler unitärer (orthogonaler) G-Modul und $U \subseteq V$ ein G-Untermodul. Dann ist auch $U^\perp$ ein G-Untermodul von V.*

Beweis. Wir müssen zeigen, dass $g.u^\perp \in U^\perp$ ist, für alle $u^\perp$ aus $U^\perp$ und g aus G. Da nach Voraussetzung der G-Modul V unitär (orthogonal) ist, ist die Gruppenaktion injektiv, denn sei $g.v = 0$ für ein v aus V und g aus G, so folgt

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

also $v = 0$. Da außerdem der Vektorraum V endlichdimensional ist, ist nach dem Rangsatz aus der linearen Algebra [2, Kap. 2.2, Kor. 3], die Gruppenaktion surjektiv. Schränken wir diese auf U ein, so haben wir $g.U = U$ für alle g aus G. Seien u aus U und g aus G. Dann gibt es aufgrund der Surjektivität ein u' aus U, sodass $u = g.u'$. Sei weiter $u^\perp$ aus $U^\perp$, dann erhalten wir

$$\langle g.u^\perp, u \rangle = \langle g.u^\perp, g.u' \rangle = \langle u^\perp, u' \rangle = 0.$$

Also gilt $g.u^\perp \in U^\perp$. Da u, $u^\perp$ und g beliebig gewählt sind, folgt, dass $U^\perp$ ein G-Untermodul von V ist. □

Bemerkung 2.10. Proposition 2.9 gilt für alle G-Moduln V, falls die Gruppe G endlich ist. Dabei kann V ein $\mathbb{C}$ ($\mathbb{R}$)-Vektorraum sein oder ein Vektorraum über einem endlichen Körper, dessen Charakteristik nicht die Ordnung der Gruppe G teilt. Diese Aussage ist auch bekannt als ***Satz von Maschke***.

Aus der Proposition 2.9 erhalten wir unmittelbar folgendes Korollar über die Zerlegung von unitären (orthogonalen) G-Moduln in irreduzible.

Korollar 2.11. *Sei V ein endlichdimensionaler unitärer (orthogonaler) G-Modul. Dann gibt es endlich viele irreduzible G-Untermoduln $U_1, \ldots, U_r$ von V, sodass*

$$V = U_1 \oplus \cdots \oplus U_r. \tag{2}$$

Beweis. Wir nehmen an, V sei ein reduzibler G-Modul, denn sonst ist nichts zu zeigen. Sei also U ein echter G-Untermodul von V und verschieden vom Nullraum. Nach Proposition 2.9 ist $U^\perp$ ebenfalls ein G-Untermodul und es gilt $V = U \oplus U^\perp$. Sind U oder $U^\perp$ reduzibel, so fahren wir induktiv für U oder $U^\perp$ fort. Da V endlichdimensional ist, sind wir nach endlich vielen Schritten fertig. □

Einen G-Modul V, den wir in endlich viele irreduzible G-Untermoduln zerlegen können, nennen wir auch ***vollständig reduzibel***. Die Frage nach der Eindeutigkeit der Zerlegung in (2), führt uns zu sogenannten ***isotypischen Zerlegungen***.

Bemerkung 2.12 (Isotypische Zerlegung). Sei V ein vollständig reduzibler G-Modul. Wir können also V in eine direkte Summe von endlich vielen irreduziblen G-Untermoduln von V zerlegen. Wir gruppieren nun irreduzible Summanden und schreiben

$$V \cong \bigoplus_{k=1}^{r} V_k, \text{ wobei } V_k \cong \underbrace{W_k \oplus \cdots \oplus W_k}_{n_k - mal}$$

und die W_k sind irreduzible G-Untermoduln von V mit $W_i \ncong W_j$ für $i \neq j$. Wir schreiben dann auch

$$V \cong \bigoplus_{k=1}^{r} W_k^{\oplus n_k} \tag{3}$$

und nennen n_k die ***Vielfachheit*** von W_k in V. Die Zerlegung (3) nennen wir die Zerlegung von V in ***isotypische Komponenten*** V_k oder einfach ***isotypische Zerlegung*** von V. Dann gilt:

(i) Jeder G-Untermodul von V, der isomorph zu W_k ist, ist in V_k enthalten. Mit anderen Worten bedeutet dies: Jeder irreduzible G-Untermodul von V ist in einem der V_k enthalten.

(ii) Die isotypische Zerlegung von V ist eindeutig, d.h. sie hängt nicht von der ursprünglichen Zerlegung in irreduzible Komponenten ab.

Für einen Beweis der Eindeutigkeit der isotypischen Zerlegung von V verweisen wir auf [1, Thm. 5.4].

2.3 Lie-Algebren

Lie-Algebren kommen in natürlicher Weise als Vektorräume von linearen Abbildungen vor, die mit einer Operation $[x, y] = xy - yx$ ausgestattet sind, wobei x und y lineare Abbildungen sind (mit der Verknüpfung als Multiplikation auf der rechten Seite). Diese Operation ist im Allgemeinen weder kommutativ noch assoziativ.

Definition 2.13. Einen Vektorraum V über einem Körper K, zusammen mit einer Operation $V \times V \to V$, $(v_1, v_2) \mapsto [v_1, v_2]$, auch ***Lie-Klammer*** von v_1 und v_2 genannt, nennen wir ***Lie-Algebra*** über K, falls die folgenden drei Eigenschaften erfüllt sind.

(i) Die Klammeroperation ist bilinear.

(ii) $[v, v] = 0$ für alle v aus V.

(iii) $[v_1, [v_2, v_3]] + [v_2, [v_3, v_1]] + [v_3, [v_1, v_2]] = 0$ für alle v_1, v_2, v_3 aus V.

Einen Unterraum U einer Lie-Algebra V nennen wir ***Lie-Unteralgebra*** von V, falls für alle u_1, u_2 aus U auch $[u_1, u_2]$ aus U ist. Eine lineare Abbildung $\varphi : V \to W$ zwischen zwei Lie-Algebren V und W nennen wir ***Lie-Algebren-Homomorphismus***, falls $\varphi([v_1, v_2]) = [\varphi(v_1), \varphi(v_2)]$ für alle v_1, v_2 aus V gilt. Ist φ sogar bijektiv, so nennen wir φ einen ***Lie-Algebren-Isomorphismus***.

Bemerkung 2.14. Aus der Eigenschaft (i) und (iii) aus Definition 2.13, angewendet auf $[x + y, x + y]$, erhalten wir Anti-Kommutativität, d.h. es gilt $[x, y] = -[y, x]$.

Vereinbarung: Im Folgenden werden wir, falls nichts anderes gesagt wird, nur Lie-Algebren betrachten, deren zugrunde liegender Vektorraum ein endlichdimensionaler Vektorraum über $\mathbb{C}$ bzw. $\mathbb{R}$ ist.

Das nächste Beispiel zeigt, dass jeder endlichdimensionale Vektorraum eine Lie-Algebra definiert.

Beispiel 2.15. Sei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$ ($\mathbb{R}$). Dann können wir V durch $[v_1, v_2] := 0$ für alle v_1, v_2 aus V als Lie-Algebra auffassen. Dabei nennen wir Lie-Algebren mit trivialer Lie-Klammer ***abelsch***.

Sei V ein n-dimensionaler Vektorraum über $\mathbb{C}$ ($\mathbb{R}$). Wir bezeichnen mit $\mathrm{End}(V)$ die Menge aller linearen Abbildungen $V \to V$. Wir wissen, dass die Vektorraumdimension von $\mathrm{End}(V)$ gleich n^2 ist. Wir definieren eine Lie-Klammer $[v_1, v_2] := v_1 v_2 - v_2 v_1$ für v_1, v_2 aus $\mathrm{End}(V)$, mit der Verknüpfung als Multiplikation auf der rechten Seite. Damit wird $\mathrm{End}(V)$ zu einer Lie-Algebra über $\mathbb{C}$ ($\mathbb{R}$). Diese Lie-Algbera bezeichnen wir mit $\mathfrak{gl}(V)$ und nennen sie ***allgemeine lineare Lie-Algebra***, um sie vom Vektorraum $\mathrm{End}(V)$ zu unterscheiden. Dabei nennen wir jede Unteralgebra der Lie-Algebra $\mathfrak{gl}(V)$ ***lineare Lie-Algebra***.

Legen wir eine Basis von V fest, so können wir $\mathfrak{gl}(V)$ als die Menge der $n \times n$-Matrizen über $\mathbb{C}$ ($\mathbb{R}$) identifizieren. Diese bezeichnen wir speziell mit $\mathfrak{gl}(n, \mathbb{C})$ ($\mathfrak{gl}(n, \mathbb{R})$). Aus der Verknüpfung linearer Abbildungen wird dann die Matrixmultiplikation. Für die Standardbasismatrizen $\mathbb{E}_{ij}$, mit einer 1 an der Stelle (i, j) und 0 sonst, verhält sich die Lie-Klammer wie folgt:

$$[\mathbb{E}_{ij}, \mathbb{E}_{kl}] = \delta_{jk}\mathbb{E}_{il} - \delta_{li}\mathbb{E}_{kj}.$$

Dabei bezeichnet δ_{ij} das Kronecker-Delta.

Definition 2.16. Einen Unterraum U einer Lie-Algebra V nennen wir ***Ideal*** von V, falls für $v \in V$ und $u \in U$ auch $[v, u] \in U$ ist.

Ideale von Lie-Algebren spielen eine ähnliche Rolle wie Normalteiler bei Gruppen oder zweiseitige Ideale bei Ringen. Wegen Bemerkung 2.14 brauchen wir nicht zwischen linksseitigen und rechtsseitigen Idealen zu unterscheiden. Jede Lie-Algebra V enthält die trivialen Ideale $\{0\}$ und V selbst und jedes Ideal ist (im Gegensatz zu Ringen) eine Unteralgebra von V. Ist $\varphi : V \to W$ ein Lie-Algebren-Homomorphismus, so kann man leicht zeigen, dass das Bild $\operatorname{im}\varphi$ ein Ideal von W und der Kern $\ker\varphi$ ein Ideal von V ist. Wie bei Gruppen oder Ringen können wir Quotienten von Lie-Algebren und Idealen bilden und erhalten wieder eine Lie-Algebra. Sei also V eine Lie-Algebra und I ein Ideal von V. Dann wird V/I mit der Lie-Klammer $[v_1 + I, v_2 + I] = [v_1, v_2] + I$ für $v_1, v_2 \in V$ zu einer Lie-Algebra, die wir die ***Faktoralgebra*** von V und I nennen.

Definition 2.17. Sei V eine Lie-Algebra. Die Menge

$$[V, V] := \operatorname{span}\{[v_1, v_2] \mid v_1, v_2 \in V\}$$

nennen wir ***abgeleitete Lie-Algebra*** von V.

Die von V abgeleitete Lie-Algebra $[V, V]$ ist ein Ideal von V. Außerdem ist die Lie-Algebra V genau dann abelsch, wenn $[V, V] = \{0\}$ gilt. Hat V nur die trivialen Ideale und ist V nicht-abelsch, so nennen wir die Lie-Algebra V ***einfach***.

Bemerkung 2.18. Die abgeleitete Lie-Algebra $[V, V]$ einer Lie-Algebra V ist das kleinste Ideal von V mit abelschem Quotienten.

Wir definieren eine Reihe von Idealen von V durch

$$V^{(0)} = V,\ V^{(1)} = [V, V],\ V^{(2)} = [V^{(1)}, V^{(1)}], \dots, V^{(i)} = [V^{(i-1)}, V^{(i-1)}].$$

Diese Reihe von Idealen nennen wir ***abgeleitete Reihe*** von V. Weiter nennen wir V ***auflösbar***, falls $V^{(m)} = \{0\}$ für ein $m \geq 1$ gilt.

Beispiel 2.19. Die Lie-Algebra $\mathfrak{t}(n, \mathbb{C})$ der oberen Dreieckmatrizen mit Einträgen aus $\mathbb{C}$ ist ausflösbar.

Beweis. Seien $\mathbb{E}_{ij}$ für $1 \leq i, j \leq n$, die Standardbasismatrizen von $\mathfrak{gl}(n, \mathbb{C})$, so ist $\mathfrak{t}(n, \mathbb{C}) = \operatorname{span}\{\mathbb{E}_{ij} \mid i \leq j\}$. Weiter definieren wir die Lie-Algebra $\mathfrak{n}(n, \mathbb{C})$ der strikten oberen komplexen Dreiecksmatrizen $\mathfrak{n}(n, \mathbb{C}) = \operatorname{span}\{\mathbb{E}_{ij} \mid i < j\} = \operatorname{span}\{\mathbb{E}_{ij} \mid i - j \geq 1\}$ und die abelsche Lie-Algebra $\mathfrak{o}(n, \mathbb{C}) = \operatorname{span}\{\mathbb{E}_{ij} \mid i = j\}$ der komplexen Diagonalmatrizen. Dann können wir $\mathfrak{t}(n, \mathbb{C})$ als Summe von $\mathfrak{n}(n, \mathbb{C})$ und $\mathfrak{o}(n, \mathbb{C})$ darstellen.

Behauptung: $\mathfrak{n}(n, \mathbb{C})$ ist auflösbar.
Beweis: Für $i < j$, $k < l$ und o.B.d.A. $i \neq l$ haben wir $[\mathbb{E}_{ij}, \mathbb{E}_{kl}] = \delta_{jk}\mathbb{E}_{il}$. Daher haben wir $[\mathfrak{n}(n, \mathbb{C}), \mathfrak{n}(n, \mathbb{C})] = \operatorname{span}\{\mathbb{E}_{ij} \mid i - j \geq 2\}$. Durch Induktion folgt, dass es ein $m \in \mathbb{N}$ gibt, mit $\mathfrak{n}(n, \mathbb{C})^{(m)} = \{0\}$. ■

Insgesamt erhalten wir aus $[\mathfrak{o}(n, \mathbb{C}), \mathfrak{o}(n, \mathbb{C})] = 0$ und $[\mathfrak{o}(n, \mathbb{C}), \mathfrak{n}(n, \mathbb{C})] = \mathfrak{n}(n, \mathbb{C})$, dass $[\mathfrak{t}(n, \mathbb{C}), \mathfrak{t}(n, \mathbb{C})] = \mathfrak{n}(n, \mathbb{C})$ gilt. Damit ist mit $\mathfrak{n}(n, \mathbb{C})$ auch $\mathfrak{t}(n, \mathbb{C})$ auflösbar. □

Lemma 2.20. *Sei V eine Lie-Algebra, L eine auflösbare Lie-Algebra und $\varphi : L \to V$ ein Lie-Algebren-Homomorphismus. Dann ist auch $\varphi(L)$ auflösbar.*

Beweis. Wir können annehmen, dass φ surjektiv ist. Dann gilt:

$$\varphi(L^{(1)}) = \varphi([L, L]) = [\varphi(L), \varphi(L)] = [V, V] = V^{(1)}.$$

Fahren wir induktiv fort, so erhalten wir

$$\varphi(L^{(i)}) = V^{(i)}.$$

Somit ist mit L auch $\varphi(L)$ auflösbar. □

Der folgende Satz sichert uns die Existenz eines gemeinsamen Eigenvektors für alle Elemente einer auflösbaren Lie-Algebra über $\mathbb{C}$.

Satz 2.21 (Satz von Lie, [3, Ch. 4.1, Theorem])**.** *Sei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$, $\dim(V) \geq 1$ und L eine auflösbare Unteralgebra von $\mathfrak{gl}(V)$. Dann gibt es einen von Null verschiedenen Vektor $v \in V$, der gemeinsamer Eigenvektor für alle $x \in L$ ist, d.h. $x(v) \in \operatorname{span}\{v\}$ für alle $x \in L$.*

Für einen Beweis verweisen wir ebenfalls auf [3, Ch. 4.1, Theorem].

Bemerkung 2.22. Häufig findet man unter dem Satz von Lie auch eine andere, aber äquivalente Aussage: Mit den Voraussetzungen von Satz 2.21 folgt, dass es eine Basis von V gibt, bezüglich der alle Elemente von L durch obere Dreiecksmatrizen dargestellt werden können.

2.4 Darstellungstheorie von Lie-Algebren

Analog zu Darstellungen von Gruppen können wir Darstellungen von Lie-Algebren betrachten.

Definition 2.23. Sei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$ ($\mathbb{R}$) und L eine Lie-Algebra. Dann nennen wir einen Lie-Algebren-Homomorphismus

$$\rho : L \to \mathfrak{gl}(V)$$

Darstellung von L.

Beispiel 2.24. Sei L eine Lie-Algebra. Dann ist die Abbildung $\operatorname{ad} : L \to \mathfrak{gl}(L)$, $\operatorname{ad}(x)(y) := [x, y]$, eine Darstellung der Lie-Algebra L.

Anstatt von Darstellungen von Lie-Algebren zu sprechen, möchten wir (wie bei den Gruppen) die äquivalente Sprache der Lie-Algebren-Moduln nutzen.

Definition 2.25. Sei L eine Lie-Algebra. Ein ***Modul*** über L, kurz ***L-Modul***, ist ein endlichdimensionaler Vektorraum V über $\mathbb{C}$ ($\mathbb{R}$), versehen mit einer bilinearen Operation $L \times V \to V$, $(x, v) \mapsto x.v$, sodass

$$[x, y].v = x.(y.v) - y.(x.v)$$

für alle x, y aus L und alle v aus V gilt.

Ein ***L-Untermodul*** U von V ist ein Unterraum von V, der invariant unter L ist, d.h. für alle $x \in L$ und alle $u \in U$ ist $x.u \in U$. Hat V nur die trivialen L-Untermoduln $\{0\}$ und V und ist V verschieden vom Nullraum, so nennen wir V ***irreduzibel***. Andernfalls nennen wir V ***reduzibel***. Wir nennen V ***vollständig reduzibel***, falls V eine direkte Summe von irreduziblen L-Untermoduln ist.

Als Resultat aus dem Satz von Lie (Satz 2.21) erhalten wir, dass jeder Modul über einer auflösbaren Lie-Algebra einen eindimensionalen Untermodul enthält.

Proposition 2.26. *Sei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$, $\dim(V) \geq 1$ und L eine auflösbare Lie-Algebra über $\mathbb{C}$. Sei V weiter ein L-Modul mit der Darstellung $\rho : L \to \mathfrak{gl}(V)$, $\rho(x)(v) := x.v$. Dann besitzt V einen eindimensionalen L-Untermodul.*

Beweis. Nach Lemma 2.20 ist $\rho(L) \subseteq \mathfrak{gl}(V)$ auflösbar. Weiter gibt es nach dem Satz von Lie (Satz 2.21) einen von Null verschiedenen Vektor $v \in V$ mit $x.v = \rho(x)(v) \in \mathbb{C}\, v$ für alle $x \in L$. Mit anderen Worten bedeutet dies: Der eindimensionale Vektorraum $\mathbb{C}\, v$ ist ein L-Untermodul von V. □

2.5 Die Lie-Algebra einer algebraischen Gruppe

In diesem Abschnitt untersuchen wir die Lie-Algebra von algebraischen Gruppen. Hierzu folgen wir Kraft [5].

Die allgemeine lineare Gruppe $GL(n, \mathbb{C})$ besitzt in natürlicher Weise die Struktur einer affinen Varietät: Ist $\mathrm{Mat}(n, \mathbb{C})$ die Menge der komplexen $n \times n$-Matrizen, so gilt

$$GL(n, \mathbb{C}) = \{g \in \mathrm{Mat}(n, \mathbb{C}) \mid \det g \neq 0\}.$$

Eine ***algebraische Gruppe*** ist eine Zariski-abgeschlossene Untergruppe G von $GL(n, \mathbb{C})$.

Definition 2.27. Sei $G \subseteq GL(n, \mathbb{C})$ eine algebraische Gruppe. Wir definieren die Lie-Algebra $\mathrm{Lie}(G)$ einer algebraischen Gruppe G, als den Tangentialraum $T_e(G)$ von G im neutralem Element e.

Bemerkung 2.28. Das ***Matrixexponential***

$$\exp : Mat(n, \mathbb{C}) \to GL(n, \mathbb{C}), \quad A \mapsto \exp(A) := \sum_{i=0}^{\infty} \frac{A^i}{i!}$$

bildet die Menge $\mathrm{Mat}(n, \mathbb{C})$ surjektiv auf die Menge $GL(n, \mathbb{C})$ ab und es gilt:

(i) $\exp(0) = \mathrm{id}$.

(ii) $\exp(A)^{-1} = \exp(-A)$.

Beispiel 2.29. Wir betrachten die glatte Kurve

$$\gamma_A : \mathbb{R} \to GL(n, \mathbb{C}), \quad \gamma_A(t) = \exp(tA),$$

für A aus $\mathrm{Mat}(n, \mathbb{C})$. Dann gilt:

(i) $\gamma_A(t) \in GL(n, \mathbb{C})$ für alle $t \in \mathbb{R}$.

(ii) $\gamma_A(0) = \mathrm{id}$.

(iii) $\frac{d}{dt}\gamma_A(0) = \frac{d}{dt}\exp(tA)|_{t=0} = A\exp(tA)|_{t=0} = A$.

Insbesondere ist $\frac{d}{dt}\gamma_A(0) \in T_0\left(GL(n, \mathbb{C})\right)$.

Es ist also $\mathrm{Lie}(GL(n, \mathbb{C})) = \mathrm{Mat}(n, \mathbb{C}) = \mathfrak{gl}(n, \mathbb{C})$. Daher ist die Lie-Algebra $\mathrm{Lie}(G)$ einer algebraischen Gruppe $G \subseteq GL(n, \mathbb{C})$ ein Untervektorraum von $\mathrm{Mat}(n, \mathbb{C})$. Es gilt sogar: $\mathrm{Lie}(G)$ ist eine Lie-Unteralgebra von $\mathfrak{gl}(n, \mathbb{C})$.

Wir beschränken uns nun auf die Lie-Algebra $\mathfrak{gl}(n, \mathbb{C})$ der Gruppe $GL(n, \mathbb{C})$ und möchten näher auf den Zusammenhang zwischen $GL(n, \mathbb{C})$- und $\mathfrak{gl}(n, \mathbb{C})$-Untermoduln eingehen. Die Grundlage hierzu bildet die folgende Aussage:

Lemma 2.30. *Sei V ein $GL(n, \mathbb{C})$-Modul und $U \subseteq V$ sei ein Untervektorraum. Dann gilt: U ist genau dann ein $GL(n, \mathbb{C})$-Untermodul, wenn U ein $\mathfrak{gl}(n, \mathbb{C})$-Untermodul ist.*

Ein Beweis steht zum Beispiel in [5, Kap. II.2.5, Folgerung 3].

Wir brauchen also nicht zwischen $\mathfrak{gl}(n, \mathbb{C})$-Untermoduln und $GL(n, \mathbb{C})$-Untermoduln zu unterscheiden.

Bemerkung 2.31. Wir geben hier eine Idee zum Beweis von Lemma 2.30 an. Sei dazu U ein $GL(n, \mathbb{C})$-Untermodul mit Darstellung $\rho : GL(n, \mathbb{C}) \to GL(V)$. Dann gibt es eine Basis von $GL(V)$, in der $\rho(g) \in GL(V)$ für alle $g \in GL(n, \mathbb{C})$ Blockdiagonalform hat, die auch beim Übergang zur Lie-Algebra $\mathfrak{gl}(V)$ erhalten bleibt und umgekehrt. Die Darstellung der Lie-Algebra $\mathfrak{gl}(n, \mathbb{C})$ ist durch das Differential $D_e\rho : \mathfrak{gl}(n, \mathbb{C}) \to \mathfrak{gl}(V)$ gegeben.

Die Umkehrung von Lemma 2.30 liefert uns unmittelbar einen Zusammenhang zwischen der Irreduzibilität von $GL(n, \mathbb{C})$- und $\mathfrak{gl}(n, \mathbb{C})$-Moduln.

Korollar 2.32. *V ist als $GL(n, \mathbb{C})$-Modul genau dann irreduzibel, wenn V als $\mathfrak{gl}(n, \mathbb{C})$-Modul irreduzibel ist.*

Aus den Vorbereitungen über Lie-Algebren-Moduln erhalten wir ein allgemeines Prinzip über irreduzible Darstellungen von Lie-Algebren.

Satz 2.33. *Sei V ein vollständig reduzibler $\mathfrak{gl}(n,\mathbb{C})$-Modul, wobei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$ ist. Weiter sei $\mathfrak{t}(n,\mathbb{C})$ die Lie-Unteralgebra von $\mathfrak{gl}(n,\mathbb{C})$ der oberen Dreiecksmatrizen. Dann gilt: V ist genau dann ein irreduzibler $\mathfrak{gl}(n,\mathbb{C})$-Modul, wenn V genau eine $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade besitzt.*

Beweis. In Beispiel 2.19 haben wir gezeigt, dass die Lie-Algebra $\mathfrak{t}(n,\mathbb{C})$ auflösbar ist. Schränken wir den $\mathfrak{gl}(n,\mathbb{C})$-Modul V auf die Lie-Unteralgebra $\mathfrak{t}(n,\mathbb{C})$ ein, so wird V zu einem $\mathfrak{t}(n,\mathbb{C})$-Modul. Nach Proposition 2.26 besitzt V (und somit auch jeder $\mathfrak{t}(n,\mathbb{C})$-Untermodul von V) einen eindimensionalen $\mathfrak{t}(n,\mathbb{C})$-Untermodul, d.h. eine $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade.
$\boxed{\Leftarrow}$ Angenommen V besitzt genau eine $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade. Da V vollständig reduzibel ist, können wir V in eine direkte Summe von irreduziblen $\mathfrak{gl}(n,\mathbb{C})$-Untermoduln V_i, $i = 1,\dots,r$, von V zerlegen. Da jedes V_i die gleiche $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade enthält, folgt $V = V_i$ für ein i und somit die Irreduzibilität. Wir zeigen nur die Rückrichtung. Für die Hinrichtung verweisen wir auf [5, Kap. III.1.4, Satz 1, a)].
$\boxed{\Rightarrow}$ Für die Eindeutigkeit der $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade verweisen wir auf Kraft [5, Kap. III.1.4, Satz 1, a)]. □

Wir können die Rückrichtung des Satzes 2.33 noch verallgemeinern, ohne dabei die Beweisschritte zu verändern.

Satz 2.34. *Sei V ein vollständig reduzibler L-Modul, wobei V ein endlichdimensionaler Vektorraum über $\mathbb{C}$ oder $\mathbb{R}$ und L eine Lie-Algebra ist. Weiter besitze jeder L-Untermodul U von V eine L-stabile Gerade. Dann gilt: Besitzt V genau eine L-stabile Gerade, so ist V ein irreduzibler L-Modul.*

Bemerkung 2.35. Satz 2.34 gilt analog für Gruppen-Moduln, wenn wir die Lie-Algebren durch Gruppen ersetzen, da sich die Beweise kaum unterscheiden.

3 Der Raum der homogenen Polynome

Sei K ein Körper und seien d und n natürliche Zahlen. Wir definieren $\mathcal{H}_d^n(K)$ als den Raum der homogenen Polynome vom Grad d in den Variablen $X_0,\dots,X_n$ mit Koeffizienten aus K. Der Raum der multivariaten homogenen Polynome $\mathcal{H}_d^n(K)$ ist ein K-Vektorraum der Dimension $\binom{d+n}{d}$. Eine Basis für $\mathcal{H}_d^n(K)$ ist zum Beispiel die Monombasis $\{X^\alpha \mid |\alpha| = d\}$. Dabei bezeichnet $\alpha := (\alpha_0,\dots,\alpha_n) \in \mathbb{N}^{n+1}$, $|\alpha| := \alpha_0 + \cdots + \alpha_n$ und $X^\alpha := X_0^{\alpha_0}\cdots X_n^{\alpha_n}$. Anders gesagt bedeutet dies: Jedes f aus $\mathcal{H}_d^n(K)$ kann in der Form $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$, wobei f_α aus K ist, geschrieben werden.

Wir untersuchen speziell den komplexen bzw. reellen Vektorraum $\mathcal{H}_d^n(\mathbb{C})$ bzw. $\mathcal{H}_d^n(\mathbb{R})$ der multivariaten homogenen Polynome. Wir möchten unitär bzw. orthogonal invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{C})$ bzw. $\mathcal{H}_d^n(\mathbb{R})$ klassifizieren. Dabei meinen wir mit unitärer Invarianz die Invarianz bezüglich der

unitären Gruppe $\mathcal{U}(n+1) := \{u \in GL(n+1, \mathbb{C}) \mid uu^* = I_{n+1}\}$, wobei u^* die ***adjungierte Matrix*** von u bezeichnet, d.h. es gilt $u^* = u^H := \overline{u}^T$. Analog meinen wir mit orthogonaler Invarianz die Invarianz bezüglich der ***orthogonalen Gruppe*** $\mathcal{O}(n+1) := \left\{u \in GL(n+1, \mathbb{R}) \mid uu^T = I_{n+1}\right\}$.

Die unitäre Gruppe $\mathcal{U}(n+1)$ bzw. ihr reelles Gegenstück, die orthogonale Gruppe $\mathcal{O}(n+1)$, sind die Mengen der unitären bzw. orthogonalen Transformationen zwischen $n+1$-dimensionalen komplexen bzw. reellen Vektorräumen, die das Standardskalarprodukt erhalten. Diese Transformationen entsprechen Rotationen und Spiegelungen und sind aus diesem Grund, von einem geometrischen Standpunkt ausgehend, von speziellem Interesse.

Die unitäre Gruppe definiert auf natürliche Weise eine (Links-)Gruppenaktion auf dem Vektorraum $\mathcal{H}_d^n(\mathbb{C})$ durch

$$uf := u.f := f \circ u^{-1}, \tag{4}$$

für u aus $\mathcal{U}(n+1)$ und f aus $\mathcal{H}_d^n(\mathbb{C})$. Der Raum $\mathcal{H}_d^n(\mathbb{C})$ wird damit zu einem $\mathcal{U}(n+1)$-Modul.

Die Frage nach der Zerlegung dieses Moduls in irreduzible Untermoduln liefert uns eine Antwort auf die Gestalt unitär invarianter Skalarprodukte. In Abschnitt 3.1.1 werden wir zeigen, dass $\mathcal{H}_d^n(\mathbb{C})$ ein irreduzibler $\mathcal{U}(n+1)$-Modul ist. Das Korollar vom Lemma von Schur (Korollar 2.6) ermöglicht es dann, die Eindeutigkeit unitär invarianter Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ (bis auf Multiplikation mit einem positiven Skalar) zu zeigen (vgl. Satz 3.8).

Analog definiert (4) eine (Links-)Gruppenaktion auf dem reellen Vektorraum $\mathcal{H}_d^n(\mathbb{R})$. Anders als $\mathcal{H}_d^n(\mathbb{C})$ wird der $\mathcal{O}(n+1)$-Modul $\mathcal{H}_d^n(\mathbb{R})$ in mehrere irreduzible $\mathcal{O}(n+1)$-Untermoduln zerfallen. Auf jedem dieser irreduziblen Moduln werden wir ein (bis auf Multiplikation mit einem positiven Skalar) eindeutiges Skalarprodukt finden. Daher werden wir mehr Möglichkeiten haben, orthogonal invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ zu beschreiben.

Bemerkung 3.1. Da jedes nicht-homogene Polynom durch Homogenisierung in ein homogenes Polynom überführt werden kann und sich umgekehrt auch jedes homogene Polynom durch Dehomogenisierung in ein nicht-homogenes Polynom überführen lässt, beschränken wir uns im Folgenden auf den homogenen Fall.

3.1 Unitär invariantes Skalarprodukt

In diesem Abschnitt untersuchen wir den $\mathbb{C}$-Vektorraum $\mathcal{H}_d^n(\mathbb{C})$ der homogenen Polynome vom Grad d in $n+1$ Variablen. Dabei wählen wir die spezielle Basis

$$\left\{\sqrt{\tbinom{d}{\alpha}}X^\alpha \mid |\alpha| = d\right\}$$

mit Multinomialkoeffizient

$$\binom{d}{\alpha} := \frac{d!}{\alpha_0! \cdots \alpha_n!},$$

welche eine Umskalierung der Monombasis ist. Diese Basis werden wir im Weiteren ***Weylsche Basis*** nennen. Wir beobachten, dass das Skalarprodukt der Koordinatenvektoren in dieser Basis ein Hermitesches Skalarprodukt $\langle \cdot,\cdot \rangle$ auf $\mathcal{H}_d^n(\mathbb{C})$ definiert. Wir definieren genauer gesagt:

$$\langle f,g \rangle := \sum_{|\alpha|=d} f_\alpha \overline{g_\alpha} \tag{5}$$

für $f = \sum_{|\alpha|=d} \sqrt{\binom{d}{\alpha}} f_\alpha X^\alpha$ und $g = \sum_{|\alpha|=d} \sqrt{\binom{d}{\alpha}} g_\alpha X^\alpha$ aus $\mathcal{H}_d^n(\mathbb{C})$. Der Benennung folgend, werden wir dieses Skalarprodukt ***Weylsches Skalarprodukt*** nennen.

Wir möchten im Folgenden zeigen, dass das Weylsche Skalarprodukt invariant unter der Gruppe $\mathcal{U}(n+1)$ der unitären Matrizen ist.

Satz 3.2. *Für alle $f, g \in \mathcal{H}_d^n(\mathbb{C})$ und alle $u \in \mathcal{U}(n+1)$ gilt*

$$\langle uf,ug \rangle = \langle f,g \rangle.$$

Den Beweis haben wir aus [7, Thm. 16.3] entnommen.

Beweis von Satz 3.2. Wir betrachten die Funktion

$$K : \mathbb{C}^{n+1} \times \mathbb{C}^{n+1} \to \mathbb{C}, \quad (x,y) \mapsto \langle x,y \rangle^d,$$

wobei $\langle x,y \rangle := \sum_{i=0}^n x_i \overline{y}_i = y^H x$ das Hermitesche Standardskalarprodukt auf $\mathbb{C}^{n+1}$ bezeichnet. Dann gilt insbesondere $K(ux,uy) = K(x,y)$ für alle $u \in \mathcal{U}(n+1)$ und alle $x, y \in \mathbb{C}^{n+1}$, da unitäre Matrizen das Standardskalarprodukt erhalten. Wir fixieren $y \in \mathbb{C}^{n+1}$ und fassen K_y als Funktion von $\mathbb{C}^{n+1}$ nach $\mathbb{C}$ auf, mit

$$K_y(x) := K(x,y) = (x_0\overline{y}_0 + \cdots + x_n\overline{y}_n)^d = \sum_{\substack{\alpha \in \mathbb{N}^{n+1} \\ |\alpha|=d}} \binom{d}{\alpha} \overline{y}^\alpha x^\alpha$$

in Multiindexnotation. Die letzte Gleichung ist das ***Multinomialtheorem***. Dann haben wir $K_y \in \mathcal{H}_d^n(\mathbb{C})$. Die Darstellung von K_y in der Weylschen Basis ist wie folgt gegeben:

$$K_y = \sum_{|\alpha|=d} \sqrt{\binom{d}{\alpha}} \cdot \sqrt{\binom{d}{\alpha}}\, \overline{y}^\alpha \cdot X^\alpha$$

Mit anderen Worten bedeutet dies: $\sqrt{\binom{d}{\alpha}}\, \overline{y}^\alpha$, $\alpha \in \mathbb{N}^{n+1}$ mit $|\alpha| = d$, sind die Koordinaten von K_y in der Weylschen Basis.

Sei $f = \sum_{|\alpha|=d} \sqrt{\binom{d}{\alpha}}\, f_\alpha X^\alpha$ aus $\mathcal{H}_d^n(\mathbb{C})$ auch in der Weylschen Basis gegeben. Dann gilt:

$$\langle f,K_y \rangle = \sum_{|\alpha|=d} f_\alpha \sqrt{\binom{d}{\alpha}}\, y^\alpha = f(y). \tag{6}$$

Sind $u \in \mathcal{U}(n+1)$ und $y \in \mathbb{C}^{n+1}$, so beobachten wir folgendes Transformationsverhalten:

$$(uK_y)(X) \underset{\text{Def.}}{=} K_y(u^{-1}X) = \langle u^{-1}X,y\rangle^d \underset{\substack{\uparrow \\ u^{-1}=u^*}}{=} \langle u^*X,y\rangle^d = \langle X,uy\rangle^d = K_{uy}(X).$$

Somit gilt $uK_y = K_{uy}$.

Sei weiter $U := \operatorname{span}\left\{K_y \mid y \in \mathbb{C}^{n+1}\right\}$. Dann ist U nach Konstruktion ein Untervektorraum von $\mathcal{H}_d^n(\mathbb{C})$ und für das Vektorraumkomplement $U^\perp := \{f \in \mathcal{H}_d^n(\mathbb{C}) \mid \langle f,h\rangle = 0 \text{ für alle } h \in U\}$ in $\mathcal{H}_d^n(\mathbb{C})$ gilt die folgende Behauptung.

Behauptung: $U^\perp = \{0\}$.
Beweis: Angenommen $f \in \mathcal{H}_d^n(\mathbb{C})$ erfüllt $\langle f,K_y\rangle = f(y) = 0$ für alle y aus $\mathbb{C}^{n+1}$. Dann folgt $f = 0$ und wegen $U = \operatorname{span}\left\{K_y \mid y \in \mathbb{C}^{n+1}\right\}$ folgt auch $U^\perp = \{0\}$. ■

Wir erhalten also $\mathcal{H}_d^n(\mathbb{C}) = U \oplus U^\perp = U$, d.h. die Menge $\left\{K_y \mid y \in \mathbb{C}^{n+1}\right\}$ erzeugt den Vektorraum $\mathcal{H}_d^n(\mathbb{C})$. Daher genügt es, die Invarianz des Weylschen Skalarprodukts unter der unitären Gruppe $\mathcal{U}(n+1)$ für die Elemente des Erzeugers $\left\{K_y \mid y \in \mathbb{C}^{n+1}\right\}$ zu zeigen.

Seien also x, y aus $\mathbb{C}^{n+1}$ beliebig, dann haben wir

$$\begin{aligned}\langle uK_x,uK_y\rangle &= \langle K_{ux},K_{uy}\rangle \\ &\underset{(6)}{=} K_{ux}(uy) = \langle ux,uy\rangle^d = \langle x,y\rangle^d = K_y(x) \\ &\underset{(6)}{=} \langle K_x,K_y\rangle\end{aligned}$$

für alle u aus $\mathcal{U}(n+1)$ und die Behauptung des Satzes folgt. □

Wir haben also ein unitär invariantes Skalarprodukt gefunden. Interessanterweise wird sich zeigen, dass das Weylsche Skalarprodukt (bis auf Skalierung) das einzige unitär invariante Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ ist. Dazu benötigen wir, dass $\mathcal{H}_d^n(\mathbb{C})$ ein irreduzibler $\mathcal{U}(n+1)$-Modul ist. D.h $\mathcal{H}_d^n(\mathbb{C})$ enthält keinen nicht-trivialen linearen Unterraum, der $\mathcal{U}(n+1)$-invariant ist.

3.1.1 Irreduzibler $\mathcal{U}(n+1)$-Modul

Wir wissen, dass die Gruppenaktion (4) der unitären Gruppe $\mathcal{U}(n+1)$ auf dem Vektorraum der homogenen Polynome $\mathcal{H}_d^n(\mathbb{C})$ einen $\mathcal{U}(n+1)$-Modul definiert. Mit dem Weylschen Skalarprodukt wird $\mathcal{H}_d^n(\mathbb{C})$ sogar zu einem endlichdimensionalen unitären $\mathcal{U}(n+1)$-Modul. Insbesondere ist dieser nach Korollar 2.11 vollständig reduzibel.

In diesem Abschnitt werden wir zeigen, dass $\mathcal{H}_d^n(\mathbb{C})$ bereits ein irreduzibler $\mathcal{U}(n+1)$-Modul ist. Um die Irreduzibilität zu zeigen, wollen wir den Satz 2.33 verwenden. Allerdings ist Satz 2.33 in der Sprache der Lie-Algebren formuliert. Der $\mathcal{U}(n+1)$-Modul $\mathcal{H}_d^n(\mathbb{C})$ induziert einen

Lie-Algebren-Modul. Wir werden daher zunächst zeigen, dass $\mathcal{H}_d^n(\mathbb{C})$ als $\mathfrak{gl}(n, \mathbb{C})$-Modul irreduzibel ist, d.h. als Lie-Algebren-Modul.

Einen Rückschluss auf die Irreduzibilität von $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul liefert Korollar 2.32, zusammen mit der Eigenschaft, die wir als nächstes beweisen wollen: $\mathcal{H}_d^n(\mathbb{C})$ ist genau dann als $GL(n+1, \mathbb{C})$-Modul irreduzibel, wenn $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul irreduzibel ist.

Lemma 3.3. *Der $\mathcal{U}(n+1)$-Modul $\mathcal{H}_d^n(\mathbb{C})$ ist genau dann irreduzibel, wenn $\mathcal{H}_d^n(\mathbb{C})$ als $GL(n+1, \mathbb{C})$-Modul irreduzibel ist.*

Beweis. $\boxed{\Rightarrow}$ Angenommen $\mathcal{H}_d^n(\mathbb{C})$ sei als $GL(n+1, \mathbb{C})$-Modul reduzibel. Dann gibt es einen echten $GL(n+1, \mathbb{C})$-Untermodul $U \neq \{0\}$ von $\mathcal{H}_d^n(\mathbb{C})$. Wegen $\mathcal{U}(n+1) \subseteq GL(n+1, \mathbb{C})$ ist U auch ein echter $\mathcal{U}(n+1)$-Untermodul. Die Umkehrung liefert das Gewünschte.
$\boxed{\Leftarrow}$ Wir zeigen die Umkehrung: Ist $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul reduzibel, so auch als $GL(n+1, \mathbb{C})$-Modul.

Sei dazu $U \neq \{0\}$ ein echter $\mathcal{U}(n+1)$-Untermodul von $\mathcal{H}_d^n(\mathbb{C})$. Dann ist die Menge $\{g \in GL(n+1, \mathbb{C}) \mid g.U = U\}$ eine Zariski-abgeschlossene Teilmenge von $GL(n+1, \mathbb{C})$, die $\mathcal{U}(n+1)$ enthält. Verwenden wir nun die Eigenschaft, dass die Gruppe $\mathcal{U}(n+1)$ Zariski-dicht in der Gruppe $GL(n+1, \mathbb{C})$ ist (vgl. [5, AII.5, Satz 4, Beweis]), so folgt

$$\big\{g \in GL(n+1, \mathbb{C}) \mid g.U = U\big\} \supseteq \overline{\mathcal{U}(n+1)} = GL(n+1, \mathbb{C}).$$

Dabei bezeichnet $\overline{\mathcal{U}(n+1)}$ den Zariski-Abschluss von $\mathcal{U}(n+1)$. Mit $\{g \in GL(n+1, \mathbb{C}) \mid g.U = U\} \subseteq GL(n+1, \mathbb{C})$ folgt die Gleichheit.

Damit ist U auch ein $GL(n+1, \mathbb{C})$-Untermodul von $\mathcal{H}_d^n(\mathbb{C})$. □

Bemerkung 3.4. Aus dem Beweis von Lemma 3.3 erhalten wir, dass $\mathcal{U}(n+1)$-Untermoduln von $\mathcal{H}_d^n(\mathbb{C})$ auch $GL(n+1, \mathbb{C})$-Untermoduln sind und umgekehrt. Für unsere Zwecke brauchen wir also nicht zwischen $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul und $\mathcal{H}_d^n(\mathbb{C})$ als $GL(n+1, \mathbb{C})$-Modul zu unterscheiden.

Wir betrachten im Folgenden $\mathcal{H}_d^n(\mathbb{C})$ als $GL(n+1, \mathbb{C})$-Modul mit der Darstellung $\rho : GL(n+1, \mathbb{C}) \to GL(\mathcal{H}_d^n(\mathbb{C}))$, $\rho(u)(f) = f \circ u^{-1}$ (vgl. (4)). Die Darstellung ρ induziert eine Darstellung von Lie-Algebren durch

$$\begin{aligned} &D_0\rho : \mathfrak{gl}(n+1, \mathbb{C}) \to \mathfrak{gl}(\mathcal{H}_d^n(\mathbb{C})), \\ &D_0\rho(A)(f)(x) = \frac{d}{dt} f(e^{-tA}x)|_{t=0}, \end{aligned} \tag{7}$$

wobei $A \in \mathfrak{gl}(n+1, \mathbb{C})$, $f \in \mathcal{H}_d^n(\mathbb{C})$ und $x \in \mathbb{C}^{n+1}$. Dazu haben wir Beispiel 2.29 sowie Bemerkung 2.28 (ii) verwendet und die Eigenschaft, dass das Matrixexponential die Lie-Algebra $\mathfrak{gl}(n+1, \mathbb{C})$ surjektiv auf die Gruppe $GL(n+1, \mathbb{C})$ abbildet (vgl. Bemerkung 2.28). Die Darstellung $D_0\rho$ nennen wir die von ρ ***induzierte Darstellung***.

Für $A \in \mathfrak{gl}(n+1, \mathbb{C})$, $f \in \mathcal{H}_d^n(\mathbb{C})$ und $x \in \mathbb{C}^{n+1}$ gilt genauer

$$D_0\rho(A)(f)(x) = \frac{d}{dt} f(\exp(-tA)x)|_{t=0} = \frac{d}{dt} f(x - tAx)|_{t=0}.$$

Für die Standardbasismatrizen $\mathbb{E}_{ij}$ von $\mathfrak{gl}(n+1,\mathbb{C})$ gilt speziell

$$\begin{aligned} D_0\rho(\mathbb{E}_{ij})(f)(x) &= \frac{d}{dt} f(x - t\mathbb{E}_{ij}x)|_{t=0} \\ &= \frac{d}{dt} f(x_0,\dots,x_i - tx_j,\dots,x_n)|_{t=0} = -x_j \frac{\partial}{\partial x_i} f(x). \end{aligned}$$

Für $A = \sum_{i,j} a_{ij}\mathbb{E}_{ij}$ erhalten wir dann:

$$D_0\rho(A)(f) = -\sum_{i,j} a_{ij} X_j \frac{\partial}{\partial X_i} f.$$

Wir möchten jetzt zeigen, dass die induzierte Darstellung tatsächlich eine Darstellung der Lie-Algebra $\mathfrak{gl}(n+1,\mathbb{C})$ auf $\mathcal{H}_d^n(\mathbb{C})$ ist.

Proposition 3.5. *Die Abbildung $D_0\rho$ in (7) ist eine Darstellung der Lie-Algebra $\mathfrak{gl}(n+1,\mathbb{C})$ auf $\mathcal{H}_d^n(\mathbb{C})$.*

Beweis. Um zu zeigen, dass die induzierte Darstellung eine Darstellung von $\mathfrak{gl}(n+1,\mathbb{C})$ auf $\mathcal{H}_d^n(\mathbb{C})$ ist, müssen wir zeigen, dass die Lie-Klammer erhalten bleibt. Es gilt:

$$\begin{aligned} &[D_0\rho(E_{ij})(f), D_0\rho(E_{kl})(f)] = \left[-X_j \frac{\partial}{\partial X_i} f, -X_l \frac{\partial}{\partial X_k} f\right] \\ &= \left(-X_j \frac{\partial}{\partial X_i} f\right)\left(-X_l \frac{\partial}{\partial X_k} f\right) - \left(-X_l \frac{\partial}{\partial X_k} f\right)\left(-X_j \frac{\partial}{\partial X_i} f\right) \\ &= X_j \frac{\partial X_l}{\partial X_i} \frac{\partial}{\partial X_k} f + X_j X_l \frac{\partial}{\partial X_i}\frac{\partial}{\partial X_k} f - \left(X_l \frac{\partial X_j}{\partial X_k}\frac{\partial}{\partial X_i} f + X_l X_j \frac{\partial}{\partial X_k}\frac{\partial}{\partial X_i} f\right) \\ &= \delta_{il} X_j \frac{\partial}{\partial X_k} f - \delta_{jk} X_l \frac{\partial}{\partial X_i} f \\ &= D_0\rho([E_{ij}, E_{kl}])(f). \end{aligned}$$

□

$\mathcal{H}_d^n(\mathbb{C})$ wird also mit der Darstellung $D_0\rho$ in (7) zu einem $\mathfrak{gl}(n+1,\mathbb{C})$-Modul. Außerdem ist $\mathcal{H}_d^n(\mathbb{C})$ als $GL(n+1,\mathbb{C})$-Modul vollständig reduzibel und somit auch als $\mathfrak{gl}(n+1,\mathbb{C})$-Modul (vgl. Lemma 2.30). Für die Irreduzibilität von $\mathcal{H}_d^n(\mathbb{C})$ als Lie-Algebren-Modul genügt es nach Satz 2.33 zu zeigen, dass $\mathcal{H}_d^n(\mathbb{C})$ genau eine $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade (eindimensionalen $\mathfrak{t}(n,\mathbb{C})$-Untermodul) besitzt.

Satz 3.6. *Der $\mathfrak{gl}(n+1,\mathbb{C})$-Modul $\mathcal{H}_d^n(\mathbb{C})$ besitzt genau eine $\mathfrak{t}(n,\mathbb{C})$-stabile Gerade.*

Beweis. Wir zeigen: $\mathbb{C}X_n^d$ ist der einzige lineare Unterraum von $\mathcal{H}_d^n(\mathbb{C})$, der unter $D_0\rho|_{\mathfrak{t}(n,\mathbb{C})}$ invariant bleibt.

Hierfür reicht es aus, zu untersuchen, wie $D_0\rho|_{\mathfrak{t}(n,\mathbb{C})}$ Monome aus $\mathcal{H}_d^n(\mathbb{C})$ abbildet. Sei X^α ein Monom aus $\mathcal{H}_d^n(\mathbb{C})$ mit $|\alpha| = d$ und $\mathbb{E}_{ij}$ wieder eine Standardbasismatrix. Dann gilt:

(i) $D_0\rho(\mathbb{E}_{ii})(X^\alpha) = -\alpha_i X^\alpha$.

(ii) $D_0\rho(\mathbb{E}_{ij})(X^\alpha)$

$$= -X_j \frac{\partial}{\partial X_i} X^\alpha$$
$$= -\alpha_i \, X_0^{\alpha_0} \cdots X_{i-1}^{\alpha_{i-1}} X_i^{\alpha_i - 1} X_{i+1}^{\alpha_{i+1}} \cdots X_{j-1}^{\alpha_{j-1}} X_j^{\alpha_j+1} X_{j+1}^{\alpha_{j+1}} \cdots X_n^{\alpha_n}$$
$$= -\alpha_i X^{(\alpha_0,\ldots,\alpha_{i-1},\alpha_i-1,\alpha_{i+1},\ldots,\alpha_{j-1},\alpha_j+1,\alpha_{j+1},\ldots,\alpha_n)}$$

für $i < j$. In Worten bedeutet dies: Der Exponent von $X_i^{\alpha_i}$ reduziert sich um eins, da wir nach X_i ableiten. Außerdem multiplizieren wir mit X_j, d.h. der Exponent von $X_j^{\alpha_j}$ erhöht sich um eins.

(iii) $D_0\rho(\mathbb{E}_{ij})(X_k^d) = -X_j \frac{\partial}{\partial X_i} X_k^d$, für $i < j$ und $0 \le k \le n$.

Behauptung 1: $D_0\rho(\mathbb{E}_{ij})(X_n^d) \in \mathbb{C}\, X_n^d$ für alle $i \le j$.
Beweis: Der Fall $i = j$ ist (i). Sei also $i < j$. Mit (iii) folgt dann:

$$D_0\rho(\mathbb{E}_{ij})(X_n^d) = -X_j \frac{\partial}{\partial X_i} X_n^d = -X_j \cdot 0 = 0 \in \mathbb{C}\, X_n^d.$$

■

Behauptung 2: $D_0\rho(\mathbb{E}_{ij})(X^\alpha) \notin \mathbb{C}\, X^\alpha$ für alle $i < j$ und $\alpha_k \neq d$, $k = 0, \ldots, n$.
Beweis: Da $\alpha_k \neq d$, gibt es $0 \le l < m \le n$ mit $1 \le \alpha_l, \alpha_m < d$. Mit (ii) folgt dann:

$$D_0\rho(\mathbb{E}_{lm})(X^\alpha) = -\alpha_l X^{(\alpha_0,\ldots,\alpha_{l-1},\alpha_l-1,\alpha_{l+1},\ldots,\alpha_{m-1},\alpha_m+1,\alpha_{m+1},\ldots,\alpha_n)} \notin \mathbb{C}\, X^\alpha.$$

■

Wir können also die Monome, die von $D_0\rho|_{\mathfrak{t}(n,\mathbb{C})}$ auf ein skalares Vielfaches von sich selbst abbgebildet werden, auf X_k^d für $k = 0, \ldots, n$ einschränken.

Behauptung 3: $D_0\rho(\mathbb{E}_{ij})(X_k^d) \notin \mathbb{C}\, X_k^d$ für alle $i < j$ und $0 \le k < n$.
Beweis: Wegen $0 \le k < n$ gibt es ein j mit $0 \le k < j \le n$ und mit (iii) folgt:

$$D_0\rho(\mathbb{E}_{kj})(X_k^d) = -X_j \frac{\partial}{\partial X_k} X_k^d = -d X_j X_k^{d-1} \notin \mathbb{C}\, X_k^d.$$

■

Die Behauptungen 1 bis 3 zeigen, dass X_n^d das einzige Monom ist, welches von $D_0\rho|_{\mathfrak{t}(n,\mathbb{C})}$ auf ein skalares Vielfaches von sich selbst abgebildet wird. Da $D_0\rho|_{\mathfrak{t}(n,\mathbb{C})}$ insbesondere $\mathbb{C}$-linear ist, ist $\mathbb{C}\, X_k^d$ der einzige eindimensionale $\mathfrak{t}(n, \mathbb{C})$-Untermodul von $\mathcal{H}_d^n(\mathbb{C})$. □

Wir haben also gezeigt, dass der $\mathfrak{gl}(n+1, \mathbb{C})$-Modul $\mathcal{H}_d^n(\mathbb{C})$ irreduzibel ist. Als direkte Folgerung aus Korollar 2.32 und Lemma 3.3 erhalten wir, dass $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul ebenfalls irreduzibel ist. Diese Aussage halten wir im folgenden Satz fest.

Satz 3.7. *$\mathcal{H}_d^n(\mathbb{C})$ ist ein irreduzibler $\mathcal{U}(n+1)$-Modul.*

3.1.2 Eindeutiges Skalarprodukt

Wir wissen, dass das Weylsche Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ invariant unter der Gruppe $\mathcal{U}(n+1)$ der unitären Matrizen ist (vgl. Satz 3.2). Im Folgenden möchten wir zeigen, dass das Weylsche Skalarprodukt (bis auf Multiplikation mit einem positiven Skalar) das einzige unitär invariante Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ ist. Die Hauptidee des Beweises der Eindeutigkeit besteht dabei aus der Anwendung des Lemmas von Schur (Lemma 2.5). Dazu benötigen wir die Irreduzibilität von $\mathcal{H}_d^n(\mathbb{C})$ als $\mathcal{U}(n+1)$-Modul, die wir in Abschnitt 3.1.1 erarbeitet haben.

Satz 3.8. *Das Weylsche Skalarprodukt ist, bis auf Multiplikation mit einem positiven Skalar, das einzige unitär invariante Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$.*

Beweis. Sei $\langle\cdot,\cdot\rangle$ das Weylsche Skalarprodukt und $\langle\cdot,\cdot\rangle'$ sei ein weiteres $\mathcal{U}(n+1)$-invariantes Hermitesches Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ mit Semilinearität im zweiten Argument. Wir betrachten die linearen Abbildungen

$$\mathrm{T} : \overline{\mathcal{H}_d^n(\mathbb{C})} \to \mathcal{H}_d^n(\mathbb{C})^*, \quad f \mapsto \langle\cdot,f\rangle$$

und

$$\mathrm{T}' : \overline{\mathcal{H}_d^n(\mathbb{C})} \to \mathcal{H}_d^n(\mathbb{C})^*, \quad f \mapsto \langle\cdot,f\rangle',$$

wobei $\mathcal{H}_d^n(\mathbb{C})^*$ den Dualraum und $\overline{\mathcal{H}_d^n(\mathbb{C})}$ den komplex konjugierten Vektorraum von $\mathcal{H}_d^n(\mathbb{C})$ bezeichnet. Die Skalarmultiplikation auf $\overline{\mathcal{H}_d^n(\mathbb{C})}$ ist durch $\lambda \cdot v = \overline{\lambda}\, v$ gegeben.

Die (Links-)Gruppenaktion (4) induziert eine (Links-)Gruppenaktion auf $\mathcal{H}_d^n(\mathbb{C})^*$ durch $u.\,\mathrm{T}(f)(g) = \langle u^{-1}g,f\rangle$, die wir die zu (4) ***duale Gruppenaktion*** nennen.

Behauptung 1: Die zu (4) duale Gruppenaktion ist eine (Links-)Gruppenaktion der Gruppe $\mathcal{U}(n+1)$ auf $\mathcal{H}_d^n(\mathbb{C})^*$.
Beweis: Seien $u, v \in \mathcal{U}(n+1)$, $f \in \overline{\mathcal{H}_d^n(\mathbb{C})}$ und $g \in \mathcal{H}_d^n(\mathbb{C})$. Dann gilt:

$$\begin{aligned}(uv).\,\mathrm{T}(f)(g) &= \langle v^{-1}u^{-1}g,f\rangle \\ &= \mathrm{T}(f)(v^{-1}u^{-1}g) = v.\,\mathrm{T}(f)(u^{-1}g) \\ &= u.(v.\,\mathrm{T}(f))(g)\end{aligned}$$

und

$$\mathrm{id}.\,\mathrm{T}(f)(g) = \mathrm{T}(f)(g).$$

■

Behauptung 2: $\mathrm{T}, \mathrm{T}' \in \mathrm{Hom}_{\mathcal{U}(n+1)}\left(\overline{\mathcal{H}_d^n(\mathbb{C})}, \mathcal{H}_d^n(\mathbb{C})^*\right)$.
Beweis: Wir zeigen die Behauptung für T. Für T′ erfolgt der Beweis analog. Seien $u \in \mathcal{U}(n+1)$, $f \in \overline{\mathcal{H}_d^n(\mathbb{C})}$ und $g \in \mathcal{H}_d^n(\mathbb{C})$. Dann gilt:

$$u.\,\mathrm{T}(f)(g) = \langle u^{-1}g,f\rangle = \langle u^*g,f\rangle = \langle g,uf\rangle = \mathrm{T}(uf)(g).$$

Der Operator T ist also $\mathcal{U}(n+1)$-linear. T ist insbesondere $\mathbb{C}$-linear und damit ein Element der Menge $\mathrm{Hom}_{\mathcal{U}(n+1)}\left(\overline{\mathcal{H}_d^n(\mathbb{C})}, \mathcal{H}_d^n(\mathbb{C})^*\right)$. ■

Wir bemerken, dass die Eigenschaft der Irreduzibilität bzgl. der dualen Gruppenaktion erhalten bleibt. Daher gilt $\overline{\mathcal{H}_d^n(\mathbb{C})} \cong \mathcal{H}_d^n(\mathbb{C})^*$. Nach dem Korollar vom Lemma von Schur (Korollar 2.6) gibt es ein $\lambda \in \mathbb{C}$, sodass

$$\mathrm{T}' = \lambda\, \mathrm{T}\,. \tag{8}$$

Weiter erhalten wir aus

$$\mathrm{T}'(f)(f) = \lambda\, \mathrm{T}(f)(f)$$

für alle $f \in \mathcal{H}_d^n(\mathbb{C})$ und alle $f \in \overline{\mathcal{H}_d^n(\mathbb{C})}$ und der positiven Definitheit der Skalarprodukte, dass λ positiv sein muss. □

Interessanterweise wird sich zeigen, dass der Satz 3.8 über $\mathcal{H}_d^n(\mathbb{R})$ nicht richtig ist und man mehr Möglichkeiten hat, $\mathcal{O}(n+1)$-invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ zu wählen.

3.2 Orthogonal invariante Skalarprodukte

In Abschnitt 3.1.2 haben wir gezeigt, dass das Weylsche Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ (bis auf Multiplikation mit einer positiven Zahl) das einzige unitär invariante Skalarprodukt ist. Dazu haben wir verwendet, dass $\mathcal{H}_d^n(\mathbb{C})$ ein irreduzibler $\mathcal{U}(n+1)$-Modul ist.

Als nächstes untersuchen wir orthogonal invariante Skalarprodukte auf dem reellen Vektorraum $\mathcal{H}_d^n(\mathbb{R})$ der homogenen Polynome vom Grad d in $n+1$ Variablen. Analog zu $\mathcal{H}_d^n(\mathbb{C})$ wird $\mathcal{H}_d^n(\mathbb{R})$ in natürlicher Weise durch $f \circ g^{-1}$, für f aus $\mathcal{H}_d^n(\mathbb{R})$ und g aus der orthogonalen Gruppe $\mathcal{O}(n+1)$, zu einem orthogonalen $\mathcal{O}(n+1)$-Modul. In Abschnitt 3.2.4 zeigen wir, dass anders als im komplexen Fall, $\mathcal{H}_d^n(\mathbb{R})$ in mehrere irreduzible $\mathcal{O}(n+1)$-Untermoduln zerfällt. Auf jedem dieser irreduziblen Untermoduln werden wir ein (bis auf Multiplikation mit einer positiven Zahl) eindeutiges orthogonal invariantes Skalarprodukt finden. Dazu nutzen wir ähnlich wie im Beweis von Satz 3.8 das Lemma von Schur (Lemma 2.5). Genauer gesagt verwenden wir Bemerkung 2.7. Die Existenz eines reellen Eigenwertes, die wir für Bemerkung 2.7 brauchen, erhalten wir aus der Selbstadjungiertheit eines G-Modul-Endomorphismus (vgl. Behauptung 3.30).

Wir beschreiben kurz, wie dann orthogonal invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ aussehen und verweisen für eine genaue Beschreibung der Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in irreduzible Untermoduln auf später. Sei also

$$\mathcal{H}_d^n(\mathbb{R}) = \bigoplus_{i=1}^{m} W_i$$

eine Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in irreduzible $\mathcal{O}(n+1)$-Untermoduln W_i. Wir wählen auf jedem der W_i ein $\mathcal{O}(n+1)$-invariantes Skalarprodukt $\langle\cdot,\cdot\rangle_i$, welches eindeutig (bis auf Multiplikation mit einem positiven Skalar) ist. Dann können wir durch

$$\mathcal{H}_d^n(\mathbb{R}) \times \mathcal{H}_d^n(\mathbb{R}) \to \mathbb{R},\ (\oplus_{i=1}^m u_i, \oplus_{i=1}^m v_i) \mapsto \sum_{i=1}^m c_i \langle u_i, v_i \rangle_i \tag{9}$$

ein $\mathcal{O}(n+1)$-invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})$ definieren, welches durch positive Zahlen $c_1, \ldots, c_m$ parametrisiert ist.

Wir suchen also eine Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in irreduzible $\mathcal{O}(n+1)$-Untermoduln. Dazu folgen wir Atkinson & Han [8].

Für $d \geq 2$ und $r^2 := (X_0^2 + \cdots + X_n^2)$ ist zunächst

$$r^2 \mathcal{H}_{d-2}^n(\mathbb{R}) = \left(X_0^2 + \cdots + X_n^2\right) \mathcal{H}_{d-2}^n(\mathbb{R}) \tag{10}$$

ein echter $\mathcal{O}(n+1)$-invarianter Unterraum von $\mathcal{H}_d^n(\mathbb{R})$.

Bemerkung 3.9. Die Konstruktion in (10) ist für $\mathcal{H}_d^n(\mathbb{C})$ nicht möglich, da der Unterraum $r^2\mathcal{H}_{d-2}^n(\mathbb{C})$ nicht $\mathcal{U}(n+1)$-invariant ist und deshalb auch kein $\mathcal{U}(n+1)$-Untermodul von $\mathcal{H}_d^n(\mathbb{C})$ sein kann. Dies ist konsistent mit unserem Ergebnis aus Abschnitt 3.1.1.

Wir möchten als nächstes die $\mathcal{O}(n+1)$-invarianten Unterräume von $\mathcal{H}_d^n(\mathbb{R})$ identifizieren, die nicht den Faktor r^2 enthalten. Dies führt uns zu den sogenannten ***harmonischen Polynomen***.

3.2.1 Harmonische Polynome

Sei ∂_i die partielle Ableitung nach der i-ten Variable. Dann definieren wir den ***Laplace-Operator***

$$\Delta_n := \partial_0^2 + \cdots + \partial_n^2$$

im $(n+1)$-dimensionalen euklidischen Raum. Wir schreiben einfach nur Δ, falls die Dimension klar ist. Darüber hinaus ist es nützlich, den ***Nabla-Operator*** $\nabla_n := (\partial_0, \ldots, \partial_n)^T$ zu definieren. Wir schreiben wieder einfach nur ∇, falls die Dimension klar ist.

Der Nabla-Operator ist bekannt aus der Physik und kann sowohl auf differenzierbare Skalarfelder als auch auf differenzierbare Vektorfelder angewendet werden. Die Anwendung des Nabla-Operators auf ein differenzierbares Skalarfeld $f : D \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$ ergibt den ***Gradienten*** $\operatorname{grad}(f) := \nabla f = (\partial_0 f, \ldots, \partial_n f)^T$. Für ein differenzierbares Vektorfeld $h : D \subseteq \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ können wir die sogenannte ***Divergenz*** $\operatorname{div}(h) := \sum_{i=0}^n \partial_i h_i$ definieren. Die Divergenz können wir auch als Standardskalarprodukt zwischen ∇ und h interpretieren, d.h. als die Summe der komponentenweisen Produkte. Wir schreiben daher $\nabla \cdot h := \sum_{i=0}^n \partial_i h_i = \operatorname{div}(h)$.

Wir beobachten, dass der Laplace-Operator einem zweimal differenzierbaren Skalarfeld f die Divergenz seines Gradienten zuordnet, d.h. $\Delta f = \operatorname{div}(\operatorname{grad} f) = \nabla \cdot \nabla f$.

Ist f aus $\mathcal{H}_d^n(\mathbb{R})$, so nennen wir f ***harmonisch***, falls

$$\Delta f = \Delta_n f = \sum_{i=0}^{n} \partial_i^2 f = 0$$

gilt.

Definition 3.10. Der Raum $\mathcal{Y}_d^n(\mathbb{R})$ der reellen homogenen ***harmonischen Polynome*** vom Grad d in $n+1$ Variablen besteht aus den Polynomen von $\mathcal{H}_d^n(\mathbb{R})$, die harmonisch sind.

Beispiel 3.11. Es ist $\mathcal{Y}_d^n(\mathbb{R}) = \mathcal{H}_d^n(\mathbb{R})$, falls $d = 0$ oder $d = 1$ gilt. Ist $n = 0$, so ist $\mathcal{Y}_d^n(\mathbb{R}) = \emptyset$ für alle $d \geq 2$. Sei $n = 1$ und $d = 2$, dann besteht $\mathcal{Y}_2^1(\mathbb{R})$ aus Polynomen der Form $a(X_0^2 - X_1^2) + bX_0X_1$, wobei a, b aus $\mathbb{R}$ sind.

Wir zeigen jetzt, dass der Raum $\mathcal{Y}_d^n(\mathbb{R})$ der harmonischen Polynome einen $\mathcal{O}(n+1)$-Untermodul bildet. Dazu nutzen wir die Eigenschaft, dass wir den Laplace-Operator als die Divergenz des Gradienten schreiben können.

Proposition 3.12. *Der Raum $\mathcal{Y}_d^n(\mathbb{R})$ ist invariant unter der orthogonalen Gruppe $\mathcal{O}(n+1)$.*

Beweis. Zu zeigen: $\Delta(g.f) = 0$ für alle $f \in \mathcal{Y}_d^n(\mathbb{R})$ und alle $g \in \mathcal{O}(n+1)$, d.h. $\Delta f(g^{-1}x) = 0$ für alle $f \in \mathcal{Y}_d^n(\mathbb{R})$ und alle $g \in \mathcal{O}(n+1)$.

Wir bezeichnen für $z \in \mathbb{R}^{n+1}$, $\nabla^{[z]} = (\frac{\partial}{\partial z_0}, \dots, \frac{\partial}{\partial z_n})^T$ und analog $\Delta^{[z]}$. Sei $y = g^{-1}x$ für $g \in \mathcal{O}(n+1)$ und $x \in \mathbb{R}^{n+1}$. Für f aus $\mathcal{Y}_d^n(\mathbb{R})$ erhalten wir:

$$\nabla^{[x]} f(g^{-1}x) = g^{-1}\nabla^{[y]} f(y)\big|_{y=g^{-1}x}.$$

Daher folgt

$$\begin{aligned} \Delta^{[x]} f(g^{-1}x) &= \nabla^{[x]} \cdot \nabla^{[x]} f(g^{-1}x) \\ &= g^{-1}\nabla^{[y]} \cdot g^{-1}\nabla^{[y]} f(y) \underset{(*)}{=} \nabla^{[y]} \cdot \nabla^{[y]} f(y) \\ &= \Delta^{[y]} f(y). \end{aligned} \tag{11}$$

Dabei haben wir in $(*)$ die orthogonale Invarianz des Standardskalarprodukts ausgenutzt. Da außerdem $f \in \mathcal{Y}_d^n(\mathbb{R})$ harmonisch ist, folgt aus der Gleichung (11), dass $\Delta(g.f) = 0$ für alle $g \in \mathcal{O}(n+1)$ gilt.

$\square$

Der Raum $\mathcal{Y}_d^n(\mathbb{R})$ wird also mit Proposition 3.12 zu einem $\mathcal{O}(n+1)$-Untermodul von $\mathcal{H}_d^n(\mathbb{R})$.

Um zeigen zu können, dass nicht-triviale Polynome aus $\mathcal{Y}_d^n(\mathbb{R})$ nicht den Faktor r^2 enthalten, machen wir von einem speziellen Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})$ Gebrauch. Wir erinnern uns, dass wir jedes $f \in \mathcal{H}_d^n(\mathbb{R})$ in der Form $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$, f_α aus $\mathbb{R}$ und α aus $\mathbb{N}^{n+1}$, schreiben können. Für ein solches f definieren wir den Differentialoperator (vgl. [8, Cp. 2.1.1])

$$f(\nabla) := \sum_{|\alpha|=d} f_\alpha \nabla^\alpha, \quad \nabla^\alpha := \frac{\partial^{|\alpha|}}{\partial_0^{\alpha_0} \cdots \partial_n^{\alpha_n}}. \tag{12}$$

Dann können wir durch

$$\langle f,h \rangle_\nabla := f(\nabla)h \tag{13}$$

eine Bilinearform auf $\mathcal{H}_d^n(\mathbb{R})$ definieren. Für $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$ und $h = \sum_{|\alpha|=d} h_\alpha X^\alpha$ aus $\mathcal{H}_d^n(\mathbb{R})$ erhalten wir aus der Definition (13) folgende Relation:

$$\langle f,h \rangle_\nabla = \sum_{|\alpha|=d} \alpha!\, f_\alpha h_\alpha, \tag{14}$$

wobei $\alpha! := \alpha_0! \cdots \alpha_n!$. Daher definiert die Bilinearform $\langle \cdot,\cdot \rangle_\nabla$ sogar ein Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})$.

Bemerkung 3.13. Das Skalarprodukt $\langle \cdot,\cdot \rangle_\nabla$ auf $\mathcal{H}_d^n(\mathbb{R})$ ist $\mathcal{O}(n+1)$-invariant.

Beweis. Wir betrachten $\langle f,h \rangle_\nabla = \sum_{|\alpha|=d} \alpha!\, f_\alpha h_\alpha$ wie in (14). Dann können wir wegen

$$\alpha! = \frac{d!}{\binom{d}{\alpha}}$$

die Gleichung (14) auch folgendermaßen schreiben:

$$\langle f,h \rangle_\nabla = \sum_{|\alpha|=d} \alpha!\, f_\alpha h_\alpha = d! \sum_{|\alpha|=d} \frac{f_\alpha}{\sqrt{\binom{d}{\alpha}}} \frac{h_\alpha}{\sqrt{\binom{d}{\alpha}}}.$$

Wir beobachten, dass $f_\alpha \binom{d}{\alpha}^{-1/2}$ und $h_\alpha \binom{d}{\alpha}^{-1/2}$ die Koeffizienten von f und h bezüglich der Weylschen Basis sind. Damit steht auf der rechten Seite der obigen Gleichung das Weylsche Skalarprodukt (über $\mathbb{R}$) zwischen f und h, skaliert mit dem Faktor $d!$ (vgl. (5)). Dieses ist nach Satz 3.2 $\mathcal{O}(n+1)$-invariant. □

Die Hauptidee bei der Konstruktion von $\langle \cdot,\cdot \rangle_\nabla$ ist die folgende Eigenschaft: Die linearen Operatoren $r^2 : \mathcal{H}_{d-2}^n(\mathbb{R}) \to \mathcal{H}_d^n(\mathbb{R})$, $f \mapsto r^2 f$ und $\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_{d-2}^n(\mathbb{R})$, $f \mapsto \Delta f$, sind bezüglich $\langle \cdot,\cdot \rangle_\nabla$ zueinander adjungiert.

Eigenschaft 3.14. *Für alle $h \in \mathcal{H}_d^n(\mathbb{R})$ und $f \in \mathcal{H}_{d-2}^n(\mathbb{R})$ gilt*

$$\langle f,\Delta h \rangle_\nabla = \langle r^2 f,h \rangle_\nabla.$$

Beweis. Aus der Definition (12) erhalten wir

$$\Delta = \left(X_0^2 + \cdots + X_n^2\right)(\nabla) = r^2\,(\nabla).$$

Daher folgt

$$\begin{aligned}\langle f,\Delta h\rangle_\nabla &= f(\nabla)\Delta h = f(\nabla)r^2(\nabla)h\\ &= (fr^2)(\nabla)h = (r^2 f)(\nabla)h\\ &= \langle r^2 f,h\rangle_\nabla.\end{aligned}$$

□

Wir verwenden jetzt die Eigenschaft 3.14 um zu zeigen, dass nicht-triviale Polynome aus $\mathcal{Y}_d^n(\mathbb{R})$ nicht den Faktor r^2 enthalten.

Lemma 3.15. *Nicht-triviale Polynome aus $\mathcal{Y}_d^n(\mathbb{R})$ enthalten nicht den Faktor r^2.*

Beweis. Wir zeigen: Ist $f = r^2 h$ aus $\mathcal{Y}_d^n(\mathbb{R})$ mit $h \in \mathcal{H}_{d-2}^n(\mathbb{R})$, dann folgt $f = 0$. Sei also $f = r^2 h \in \mathcal{Y}_d^n(\mathbb{R})$. Mit der Eigenschaft 3.14 erhalten wir:

$$\langle f,f\rangle_\nabla = \langle r^2 h,f\rangle_\nabla = \langle h, \underbrace{\Delta f}_{=0}\rangle_\nabla = 0. \tag{15}$$

Insbesondere folgt $f = 0$. □

Zum Abschluss dieses Abschnitts möchten wir zeigen, dass bezüglich eines orthogonal invarianten Skalarprodukts auf $\mathcal{H}_d^n(\mathbb{R})$, die Räume $\mathcal{Y}_d^n(\mathbb{R})$ für verschiedene d orthogonal zueinander sind. Wir haben schon ein orthogonal invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})$ kennengelernt, nämlich $\langle\cdot,\cdot\rangle_\nabla$. Allgemeiner können wir ein auf $\mathcal{H}_d^n(\mathbb{R})$ orthogonal invariantes Skalarprodukt über $\langle f,h\rangle_{\mathbb{S}^n} := \int_{\mathbb{S}^n} f(\sigma)h(\sigma)\,\mathrm{d}\sigma$ definieren. Dabei bezeichne $\mathbb{S}^n$ die ***Einheitssphäre*** im $\mathbb{R}^{n+1}$ und $\mathrm{d}\sigma$ ein orthogonal invariantes Maß (z.B. das Haar-Maß).

Die Einschränkung von Polynomen aus $\mathcal{H}_d^n(\mathbb{R})$ auf die Einheitssphäre $\mathbb{S}^n$ bei Skalarprodukten der Form $\langle f,h\rangle_{\mathbb{S}^n} = \int_{\mathbb{S}^n} fh\,\mathrm{d}\sigma$ ist gerechtfertigt, da Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ genau dann das Nullpolynom sind, wenn sie bereits auf der Einheitssphäre $\mathbb{S}^n$ verschwinden.

Proposition 3.16. *Sei $\langle f,h\rangle_{\mathbb{S}^n} := \int_{\mathbb{S}^n} f(\sigma)h(\sigma)\,\mathrm{d}\sigma$ ein $\mathcal{O}(n+1)$-invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})$. Dann gilt: Bezüglich $\langle\cdot,\cdot\rangle_{\mathbb{S}^n}$ sind die Räume $\mathcal{Y}_{d_1}^n(\mathbb{R})$ und $\mathcal{Y}_{d_2}^n(\mathbb{R})$ für $d_1 \neq d_2$ orthogonal zueinander.*

Beweis. Sei $f \in \mathcal{Y}_{d_1}^n(\mathbb{R})$ und $h \in \mathcal{Y}_{d_2}^n(\mathbb{R})$. Aufgrund der Homogenität gilt $f(x) = \|x\|^{d_1} f(\frac{x}{\|x\|}) = r^{d_1} f(\xi)$ für $r := \|x\|$ und $\xi := \frac{x}{\|x\|}$. Für die partielle Ableitung von f nach r gilt:

$$r\,\frac{\partial}{\partial r} f(r\xi) = r\,\frac{\partial f}{\partial r} r^{d_1}\,f(\xi) = r\,d_1\,r^{d_1-1}\,f(\xi) = d_1\,r^{d_1}\,f(x) = d_1\,f(r\xi).$$

Wir schränken f auf die Einheitssphäre $\mathbb{S}^n$ ein und erhalten $\frac{\partial}{\partial r} f(\xi) = d_1\,f(\xi)$ für $\xi \in \mathbb{S}^n$. Für h erhalten wir analog $\frac{\partial h}{\partial r}(\xi) = d_2\,h(\xi)$, $\xi \in \mathbb{S}^n$. Da $\Delta f = 0$ und $\Delta h = 0$ gilt, folgt mit der ***Green'schen Formel***:

$$(d_1 - d_2)\int_{\mathbb{S}^n} fh\,\mathrm{d}\sigma = \int_{\mathbb{S}^n} (h\frac{\partial f}{\partial r} - f\frac{\partial h}{\partial r})\,\mathrm{d}\sigma = \int_{\mathbb{B}^{n+1}} (h\Delta f - f\Delta h)\,\mathrm{d}x = 0.$$

Nach Voraussetzung gilt $d_1 - d_2 \neq 0$. Daher folgt $\int_{\mathbb{S}^n} fh\,\mathrm{d}\sigma = 0$. Dabei bezeichnet $\mathbb{B}^{n+1}$ die Einheitskugel im $\mathbb{R}^{n+1}$. □

3.2.2 Die Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln

In diesem Abschnitt möchten wir eine Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln finden. Dazu verwenden wir die spezielle Eigenschaft des Skalarprodukts $\langle\cdot,\cdot\rangle_\nabla$ (vgl. Eigenschaft 3.14). Wir werden sehen, dass $\mathcal{H}_d^n(\mathbb{R})$ eine Zerlegung in harmonische Polynome besitzt. Genauer gesagt: $\mathcal{H}_d^n(\mathbb{R})$ zerfällt in $\lfloor d/2 \rfloor$-viele $\mathcal{O}(n+1)$-Untermoduln $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$, $k = 0, \ldots, \lfloor d/2 \rfloor$. Dabei bezeichnen wir mit $\lfloor x \rfloor$ den ganzzahligen abgerundeten Wert von $x \in \mathbb{R}$. Den Faktor r^{2k} brauchen wir, damit der Grad erhalten bleibt. Im nächsten Abschnitt wird sich herausstellen, dass die Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in harmonische Polynome bereits eine Zerlegung in irreduzible $\mathcal{O}(n+1)$-Untermoduln ist.

Die zentrale Aussage dieses Abschnitts ist die Folgende: Bezüglich $\langle\cdot,\cdot\rangle_\nabla$ hat $\mathcal{H}_d^n(\mathbb{R})$ eine orthogonale Zerlegung in $\mathcal{O}(n+1)$-Untermoduln $\mathcal{Y}_d^n(\mathbb{R})$ und $r^2\mathcal{H}_{d-2}^n(\mathbb{R})$, d.h. es gilt $\mathcal{H}_d^n(\mathbb{R}) = \mathcal{Y}_d^n(\mathbb{R}) \oplus r^2\mathcal{H}_{d-2}^n(\mathbb{R})$.

Lemma 3.17. *Die lineare Abbildung $\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_{d-2}^n(\mathbb{R})$, $f \mapsto \Delta f$, ist surjektiv und bezüglich $\langle\cdot,\cdot\rangle_\nabla$ sind harmonische Polynome $\mathcal{Y}_d^n(\mathbb{R})$ orthogonal zu Polynomen aus $r^2\mathcal{H}_{d-2}^n(\mathbb{R})$.*

Beweis. $\boxed{\text{Surjektivität}}$ Angenommen $\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_{d-2}^n(\mathbb{R})$, $f \mapsto \Delta f$, sei nicht surjektiv. Dann gibt es ein Polynom $h \in \mathcal{H}_{d-2}^n(\mathbb{R})$ verschieden von Null, sodass

$$\langle \Delta f, h\rangle_\nabla = 0$$

für alle $f \in \mathcal{H}_d^n(\mathbb{R})$ gilt. Wir wählen $f = r^2 h$. Mit (15) folgt:

$$\langle f,f\rangle_\nabla = \langle r^2 h, f\rangle_\nabla = \langle h, \Delta f\rangle_\nabla = 0$$

und daraus $f = 0$. Insbesondere folgt wegen $f = r^2 h$ auch $h = 0$, im Widerspruch zur Annahme.

$\boxed{\text{Orthogonalität}}$ Die Orthogonalität erhalten wir sofort aus der Gleichung (15). □

Sei $d \geq 2$. Aus Lemma 3.17 erhalten wir wegen $\dim \mathcal{H}_{d-2}^n(\mathbb{R}) = \dim r^2\mathcal{H}_{d-2}^n(\mathbb{R})$ und $\ker \Delta = \mathcal{Y}_d^n(\mathbb{R})$ eine bezüglich $\langle\cdot,\cdot\rangle_\nabla$ orthogonale Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in zwei $\mathcal{O}(n+1)$-Untermoduln

$$\mathcal{H}_d^n(\mathbb{R}) = \mathcal{Y}_d^n(\mathbb{R}) \oplus r^2\mathcal{H}_{d-2}^n(\mathbb{R}). \tag{16}$$

Bemerkung 3.18. Aus der Relation (16) erhalten wir eine Formel für die Dimension von $\mathcal{Y}_d^n(\mathbb{R})$. Es gilt:

$$\dim \mathcal{Y}_d^n(\mathbb{R}) = \dim \mathcal{H}_d^n(\mathbb{R}) - \dim \mathcal{H}_{d-2}^n(\mathbb{R}) = \binom{d+n}{d} - \binom{d-2+n}{d-2}.$$

Eine rekursive Anwendung von Lemma 3.17 liefert uns die gewünschte Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln.

Satz 3.19. *Sei $d \geq 2$. Dann gilt:*

$$\mathcal{H}_d^n(\mathbb{R}) = \bigoplus_{k=0}^{\lfloor d/2 \rfloor} r^{2k} \mathcal{Y}_{d-2k}^n(\mathbb{R}) \tag{17}$$

ist eine bezüglich $\langle \cdot, \cdot \rangle_\nabla$ orthogonale Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermodul.

Bemerkung 3.20. Wegen Proposition 3.16 ist die Zerlegung (17) sogar bezüglich jedem $\mathcal{O}(n+1)$-invarianten Skalarprodukt orthogonal.

Bemerkung 3.21. Schränken wir die Zerlegung in (17) auf die Einheitssphäre $\mathbb{S}^n$ ein, so ist

$$\mathcal{H}_d^n(\mathbb{R})\big|_{\mathbb{S}^n} = \bigoplus_{k=0}^{\lfloor d/2 \rfloor} \mathcal{Y}_{d-2k}^n(\mathbb{R})\big|_{\mathbb{S}^n}$$

eine Zerlegung von $\mathcal{H}_d^n(\mathbb{R})\big|_{\mathbb{S}^n}$ in sogenannte ***spherical Harmonics*** $\mathcal{Y}_{d-2k}^n(\mathbb{R})\big|_{\mathbb{S}^n}$.

Wir möchten die Räume $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ etwas genauer untersuchen. Sei dazu im Folgenden f aus $\mathcal{H}_s^n(\mathbb{R})$. Es gelten folgende Relationen:

(i) $\displaystyle\sum_{i=0}^{n} x_i \frac{\partial f}{\partial x_i}(x) = s\, f(x).$ (Euler-Relation)

(ii) $\displaystyle\frac{\partial r}{\partial x_i} = \frac{x_i}{r}$ für $r := \|x\|$ und $0 \leq i \leq n$.

Seien s und m natürliche Zahlen und $r := \|x\|$. Für die partiellen Ableitungen von $r^m f(x)$ nach x_i, $0 \leq i \leq n$, gilt:

$$\frac{\partial}{\partial x_i}\big(r^m f(x)\big) = m\, r^{m-2}\, x_i\, f(x) + r^m \frac{\partial f}{\partial x_i}(x)$$

und

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2}\big(r^m f(x)\big) &= \big(m\,(m-2)\, r^{m-4}\, x_i^2 + m\, r^{m-2}\big)\, f(x) \\ &\quad + 2\, m\, r^{m-2}\, x_i \frac{\partial f}{\partial x_i}(x) + r^m \frac{\partial^2 f}{\partial x_i^2}(x).\end{aligned}$$

Mit der Euler-Relation erhalten wir

$$\Delta\big(r^m f(x)\big) = m\,(n + 2s + m - 1)\, r^{m-2}\, f(x) + r^m \Delta f(x).$$

Ist f harmonisch, so gilt insbesondere

$$\Delta\big(r^m f(x)\big) = m\,(n + 2s + m - 1)\, r^{m-2}\, f(x),$$

bzw.

$$r^2 \Delta\big(r^m f(x)\big) = m\,(n + 2s + m - 1)\, r^m\, f(x). \tag{18}$$

Setzen wir $s := d - 2k$, $m := 2k$ und $r^2 := X_0^2 + \cdots + X_n^2$, so erhalten wir aus der Gleichung (18), dass der Raum $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ der Eigenraum des linearen Operators $r^2\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_d^n(\mathbb{R})$, $f \mapsto r^2\Delta f$, zum Eigenwert $2k\,(n + 2d - 2k - 1)$ ist, $0 \leq k \leq \lfloor d/2 \rfloor$.

Proposition 3.22. *Die Räume $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ sind die Eigenräume des linearen Operators $r^2\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_d^n(\mathbb{R})$ zu den Eigenwerten $2k\,(n+2d-2k-1)$, $0 \leq k \leq \lfloor d/2 \rfloor$.*

Bemerkung 3.23. Wir hätten auch Proposition 3.22 nutzen können, um eine orthogonale Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln, wie in Satz 3.19 zu erhalten. Es genügt dazu die Eigenschaft, dass der lineare Operator $r^2\Delta : \mathcal{H}_d^n(\mathbb{R}) \to \mathcal{H}_d^n(\mathbb{R})$ selbstadjungiert ist. Es gilt: Die Eigenräume eines selbstadjungierten Operators zu verschiedenen Eigenwerten sind orthogonal zueinander.

3.2.3 Die Legendre-Harmonischen

In diesem Abschnitt werden wir eine spezielle Klasse von harmonische Polynomen aus $\mathcal{Y}_d^n(\mathbb{R})$ betrachten, die sogenannten ***Legendre-Harmonischen***. Legendre-Harmonische erfüllen eine bestimmte Invarianz-Bedingung: Sie sind invariant unter dem Stabilisator des $(n+1)$-ten Einheitsvektors $e_{n+1} \in \mathbb{R}^{n+1}$ der orthogonalen Gruppe. Wir werden zeigen, dass Legendre-Harmonische bis auf eine Konstante eindeutig bestimmt sind. Diese Eigenschaft wird später beim Beweis der Irreduzibilität des $\mathcal{O}(n+1)$-Moduls $\mathcal{Y}_d^n(\mathbb{R})$ eine wichtige Rolle spielen.

Sei zunächst f aus $\mathcal{H}_d^n(\mathbb{R})$. Wir können f in der Form

$$f\,(x_0, \ldots, x_n) = \sum_{i=0}^{d} (x_n)^i f_{d-i}\,(x_1, \ldots, x_{n-1})\,, \quad f_{d-i} \in \mathcal{H}_{d-i}^{n-1}(\mathbb{R}), \tag{19}$$

schreiben. Wenden wir den Laplace-Operator auf f an, so erhalten wir

$$\begin{aligned}
\Delta f(x) &= \sum_{i=0}^{d} \left(\Delta (x_n)^i\right) f_{d-i}\,(x_0, \ldots, x_{n-1}) + (x_n)^i\, \Delta f_{d-i}\,(x_0, \ldots, x_{n-1}) \\
&= \sum_{i=2}^{d} i\,(i-1)\,(x_n)^{i-2}\, f_{d-i}\,(x_0, \ldots, x_{n-1}) \\
&\qquad + \sum_{i=0}^{d-2} (x_n)^i\, \Delta f_{d-i}\,(x_0, \ldots, x_{n-1})\,.
\end{aligned}$$

Wir machen eine Indexverschiebung: Wir setzen $j := i - 2$ für $2 \leq i \leq d$. Dann gilt:

$$\begin{aligned}\Delta f(x) &= \sum_{j=0}^{d-2}(j+2)\,(j+1)\,(x_n)^j\, f_{d-j-2}\,(x_0,\dots,x_{n-1}) \\ &\qquad + \sum_{i=0}^{d-2}(x_n)^i\,\Delta f_{d-i}\,(x_0,\dots,x_{n-1}) \\ &= \sum_{j=0}^{d-2}(x_n)^j\Big[(j+2)(j+1)f_{d-j-2}\,(x_0,\dots,x_{n-1}) \\ &\qquad + \Delta f_{d-j}\,(x_0,\dots,x_{n-1})\Big].\end{aligned}$$

Für $f \in \mathcal{Y}_d^n(\mathbb{R})$ gilt $\Delta f = 0$ und daher folgt

$$f_{d-j-2} = -\frac{1}{(j+2)(j+1)}\Delta f_{d-j}, \quad 0 \le j \le d-2. \tag{20}$$

Folglich ist ein harmonisches Polynom $f \in \mathcal{Y}_d^n(\mathbb{R})$ durch $f_d \in \mathcal{H}_d^{n-1}(\mathbb{R})$ und $f_{d-1} \in \mathcal{H}_{d-1}^{n-1}(\mathbb{R})$ in (19) eindeutig bestimmt.

Wir bezeichnen mit $\mathcal{O}_{e_{n+1}}(n+1)$ den Stabilisator des Einheitsvektors $e_{n+1} \in \mathbb{R}^{n+1}$ der Gruppe $\mathcal{O}(n+1)$. Mit anderen Worten bedeutet dies:

$$\mathcal{O}_{e_{n+1}}(n+1) := \{g \in \mathcal{O}(n+1) \mid g\, e_{n+1} = g\}\,.$$

Wir können jedes $g \in \mathcal{O}_{e_{n+1}}(n+1)$ durch

$$g = \begin{bmatrix} g_1 & 0 \\ 0 & 1 \end{bmatrix},\ g_1 \in \mathcal{O}(n) \tag{21}$$

darstellen und daher $\mathcal{O}_{e_{n+1}}(n+1)$ mit $\mathcal{O}(n)$ identifizieren.

Definition 3.24. Wir nennen $L \in \mathcal{H}_d^n(\mathbb{R})$ eine ***Legendre-Harmonische*** vom Grad d in $n+1$ Variablen, falls L die folgenden drei Eigenschaften erfüllt:

(i) L ist harmonisch, d.h. es gilt $\Delta L = 0$.

(ii) Es gilt $g.L = L$ für alle $g \in \mathcal{O}_{e_{n+1}}(n+1)$.

(iii) Es gilt $L\,(e_{n+1}) = 1$. (Normierung)

Wir möchten jetzt zeigen, dass die Legendre-Harmonischen durch die Eigenschaften (i) bis (iii) aus Definition 3.24 bereits eindeutig bestimmt sind. Sei $L \in \mathcal{H}_d^n(\mathbb{R})$ eine Legendre-Harmonische. Schreiben wir L in der Form (19), so gilt

$$L\,(x_0,\dots,x_n) = \sum_{i=0}^{d}(x_n)^i L_{d-i}\,(x_1,\dots,x_{n-1})\,, \quad L_{d-i} \in \mathcal{H}_{d-i}^{n-1}(\mathbb{R}).$$

Schreiben wir außerdem $g \in \mathcal{O}_{e_{n+1}}(n+1)$ in der Form (21), so impliziert Definition 3.24 (ii), dass

$$L_{d-i}\left(g_1^{-1}(x_0,\dots,x_{n-1})\right) = g_1.L_{d-i}(x_0,\dots,x_{n-1}) = L_{d-i}(x_0,\dots,x_{n-1})$$

für alle $g_1 \in \mathcal{O}(n)$ gilt. Daraus folgt, dass L_{d-i} nur von $r_{n-1} := \sqrt{x_0^2+\dots+x_{n-1}^2}$ abhängt. Da L_{d-i} insbesondere homogen vom Grad $d-i$ ist, ist dies nur möglich, falls $d-i$ eine gerade Zahl ist. Daher gilt:

$$L_{d-i}(x_0,\dots,x_{n-1}) = \begin{cases} c_k \cdot r_{n-1}^{2k} & \text{, falls } d-i = 2k \\ 0 & \text{, falls } d-i = 2k+1 \end{cases}$$

für $c_k \in \mathbb{R}$. Insgesamt erhalten wir

$$L(x) = \sum_{k=0}^{\lfloor \frac{d}{2} \rfloor} c_k\, r_{n-1}^{2k}\,(x_n)^{d-2k}. \tag{22}$$

Wir wollen die Koeffizienten c_k bestimmen. Wir wenden (20) auf die $L_{d-i}(x_0,\dots,x_{n-1}) = c_k \cdot r_{n-1}^{2k}$ für gerade Zahlen $d-i$ an und erhalten:

$$c_k = -\frac{(d-2k+2)(d-2k+1)}{2k(2k+n-2)}\, c_{k-1}, \quad 1 \le k \le \lfloor {}^{d}\!/\!{}_{2} \rfloor.$$

Mit der Normierung $c_0 = L(e_{n+1}) = 1$ sind die c_k und damit auch die Legendre-Harmonische $L \in \mathcal{H}_d^n(\mathbb{R})$ eindeutig bestimmt.

Bemerkung 3.25. Die Existenz von Legendre-Harmonischen ist leicht zu sehen: Wir brauchen einfach nur die obigen Schritte rückwärts zu laufen. Definieren wir ein Polynom wie in (22) und die c_k wie oben, so erfüllt dieses Polynom die Eigenschaften aus der Definition 3.24.

3.2.4 Irreduzibilität von $\mathcal{Y}_d^n(\mathbb{R})$

In Abschnitt 3.2.2 haben wir eine orthogonale Zerlegung von $\mathcal{H}_d^n(\mathbb{R})$ in $\mathcal{O}(n+1)$-Untermoduln $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ gefunden. In diesem Abschnitt werden wir zeigen, dass diese Zerlegung bereits eine Zerlegung in irreduzible Untermoduln gewesen ist. Hierzu nutzen wir das allgemeine Prinzip (Satz 2.34), welches wir schon für den Beweis der Irreduzibilität von $\mathcal{H}_d^n(\mathbb{C})$ verwendet haben (vgl. Satz 3.6). Genauer gesagt: Wir zeigen die Existenz und die Eindeutigkeit einer $\mathcal{O}(n)$-stabilen Gerade in $\mathcal{Y}_d^n(\mathbb{R})$. Dabei identifizieren wir $\mathcal{O}_{e_{n+1}}(n+1)$ mit $\mathcal{O}(n)$ (vgl. (21)).

Wir zeigen zunächst die Existenz einer $\mathcal{O}(n)$-stabilen Gerade.

Lemma 3.26 (Existenz)**.** *Jeder $\mathcal{O}(n+1)$-Untermodul $U \neq \{0\}$ von $\mathcal{H}_d^n(\mathbb{R})$ besitzt eine $\mathcal{O}(n)$-stabile Gerade.*

Beweis. Sei U ein $\mathcal{O}(n+1)$-Untermodul von $\mathcal{H}_d^n(\mathbb{R})$, der nicht der Nullraum ist. Wir wählen ein $\mathcal{O}(n+1)$-invariantes Skalarprodukt $\langle\cdot,\cdot\rangle_U$ auf U.

Betrachten wir das (stetige) lineare Funktional

$$\delta_{e_{n+1}} : U \to \mathbb{R},\ f \mapsto f(e_{n+1}),$$

so ist $\delta_{e_{n+1}}$ invariant unter der Gruppe $\mathcal{O}(n)$, denn es gilt

$$g.\delta_{e_{n+1}}(f) = g.f(e_{n+1}) = f\left(g^{-1}e_{n+1}\right) = f(e_{n+1})$$

für alle g aus $\mathcal{O}(n)$ und f aus U. Insbesondere gilt $g.\delta_{e_{n+1}}(f) = \delta_{e_{n+1}}(g.f)$. Der ***Darstellungssatz von Fréchet-Riesz*** liefert uns genau ein $\zeta \in U$, sodass

$$f(e_{n+1}) = \delta_{e_{n+1}}(f) = \langle f,\zeta\rangle_U$$

für alle f aus U gilt.

Behauptung: ζ ist $\mathcal{O}(n)$-invariant.
Beweis: Aus der $\mathcal{O}(n)$-Invarianz des linearen Funktionals $\delta_{e_{n+1}}$ und der $\mathcal{O}(n+1)$-Invarianz des Skalarprodukts $\langle\cdot,\cdot\rangle_U$ folgt, dass

$$\langle f,\zeta\rangle_U = g.\delta_{e_{n+1}}(f) = \langle g.f,\zeta\rangle_U = \langle g^{-1}g.f,g^{-1}.\zeta\rangle_U = \langle f,g^{-1}.\zeta\rangle_U$$

für alle f aus U und g aus $\mathcal{O}(n)$ gilt. Also gilt: $\zeta = g^{-1}.\zeta$ für alle $g \in \mathcal{O}(n)$. ■

Wir erhalten: $\mathbb{R}\,\zeta$ ist eine $\mathcal{O}(n)$-stabile Gerade in U.

□

Sei U ein $\mathcal{O}(n+1)$-Untermodul von $\mathcal{Y}_d^n(\mathbb{R})$, der nicht der Nullraum ist und $\mathbb{R}\,\zeta$ die $\mathcal{O}(n)$-stabile Gerade in U aus dem obigen Beweis. Wir zeigen im Folgenden, dass $\mathbb{R}\,\zeta$ durch die $\mathcal{O}(n)$-Invarianz und die Eigenschaft, dass ζ harmonisch ist, bereits eindeutig bestimmt ist.

Lemma 3.27 (Eindeutigkeit)**.** *Sei $U \neq \{0\}$ ein $\mathcal{O}(n+1)$-Untermodul von $\mathcal{Y}_d^n(\mathbb{R})$ und $\mathbb{R}\,\zeta$ eine $\mathcal{O}(n)$-stabile Gerade in U. Dann gilt: Die $\mathcal{O}(n)$-stabile Gerade $\mathbb{R}\,\zeta$ ist eindeutig bestimmt.*

Beweis. Wir zeigen: ζ ist bis auf Multiplikation mit einem Skalar eindeutig bestimmt. Wir können zunächst annehmen, dass $\zeta(e_{n+1}) = 1$ gilt, ansonsten normieren wir ζ. Aus der $\mathcal{O}(n)$-Invarianz von ζ und der Voraussetzung, dass ζ harmonisch ist, folgt: ζ ist eine Legendre-Harmonische. Nach Abschnitt 3.2.3 ist die Legendre-Harmonische ζ eindeutig bestimmt.

Die einzige Freiheit, die uns bleibt, liegt in der Wahl der Normierungskonstante $c_0 = \zeta(e_{n+1})$. Daher ist ζ bis auf Multiplikation mit einem Skalar eindeutig bestimmt. □

Mit Lemma 3.26 und 3.27 folgt, dass jeder $\mathcal{O}(n+1)$-Untermodul von $\mathcal{Y}_d^n(\mathbb{R})$ genau eine $\mathcal{O}(n)$-stabile Gerade enthält. Aus Satz 2.34 (zusammen mit Bemerkung 2.35) erhalten wir die Irreduzibilität von $\mathcal{Y}_d^n(\mathbb{R})$ als $\mathcal{O}(n+1)$-Modul. Dieses Ergebnis halten wir in dem folgenden Satz fest.

Satz 3.28. *Die $\mathcal{O}(n+1)$-Moduln $\mathcal{Y}^n_{d-2k}(\mathbb{R})$ sind irreduzibel für* $0 \leq k \leq \lfloor \frac{d}{2} \rfloor$.

Nach Satz 3.19 und 3.28 ist daher

$$\mathcal{H}^n_d(\mathbb{R}) = \bigoplus_{k=0}^{\lfloor d/2 \rfloor} r^{2k} \mathcal{Y}^n_{d-2k}(\mathbb{R})$$

eine Zerlegung von $\mathcal{H}^n_d(\mathbb{R})$ in irreduzible $\mathcal{O}(n+1)$-Untermoduln $r^{2k}\mathcal{Y}^n_{d-2k}(\mathbb{R})$, $1 \leq k \leq \lfloor d/2 \rfloor$.

Wir möchten jetzt zeigen, dass Skalarprodukte auf den irreduziblen $\mathcal{O}(n+1)$-Untermoduln $r^{2k}\mathcal{Y}^n_{d-2k}(\mathbb{R})$ (bis auf Multiplikation mit einem positiven Skalar) eindeutig bestimmt sind.

Für die Eindeutigkeit des Weylschen Skalarprodukts (bis auf Multiplikation mit einem positiven Skalar) auf $\mathcal{H}^n_d(\mathbb{C})$ haben wir das Lemma von Schur, speziell Korollar 2.6, verwendet. Das Lemma von Schur (Lemma 2.5) ist für irreduzible G-Moduln über den komplexen Zahlen formuliert, da man für den Beweis die Existenz eines Eigenwertes benötigt. Daher können wir das Lemma von Schur nicht ohne Weiteres auf die irreduziblen reellen $\mathcal{O}(n+1)$-Untermoduln $r^{2k}\mathcal{Y}^n_{d-2k}(\mathbb{R})$ anwenden: Wir benötigen die Existenz eines reellen Eigenwertes (vgl. Bemerkung 2.7).

Wir definieren den linearen Operator

$$\mathrm{T} : \mathcal{Y}^n_d(\mathbb{R}) \to \mathcal{Y}^n_d(\mathbb{R})^*, \quad f \mapsto \langle \cdot, f \rangle$$

wobei $\langle \cdot,\cdot \rangle$ ein $\mathcal{O}(n+1)$-invariantes Skalarprodukt auf $\mathcal{Y}^n_d(\mathbb{R})$ ist. Sei außerdem ein weiteres $\mathcal{O}(n+1)$-invariantes Skalarprodukt $\langle \cdot,\cdot \rangle'$ auf $\mathcal{Y}^n_d(\mathbb{R})$ gegeben. Ist $f \in \mathcal{Y}^n_d(\mathbb{R})$, so gilt: $T(f) : \mathcal{Y}^n_d(\mathbb{R}) \to \mathbb{R}$, $h \mapsto \langle h, f \rangle$, ist ein lineares Funktional. Nach dem Darstellungssatz von Fréchet-Riesz gibt es genau ein $\zeta_f \in \mathcal{Y}^n_d(\mathbb{R})$, sodass $T(f)(h) = \langle h, \zeta_f \rangle'$ für alle $h \in \mathcal{Y}^n_d(\mathbb{R})$ gilt. Wir können daher eine $\mathbb{R}$-lineare Abbildung $\varphi : \mathcal{Y}^n_d(\mathbb{R}) \to \mathcal{Y}^n_d(\mathbb{R})$ definieren, die durch

$$T(f)(h) = \langle h, \varphi(f) \rangle'$$

für alle $f, h \in \mathcal{Y}^n_d(\mathbb{R})$ gegeben ist. Wir zeigen jetzt, dass φ sogar $\mathcal{O}(n+1)$-linear ist, d.h. ein $\mathcal{O}(n+1)$-Modul-Homomorphismus.

Behauptung 3.29. *φ ist $\mathcal{O}(n+1)$-linear.*

Beweis. Es ist zu zeigen: $g.\varphi(f) = \varphi(g.f)$ gilt für alle $f \in \mathcal{Y}^n_d(\mathbb{R})$ und $g \in \mathcal{O}(n+1)$.

Für $f, h \in \mathcal{Y}^n_d(\mathbb{R})$ und $g \in \mathcal{O}(n+1)$ erhalten wir

$$\begin{aligned} \langle h, \varphi(g.f) \rangle' &= T(g.f)(h) = \langle h, g.f \rangle = g.\langle h, f \rangle \\ &= g.T(f)(h) = g.\langle h, \varphi(f) \rangle' = \langle h, g.\varphi(f) \rangle'. \end{aligned}$$

Aus $\langle h, \varphi(g.f) \rangle' = \langle h, g.\varphi(f) \rangle'$ folgt, dass $\varphi(g.f) = g.\varphi(f)$ für alle f aus $\mathcal{Y}^n_d(\mathbb{R})$ und g aus $\mathcal{O}(n+1)$ gilt. $\square$

Als nächtes zeigen wir, dass φ einen reellen Eigenwert $\lambda \in \mathbb{R}$ besitzt. Dazu zeigen wir, dass φ ein selbstadjungierter linearer Operator ist. Selbstadjungierte lineare Operatoren besitzen nur reelle Eigenwerte.

Behauptung 3.30. *Der lineare Operator $\varphi : \mathcal{Y}_d^n(\mathbb{R}) \to \mathcal{Y}_d^n(\mathbb{R})$ ist selbstadjungiert.*

Beweis. Sei $f, h \in \mathcal{Y}_d^n(\mathbb{R})$. Dann gilt:

$$\begin{aligned}\langle h,\varphi(f)\rangle' &= T(f)(h) = \langle h,f\rangle = \langle f,h\rangle \\ &= T(h)(f) = \langle f,\varphi(h)\rangle' = \langle \varphi(h),f\rangle'.\end{aligned}$$

Für alle $f, h \in \mathcal{Y}_d^n(\mathbb{R})$ gilt also $\langle h,\varphi(f)\rangle' = \langle \varphi(h),f\rangle'$. Daher ist φ selbstadjungiert. $\square$

Die Existenz eines Eigenwertes von $\varphi : \mathcal{Y}_d^n(\mathbb{R}) \to \mathcal{Y}_d^n(\mathbb{R})$ ist gesichert: Ein linearer Operator $\phi : V \to V$, V endlichdimensionaler Vektorraum über $\mathbb{C}$ oder $\mathbb{R}$ mit $\dim(V) \geq 1$, besitzt einen Eigenwert. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von φ, dann gilt $\lambda \in \mathbb{R}$, da φ selbstadjungiert ist.

Sei $\lambda \in \mathbb{R}$ ein reeller Eigenwert von φ, so gilt $\varphi = \lambda\,\mathrm{id}$ nach Bemerkung 2.7. Daher erhalten wir

$$\langle h,f\rangle = \langle h,\varphi(f)\rangle' = \lambda\,\langle h,f\rangle'$$

für alle f, h aus $\mathcal{Y}_d^n(\mathbb{R})$. Aus der positiven Definitheit der Skalarprodukte folgt, dass λ positiv ist. Insgesamt haben wir also gezeigt, dass orthogonal invariante Skalarprodukte auf $\mathcal{Y}_d^n(\mathbb{R})$ (bis auf Multiplikation mit einem positiven Skalar) eindeutig bestimmt sind. Diese Aussage halten wir in einem Satz fest.

Satz 3.31. *Auf jedem irreduziblen $\mathcal{O}(n+1)$-Untermodul $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ von $\mathcal{H}_d^n(\mathbb{R})$ gibt es ein (bis auf Multiplikation mit einem positiven Skalar) eindeutiges Skalarprodukt.*

Aus der Vorbemerkung (9) erhalten wir, dass $\mathcal{O}(n+1)$-invariante Sklarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ durch

$$\mathcal{H}_d^n(\mathbb{R}) \times \mathcal{H}_d^n(\mathbb{R}) \to \mathbb{R},\ (\oplus_{k=0}^{\lfloor d/2 \rfloor} f_k, \oplus_{k=0}^{\lfloor d/2 \rfloor} h_k) \mapsto \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \langle f_k,h_k\rangle_k \tag{23}$$

gegeben sind, wobei $c_0, \ldots, c_{\lfloor d/2 \rfloor} > 0$ gilt. Dabei bezeichnet $\langle \cdot,\cdot\rangle_k$ ein orthogonal invariantes Skalarprodukt auf $r^{2k}\mathcal{Y}_{d-2i}^n(\mathbb{R})$ und $\oplus_{i=1}^{\lfloor d/2 \rfloor} f_k$ die Zerlegung von $f \in \mathcal{H}_d^n(\mathbb{R})$ in Elemente f_k der irreduziblen Untermoduln $r^{2k}\mathcal{Y}_{d-2i}^n(\mathbb{R})$, $0 \leq k \leq \lfloor d/2 \rfloor$. Mit anderen Worten bedeutet dies: Orthogonal invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ sind durch $\lfloor d/2 \rfloor$-viele positive Zahlen parametrisiert.

4 Invariante normalverteilte Vektoren und Matrizen

In diesem Abschnitt möchten wir einige Eigenschaften von invarianten zufälligen (normalverteilten) Vektoren und Matrizen erarbeiten. Dabei meinen wir mit Invarainz, die Invarianz der Wahrscheinlichkeitsverteilungen unter der orthogonalen Gruppe. Diese Eigenschaften werden wir in Kapitel 5 und 6 wieder benötigen. Die Sätze und Definitionen, die hier vorkommen, haben wir zu großen Teilen aus [9, Ch. 3] entnommen.

4.1 Invariante zufällige Vektoren

Wir sagen, eine n-dimensionale reelle Zufallsvariable (oder Zufallsvektor) $Y = (Y_1, \dots, Y_n) \in \mathbb{R}^n$ ist ***normalverteilt*** mit Erwartungsvektor $\mu = (\mathbb{E}[Y_1], \dots, \mathbb{E}[Y_n])$ und positiv definiter Kovarianzmatrix $\Sigma = [\mathrm{Cov}(Y_i, Y_j)]_{1 \le i,j \le n}$, falls Y eine Dichtefunktion der Form

$$\rho_Y(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \tag{24}$$

besitzt. Wir schreiben dann $Y \sim N(\mu, \Sigma)$. Manchmal nennen wir Y nur ***zufällig*** und meinen damit, dass Y normalverteilt ist, falls nichts anderes gesagt wird. Ist die Erwartung μ gleich Null, so nennen wir den Zufallsvektor Y ***zentriert***.

Bemerkung 4.1. Ist $Y = (Y_1, \dots, Y_n) \in \mathbb{R}^n$ ein zufälliger Vektor und sind die $Y_1, \dots, Y_n$ paarweise unkorrelliert, so sind sie auch (stochastisch) unabhängig.

Wir nennen zwei zufällige Vektoren $Y^{(i)} \sim N(\mu_i, \Sigma_i)$, $i = 1, 2$, ***äquivalent***, falls beide die gleiche Wahrscheinlichkeitsverteilung besitzen, d.h. falls $\mu_1 = \mu_2$ und $\Sigma_1 = \Sigma_2$ gilt.

Definition 4.2. Sei $Y \in \mathbb{R}^n$ ein normalverteilter Zufallsvektor. Der Zufallsvektor Y heißt $\mathcal{O}(n)$***-invariant***, falls gY äquivalent zu Y ist für alle $g \in \mathcal{O}(n)$. In diesem Fall sagen wir einfach nur, dass Y ***invariant*** ist, falls die Dimension klar ist.

Ein invarianter Zufallsvektor $Y = (Y_1, \dots, Y_n) \in \mathbb{R}^n$ ist schon dadurch gekennzeichnet, dass er zentriert ist und die $Y_1, \dots, Y_n$ paarweise unabhängig sind und die gleiche Varianz haben.

Lemma 4.3. *Sei $Y = (Y_1, \dots, Y_n)$ ein reeller normalverteilter Zufallsvektor. Dann gilt folgende Äquivalenz: Y ist genau dann $\mathcal{O}(n)$-invariant, wenn Y zentriert ist und die $Y_1, \dots, Y_n$ paarweise unabhängig sind und die gleiche Varianz haben.*

Den Beweis haben wir aus [9, Ch. 3, Lemma 3.2] entnommen.

Beweis von Lemma 4.3. $\boxed{\Rightarrow}$ Angenommen Y ist $\mathcal{O}(n)$-invariant. Dann gilt: $-Y$ ist äquivalent zu Y, da nach Annahme gY äquivalent zu Y ist für alle $g \in \mathcal{O}(n)$. Wir wählen $g = -I_n$. Folglich gilt $\mathbb{E}\big[Y\big] = 0$.

Nach Annahme ist gYY^Tg^T äquivalent zu YY^T für alle $g \in \mathcal{O}(n)$. Die Kovarianzmatrix $A := \mathbb{E}\big[YY^T\big]$ erfüllt daher $gAg^T = A$ für alle $g \in \mathcal{O}(n)$. Eine direkte Rechnung zeigt, dass A ein Vielfaches der Einheitsmatrix I_n ist.

$\boxed{\Leftarrow}$ Diese Richtung erhalten wir aus der Eigenschaft, dass normalverteilte Vektoren durch ihre Erwartung und Kovarianzmatrix charakterisiert sind: Sei Y zentriert und seien die $Y_1, \dots, Y_n$ paarweise unabhängig mit gleicher Varianz. Bezeichnen wir mit σ die Varianz von Y_1, so gilt $A := \mathrm{Cov}(Y) = \sigma\, I_n$. Insbesondere gilt $gAg^T = A$ für alle g aus $\mathcal{O}(n)$. Der Zufallsvektor Y ist nach Voraussetzung zentriert, daher ist Y eindeutig durch seine Kovarianzmatrix A bestimmt. $\square$

4.2 Invariante zufällige symmetrische Matrizen

Eine ***zufällige*** Matrix ist eine Matrix $W \in \mathbb{R}^{n\times n}$, deren Einträge normalverteilte Zufallsvariablen sind. Bezeichne Σ_n den Raum der reellen, symmetrischen $(n \times n)$-Matrizen. Für W aus Σ_n definieren wir die ***Frobenius-Norm*** $\|W\|_F := \big(\sum_{i,j} W_{ij}^2\big)^{1/2}$. Eine zufällige Matrix $W \in \Sigma_n$ nennen wir ***$\mathcal{O}(n)$-invariant*** (kurz: ***invariant***), falls gWg^T für alle $g \in \mathcal{O}(n)$ äquivalent zu W ist. Im Folgenden nehmen wir $n > 1$ an.

Definition 4.4. Den ***Parameter*** $\delta(W)$ einer zufälligen Matrix $W \in \Sigma_n$ definieren wir als

$$\delta(W) := \frac{1}{n\,(n-1)}\Big(\mathbb{E}\Big[\,\mathrm{spur}(W)^2\Big] - \mathbb{E}\Big[\|W\|_F^2\Big]\Big).$$

Lemma 4.5. *Sei $(u, V, W) \in \mathbb{R} \times \mathbb{R}^n \times \Sigma_n$ normalverteilt mit gemeinsamer Verteilung, sodass (u, gV, gWg^T) äquivalent zu (u, V, W) ist für alle $g \in \mathcal{O}(n)$. Dann ist V unabhängig von u und W. In diesem Fall nennen wir (u, V, W) auch $\mathcal{O}(n)$-invariant.*

Beweis. Aufgrund der $\mathcal{O}(n)$-Invarianz sind zunächst V und W zentriert. Wählen wir $g = -I_n$, so erhalten wir: $(u, -V, W)$ ist äquivalent zu (u, V, W). Daher folgt, dass $\mathbb{E}\big[u\,V_i\big] = 0$ und $\mathbb{E}\big[V_i\,W_{jk}\big] = 0$ gilt für $i, j, k = 1, \dots, n$. $\square$

4.3 Die erwartete Determinante einer standardnormalverteilten Matrix

In diesem Abschnitt möchten wir die erwartete Determinante $\mathbb{E}\big[|\det W|\big]$ einer zufälligen Matrix $W \in \mathbb{R}^{n\times n}$ mit unabhängigen standardnormalverteilten Einträgen berechnen.

Definition 4.6. Seien $A_1, \ldots, A_n$ Vektoren im $\mathbb{R}^n$. Das von $A_1, \ldots, A_n$ aufgespannte ***Parallelotop*** $P(A_1, \ldots, A_n)$ definieren wir als

$$P(A_1, \ldots, A_n) := \left\{ \sum_{i=1}^{n} a_i A_i \mid 0 \leq a_i \leq 1, \text{ für alle } i = 1, \ldots, n \right\}.$$

Außerdem bezeichnen wir mit $\operatorname{Vol}(A_1, \ldots, A_n)$ das Volumen von $P(A_1, \ldots, A_n)$ in $\mathbb{R}^n$.

Sei im Folgenden $W \in \mathbb{R}^{n \times n}$ eine Matrix und die $W_1, \ldots, W_n \in \mathbb{R}^n$ bezeichen die Zeilenvektoren von W. Wir können die Determinante $|\det W|$ von W durch das Volumen des Paralletops, das von den Zeilenvektoren $W_1, \ldots, W_n$ von W aufgespannt wird, angeben. Bekanntlich gilt:

$$\operatorname{Vol}(W_1, \ldots, W_n)^2 = \det\left(W^T W\right).$$

Insbesondere gilt: $\operatorname{Vol}(W_1, \ldots, W_n) = |\det W|$.

Außerdem bezeichnen wir mit $W_i^{\perp}$ die orthogonale Projektion von W_i auf $\operatorname{span}\{W_1, \ldots, W_{i-1}\}^{\perp}$ für $i = 1, \ldots, n$. Dabei bezeichnet $\operatorname{span}\{W_1, \ldots, W_{i-1}\}^{\perp}$ das orthogonale Komplement von $\operatorname{span}\{W_1, \ldots, W_{i-1}\}$ in $\mathbb{R}^n$.

Wir möchten nun das Volumen des Parallelotops $P(W_1, \ldots, W_n)$ über die orthogonalen Projektionen $W_i^{\perp}$ bestimmen. Dazu nutzen wir das Prinzip "Volumen = (Volumen der Basis) × Höhe". Es gilt:

$$\operatorname{Vol}(W_1, \ldots, W_n) = \operatorname{Vol}(W_1, \ldots, W_{n-1}) \cdot \left\|W_n^{\perp}\right\|. \tag{25}$$

Sei $X \in \mathbb{R}^m$ standard-normalverteilt. Dann definieren wir $K_m := \mathbb{E}\big[\|X\|\big]$. Eine einfache Rechnung ergibt dann (vgl. [11]):

$$K_m = \sqrt{2}\,\frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}, \quad L_m := K_1 K_2 \cdots K_m = \frac{1}{\sqrt{\pi}}\, 2^{\frac{m}{2}}\, \Gamma\left(\frac{m+1}{2}\right). \tag{26}$$

Dabei bezeichnet $\Gamma(x)$ die ***Eulersche Gammafunktion***.

Die erwartete Determinante $\mathbb{E}\big[|\det W|\big]$ einer zufälligen Matrix $W \in \mathbb{R}^{n \times n}$ mit unabhängigen standard-normalverteilten Einträgen ist dann das Produkt $K_1 \cdots K_n$.

Lemma 4.7. *Sei $W \in \mathbb{R}^{n \times n}$ eine zufällige Matrix mit unabhängigen, standard-normalverteilten Einträgen. Dann gilt:*

$$\mathbb{E}\Big[|\det W|\Big] = K_n K_{n-1} \cdots K_1 = L_n.$$

Beweis. Wir bezeichen mit $W_1, \ldots, W_n \in \mathbb{R}^n$ die Zeilenvektoren von W. Aufgrund der $\mathcal{O}(n)$-Invarianz der Standardnormalverteilung folgt, dass $W_i^{\perp}$ bedingt auf $W_1, \ldots, W_{i-1}$ ebenfalls standard-normalverteilt in $\operatorname{span}\{W_1, \ldots, W_{i-1}\}^{\perp}$ für $i = 1, \ldots, n$ ist. Zudem hat der Raum

$\mathrm{span}\{W_1, \ldots, W_{i-1}\}^\perp$ die Dimension $n - i + 1$ mit Wahrscheinlichkeit 1 für $i = 1, \ldots, n$. Daher gilt:

$$\mathbb{E}\Big[\|W_i^\perp\| \Big| W_1, \ldots, W_{i-1}\Big] = K_{n-i+1}, \quad i = 1, \ldots, n. \tag{27}$$

Dabei bezeichnet $\mathbb{E}[\,\cdot\,|\,\cdot\,]$ die bedingte Erwartung. Mit (25) und (27) folgt

$$\begin{aligned}\mathbb{E}\Big[|\det W|\Big] &= \mathbb{E}\Big[\operatorname{Vol}(W_1, \ldots, W_{n-1})\ \mathbb{E}\Big[\|W_n^\perp\| \Big| W_1, \ldots, W_{n-1}\Big]\Big] \\ &= \mathbb{E}\Big[\operatorname{Vol}(W_1, \ldots, W_{n-1})\Big]\, K_1.\end{aligned}$$

Wir fahren rekursiv fort und erhalten

$$\mathbb{E}\Big[|\det W|\Big] = K_n K_{n-1} \cdots K_1 = L_n.$$

□

5 Invariante zufällige Polynome

In diesem Kapitel möchten wir invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{C})$ und $\mathcal{H}_d^n(\mathbb{R})$ klassifizieren. Wir sagen, dass ein Polynom f aus $\mathcal{H}_d^n(\mathbb{C})$ bzw. $\mathcal{H}_d^n(\mathbb{R})$ ***zufällig*** ist, falls die Koeffizienten von f normalverteilt sind. Unitär bzw. orthogonal invariante zufällige Polynome sind dadurch gekennzeichnet, dass ihre Wahrscheinlichkeitsverteilungen unter der Gruppenaktion der unitären bzw. orthogonalen Gruppe invariant bleiben. Um nun invariante zufällige zentrierte Polynome zu klassifizieren, verwenden wir die Ergebnisse aus Kapitel 3. Wir bemerken, dass ein invariantes Skalarprodukt $\langle\cdot,\cdot\rangle$ auf einem Vektorraum V über die Dichte

$$\rho(v) = K\, \exp\Big(-\frac{1}{2}\langle v, v\rangle\Big), \quad K \in \mathbb{R}, \tag{28}$$

eine invariante zentrierte Normalverteilung auf V definiert.

In Abschnitt 3.1 haben wir gezeigt, dass ein unitär invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ bis auf Skalierung dem Weylschen Skalarprodukt entspricht. In Abschnitt 5.2.1 werden wir sehen, dass wir dadurch bis auf Skalierung nur eine Möglichkeit haben, unitär invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{C})$ zu wählen. Daher beschränken wir uns im Folgenden auf reelle zufällige Polynome.

Auf $\mathcal{H}_d^n(\mathbb{R})$ sind orthogonal invariante Skalarprodukte durch $\lfloor d/2 \rfloor$-viele positive Zahlen parametrisiert (vgl. Abschnitt 3.2). Daraus erhalten wir, dass orthogonal invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ durch $\lfloor d/2 \rfloor$-viele positive Zahlen parametrisiert sind.

Definition 5.1. Ein zufälliges Polynom $f \in \mathcal{H}_d^n(\mathbb{R})$ nennen wir $\mathcal{O}(n+1)$-***invariant*** (kurz: ***invariant***), falls $g.f$ äquivalent zu f für alle $g \in \mathcal{O}(n+1)$ ist.

Ein zufälliger Vektor $Y \in \mathbb{R}^n$ ist durch seinen Erwartungvektor und seine Kovarianzmatrix gegeben (vgl. Abschnitt 4.1). Wählen wir eine Basis von $\mathcal{H}_d^n(\mathbb{R})$, so ist ein zufälliges Polynom f aus $\mathcal{H}_d^n(\mathbb{R})$ eindeutig durch seinen Erwartungsvektor und seine Kovarianzmatrix bestimmt. Ist f zentriert, so ist f bereits durch seine Kovarianzmatrix eindeutig bestimmt.

Definition 5.2. Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein zufälliges Polynom. Die ***Kovarianzfunktion*** r_f von f ist definiert durch

$$r_f : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}, \quad (x,y) \mapsto \mathbb{E}\Big[f(x)f(y)\Big].$$

Bemerkung 5.3. Die Kovarianzmatrix eines invarianten zufälligen zentrierten Polynoms f aus $\mathcal{H}_d^n(\mathbb{R})$ ist schon eindeutig durch seine Kovarianzfunktion r_f bestimmt. Daher sind invariante zufällige zentrierte Polynome f aus $\mathcal{H}_d^n(\mathbb{R})$ eindeutig durch ihre Kovarianzfunktionen r_f bestimmt.

Wir geben jetzt ein Beispiel für ein invariantes zufälliges zentriertes Polynom.

Beispiel 5.4. Wir erinnern an die Weylsche Basis und an das Weylsche Skalarprodukt aus Abschnitt 3.1, welche jetzt auf $\mathcal{H}_d^n(\mathbb{R})$ gegeben sind. Dazu schreiben wir f aus $\mathcal{H}_d^n(\mathbb{R})$ in der Form $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$ mit $f_\alpha \in \mathbb{R}$. Wählen wir die Koeffizienten f_α normalverteilt, zentriert und paarweise unabhängig mit Varianz $\mathrm{Var}(f_\alpha) = \binom{d}{\alpha}$, so erhalten wir für die Kovarianzfunktion von f

$$\begin{aligned} r_f(x,y) &= \mathbb{E}\Big[f(x)f(y)\Big] \\ &= \sum_{|\alpha|,|\beta|=d} \mathbb{E}\Big[f_\alpha f_\beta\Big] x^\alpha y^\beta = \sum_{|\alpha|=d} \tbinom{d}{\alpha} x^\alpha y^\alpha \\ &= \left(\sum_{i=0}^{n} x_i y_i\right)^d = \langle x,y \rangle^d. \end{aligned}$$

Dabei bezeichnet $\langle \cdot,\cdot \rangle$ das Standardskalarprodukt auf $\mathbb{R}^{n+1}$. Daher ist die Kovarianzfunktion von f $\mathcal{O}(n+1)$-invariant. Damit folgt insbesondere, dass f ein $\mathcal{O}(n+1)$-invariantes zufälliges zentriertes Polynom ist. Ein Polynom dieser Gestalt nennen wir ***Kostlan-verteilt***.

5.1 Der Parameter eines invarianten zufälligen zentrierten Polynoms

Wir möchten einem invarianten zufälligen zentrierten Polynom f aus $\mathcal{H}_d^n(\mathbb{R})$ einen Wert $\delta(f)$ zuordnen, den wir den ***Parameter*** des Polynoms f nennen und dessen Quadratwurzel $\sqrt{\delta(f)}$ als der erwartete Grad von f interpretiert werden kann.

Wir beobachten zunächst, dass ein Polynom $f \in \mathcal{H}_d^n(\mathbb{R})$ eine differenzierbare Abbildung $f : \mathbb{S}^n \to \mathbb{R}$ definiert. Weiter bezeichnen wir mit

$\mathrm{D}f(x) : \mathrm{T}_x\,\mathbb{S}^n \to \mathbb{R}$ und $\mathrm{D}^2f(x) : \mathrm{T}_x\,\mathbb{S}^n \times \mathrm{T}_x\,\mathbb{S}^n \to \mathbb{R}$ die erste und zweite Ableitung von f an der Stelle $x \in \mathbb{S}^n$, welche durch

$$f\left(\frac{x+\lambda\,v}{\|x+\lambda\,v\|}\right) = f(x) + \lambda\;\mathrm{D}f(x)(v) + \frac{\lambda^2}{2}\,\mathrm{D}^2f(x)(v,v) + o(\lambda^2) \qquad (29)$$

für $v \in \mathrm{T}_x\,\mathbb{S}^n$ und $\lambda \in \mathbb{R}$, wobei λ geht gegen Null strebt, charakterisiert sind. Im Punkt $q := (1,0,\ldots,0) \in \mathbb{S}^n$ können wir den Tangentialraum $\mathrm{T}_q\,\mathbb{S}^n = \{0\}\times\mathbb{R}^n$ mit $\mathbb{R}^n$ identifizieren. In diesem Fall ist die erste Ableitung $\mathrm{D}f(q)$ durch den Gradienten $(\partial_1 f(q),\ldots,\partial_n f(q))$ gegeben.

Lemma 5.5. *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein Polynom. Dann gilt:* $\mathrm{D}^2f(q)$ *ist durch die Matrix* $[\partial^2_{kl} f(q)]_{1\le k,l\le n} - d\,f(q)\,I_n$ *gegeben.*

Beweis. Wir betrachten $g(t) := f\left((1+t^2)^{-1/2}(1,tv)\right)$ für $v \in \mathbb{R}^n$ mit $\|v\| = 1$. Da f homogen vom Grad d ist, gilt $g(t) = (1+t^2)^{-d/2} f(1,tv)$. Nach der Definition der zweiten Ableitung in (29) erhalten wir mit $\|(1,tv)\| = \sqrt{1+t^2}$, dass $\mathrm{D}^2f(q)(v,v) = g''(0)$ gilt. Für die ersten beiden Ableitungen von g nach t gilt:

$$\frac{d}{dt}g(t) = -t\,d\,(1+t^2)^{-\frac{d}{2}-1}\,f(1,tv) + (1+t^2)^{-d/2}\sum_{k=1}^{n}\partial_k\,f(1,tv)\,v_k$$

und

$$\begin{aligned}\frac{d^2}{dt^2}g(t) = &-d\,(1+t^2)^{-\frac{d}{2}-1}\,f(1,tv)\\ &+(1+t^2)^{-d/2}\sum_{k,l=1}^{n}\partial_l\,\partial_k\,f(1,tv)\,v_k\,v_l + \text{Terme in } t.\end{aligned}$$

Daher erhalten wir

$$\frac{d^2}{dt^2}g(0) = \left(\sum_{k,l=1}^{n}\partial^2_{kl}\,f(q)\,v_k\,v_l\right) - d\,f(q).$$

□

Wir definieren nun den Parameter eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$.

Definition 5.6. Den ***Parameter*** $\delta(f)$ eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ definieren wir durch

$$\delta(f) := \frac{\|x\|^2\,\mathbb{E}\Big[\|\mathrm{D}f(x)\|^2\Big]}{n\,\mathbb{E}\Big[f(x)^2\Big]}$$

für $x \in \mathbb{R}^{n+1}$ verschieden von Null.

Bemerkung 5.7. Aufgrund der $\mathcal{O}(n+1)$-Invarianz der Normalverteilung und der Homogenität von f ist der Parameter von f unabhängig vom gewählten $x \in \mathbb{R}^{n+1} \setminus \{0\}$.

Es ist klar, dass die Einschränkung f' eines $\mathcal{O}(n+1)$-invarianten zufälligen Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ auf $\mathcal{H}_d^{n'}(\mathbb{R})$, $n' \leq n$, ebenfalls $\mathcal{O}(n'+1)$-invariant ist. Dabei ensteht die Einschränkung f' von f auf $\mathcal{H}_d^{n'}(\mathbb{R})$ durch das Ersetzen der Variablen $X_{n'+1}, \dots, X_n$ durch Null.

Wir zeigen jetzt, dass die Invarianz von $f \in \mathcal{H}_d^n(\mathbb{R})$ auch beim Übergang zur ersten und zweiten Ableitung von f erhalten bleibt.

Lemma 5.8. *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein $\mathcal{O}(n+1)$-invariantes zufälliges zentriertes Polynom. Dann ist $\big(f(q), \mathrm{D}f(q), \mathrm{D}^2 f(q)\big) \in \mathbb{R} \times \mathbb{R}^n \times \Sigma_n$ $\mathcal{O}(n)$-invariant im Sinne von Lemma 4.5.*

Für den Beweis folgen wir [9, Ch. 4.2, Lemma 4.8].

Beweis von Lemma 5.8. Wir fixieren zunächst ein $g \in \mathcal{O}(n)$. Wir betrachten das Polynom $h := g.f$. Nach Annahme hat h dieselbe Wahrscheinlichkeitsverteilung wie f. Betrachten wir die Darstellung von f und h bezüglich einer festen Basis von $\mathcal{H}_d^n(\mathbb{R})$ mit zufälligen reellen Koeffizienten, so sehen wir, dass $(h, \partial_i h, \partial_{ij}^2 h)$ äquivalent zu $(f, \partial_i f, \partial_{ij}^2 f)$ ist für alle $i, j = 1, \dots, n$. Damit erhalten wir, dass $\big(h(q), \mathrm{D}h(q), \mathrm{D}^2 h(q)\big) = \big(f(q), g\,\mathrm{D}f(q), g\,\mathrm{D}^2 f(q) g^T\big)$ äquivalent zu $\big(f(q), \mathrm{D}f(q), \mathrm{D}^2 f(q)\big)$ ist. $\square$

Aus Lemma 5.8 können wir folgern, dass $\mathrm{D}f(q)$ (stochastisch) unabhängig von $f(q)$ und $\mathrm{D}^2 f(q)$ ist.

Korollar 5.9. *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein $\mathcal{O}(n+1)$-invariantes zufälliges zentriertes Polynom. Dann gilt: $\mathrm{D}f(q)$ ist unabhängig von $f(q)$ und $\mathrm{D}^2 f(q)$.*

Beweis. Nach Lemma 5.8 ist $\big(f(q), \mathrm{D}f(q), \mathrm{D}^2 f(q)\big)$ $\mathcal{O}(n)$-invariant. Daher ist nach Lemma 4.5 $\mathrm{D}f(q)$ unabhängig von $f(q)$ und $\mathrm{D}^2 f(q)$. $\square$

Wir möchten zeigen, dass die Kovarianzmatrix $\mathrm{Cov}\,(\mathrm{D}f(q))$ der Ableitung $\mathrm{D}f(q)$ eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ ein skalares Vielfaches der Matrix $\delta(f)\, I_n$ ist.

Proposition 5.10. *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein invariantes zufälliges zentriertes Polynom und $q := (1, 0, \dots, 0) \in \mathbb{S}^n$. Dann ist $\mathrm{D}f(q)$ normalverteilt mit Kovarianzmatrix $\mathrm{Cov}\,(\mathrm{D}f(q)) = \delta(f)\, \mathbb{E}\big[f(q)^2\big]\, I_n$.*

Beweis. Für die partielle Ableitung ∂_i nach X_i von $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$ an der Stelle q gilt zunächst

$$\partial_i f(q) = f_{(d-1,0,\dots,0,\underset{\substack{\uparrow \\ i\text{-te Stelle}}}{1},0,\dots,0)}, \qquad 1 \leq i \leq n.$$

Damit ist $\mathrm{D}f(q) = (\partial_1 f(q), \dots, \partial_n f(q))$ offensichtlich normalverteilt. Mit Lemma 5.8 folgt sogar, dass $\mathrm{D}f(q)$ $\mathcal{O}(n)$-invariant ist. Nach Lemma

4.3 ist $\mathrm{D}f(q)$ zentriert und die $\partial_1 f(q), \ldots, \partial_n f(q)$ sind unabhängig und haben dieselbe Varianz. D.h. es gilt $\mathbb{E}\Big[\partial_i f(q)\partial_j f(q)\Big] = 0$ für $i \neq j$ und $\mathbb{E}\Big[(\partial_1 f(q))^2\Big] = \mathbb{E}\Big[(\partial_i f(q))^2\Big]$ für $i = 1, \ldots, n$. Somit erhalten wir

$$\begin{aligned} \mathbb{E}\Big[(\partial_i f(q))^2\Big] &= \frac{1}{n}\,\mathbb{E}\Big[(\partial_1 f(q))^2 + \cdots + (\partial_n f(q))^2\Big] \\ &= \frac{1}{n}\,\mathbb{E}\Big[\|\mathrm{D}f(q)\|^2\Big] \\ &= \delta(f)\,\mathbb{E}\Big[f(q)^2\Big]. \end{aligned} \tag{30}$$

Insbesondere folgt damit, dass die Kovarianzmatrix von $\mathrm{D}f(q)$ durch $\delta(f)\,\mathbb{E}\Big[f(q)^2\Big]\, I_n$ gegeben ist.

□

Aus Gleichung (30) erhalten wir sofort eine einfache Formel zur Berechnung des Parameters eines Polynoms.

Korollar 5.11. *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein invariantes zufälliges zentriertes Polynom und $q := (1, 0, \ldots, 0) \in \mathbb{S}^n$. Für den Parameter $\delta(f)$ von f gilt:*

$$\delta(f) = \frac{\mathbb{E}\Big[\big(\partial_i f(q)\big)^2\Big]}{\mathbb{E}\Big[f(q)^2\Big]}, \quad i = 1, \ldots, n.$$

Wir werden jetzt den Parameter eines Kostlan-verteilten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ bestimmen.

Beispiel 5.12. Ist $f \in \mathcal{H}_d^n(\mathbb{R})$ Kostlan-verteilt, so gilt $\delta(f) = d$.

Beweis. Nach Korollar 5.11 brauchen wir nur $\mathbb{E}\big[(\partial_1 f(q))^2\big]$ und $\mathbb{E}\big[f(q)^2\big]$ berechnen. Dazu schreiben wir f wieder in der Form $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$. Dann gilt:

$$f(q) = f_{(d,0,\ldots,0)} \text{ und } \partial_1 f(q) = f_{(d-1,1,0\ldots,0)}.$$

Damit folgt:

$$\mathbb{E}\Big[(\partial_1 f(q))^2\Big] = \mathrm{Var}(f_{(d-1,1,0\ldots,0)}) = \binom{d}{(d-1,1,0\ldots,0)} = \frac{d!}{(d-1)!} = d$$

und

$$\mathbb{E}\Big[f(q)^2\Big] = \mathrm{Var}(f_{(d,0,\ldots,0)}) = \binom{d}{d} = \frac{d!}{d!} = 1.$$

Wir erhalten:

$$\delta(f) = \frac{\mathbb{E}\Big[(\partial_1 f(q))^2\Big]}{\mathbb{E}\Big[f(q)^2\Big]} = \frac{d}{1} = d.$$

□

Zum Abschluss dieses Abschnitts möchten wir noch bemerken, dass der Parameter $\delta(f)$ eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$, bis auf Skalierung, gleich dem Parameter $\delta(\mathrm{D}^2 f(q))$ der Matrix $\mathrm{D}^2 f(q)$ ist (vgl. Definition 4.4).

Lemma 5.13 ([9, Ch. 4.2, Prop. 4.9])**.** *Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein invariantes zufälliges zentriertes Polynom. Dann gilt: $\delta\big(\mathrm{D}^2 f(q)\big) = \delta(f)\,\mathbb{E}\big[f(q)^2\big]$.*

Beweis. Einen vollständigen Beweis findet man in [9, Appendix]. Wir geben hier nur die Beweisschritte an.

Wir bezeichnen mit $\partial_{x_i} f(x)$ die partielle Ableitung von $f \in \mathcal{H}_d^n(\mathbb{R})$ nach x_i, $1 \leq i \leq n$. Seien $a_{ij} := \partial^2_{x_i x_j} f(q)$ und $A := [a_{ij}]_{1 \leq i \leq n}$. Nach Lemma 5.5 gilt $W := \mathrm{D}^2 f(q) = A - d\, f(q)\, I_n$. Wir berechnen zunächst $\mathbb{E}\big[\operatorname{spur}(W)^2\big]$ und $\mathbb{E}\big[\|W\|_F^2\big]$. Es gilt:

$$\mathbb{E}\Big[\operatorname{spur}\big(W\big)^2\Big] = \mathbb{E}\Big[\operatorname{spur}\big(A\big)^2\Big] - 2\,n\,d\,\mathbb{E}\Big[\operatorname{spur}\big(A\big) f(q)\Big] + n^2 d^2\,\mathbb{E}\Big[f(q)^2\Big]$$

und

$$\mathbb{E}\Big[\|W\|_F^2\Big] = \mathbb{E}\Big[\|A\|_F^2\Big] - 2\,d\,\mathbb{E}\Big[\operatorname{spur}\big(A\big) f(q)\Big] + n\,d^2\,\mathbb{E}\Big[f(q)^2\Big].$$

Eingesetzt in Definition 4.4 ergibt dies

$$\begin{aligned}
\delta(W) &= \frac{1}{n\,(n-1)} \Big(\mathbb{E}\Big[\operatorname{spur}\big(W\big)^2\Big] - \mathbb{E}\Big[\|W\|_F^2\Big]\Big) \\
&= \frac{1}{n\,(n-1)} \sum_{i \neq j} \mathbb{E}\Big[a_{ii} a_{jj} - a_{ij}^2\Big] - \frac{2}{n}\, d\,\mathbb{E}\Big[\operatorname{spur}\big(A\big) f(q)\Big] + d^2\,\mathbb{E}\Big[f(q)^2\Big].
\end{aligned}$$

Nach Gleichung (30) gilt

$$\delta(f)\,\mathbb{E}\Big[f(q)^2\Big] = \frac{1}{n} \sum_i \mathbb{E}\Big[\big(\partial_{x_i} f(q)\big)^2\Big].$$

Berechnen wir die Erwartungen wie in (36) mit der Kovarianzfunktion (35) und vergleichen die oberen beiden Gleichungen, so sieht man, dass $\delta\big(\mathrm{D}^2 f(q)\big) = \delta(f)\,\mathbb{E}\big[f(q)^2\big]$ gilt.

□

5.2 Klassifikation invarianter zufälliger zentrierter Polynome

Wir werden in diesem Abschnitt orthogonal invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ klassifizieren. Die Klassifikation aller orthogonal invarianten zufälligen zentrierten Polynome f aus $\mathcal{H}_d^n(\mathbb{R})$ ermöglicht uns die Erwartungen aus dem Parameter $\delta(f)$ von f (vgl. Definition 5.6) zu berechnen.

Haben wir allgemein einen endlichdimensionalen reellen Vektorraum V, so ist die Wahl einer zentrierten Normalverteilung auf V äquivalent zur Wahl eines Skalarprodukts auf V (vgl. (28)). Üblicherweise werden zentrierte Normalverteilungen durch (positiv definite) Kovarianzmatrizen beschrieben. In Koordinaten-freier Sprache korrespondieren diese Kovarianzmatrizen zu einem Skalarprodukt auf dem Dualraum V^* von V, definiert durch

$$V^* \times V^* \to \mathbb{R}, \quad (\lambda, \mu) \mapsto \mathbb{E}_v\Big[\lambda(v)\mu(v)\Big]. \tag{31}$$

Um orthogonal invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ zu klassifizieren bemerken wir, dass das Skalarprodukt auf dem Dualraum $\mathcal{H}_d^n(\mathbb{R})^*$ in (31) der Kovarianzfunktion aus Definition 5.2 entspricht: Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ und definieren wir $\lambda, \mu \in \mathcal{H}_d^n(\mathbb{R})^*$ durch $\lambda(f) := f(y)$ und analog $\mu(f) := f(y)$ für $x, y \in \mathbb{R}^{n+1}$, so erhalten wir ein Skalarprodukt $\mathcal{H}_d^n(\mathbb{R})^* \times \mathcal{H}_d^n(\mathbb{R})^* \to \mathbb{R}$ mit

$$(\lambda, \mu) \mapsto \mathbb{E}\Big[\lambda(f)\mu(f)\Big] = \mathbb{E}\Big[f(x)f(y)\Big] = r_f(x, y). \tag{32}$$

Daher können wir orthogonal invariante zufällige zentrierte Polynome mithilfe ihrer Kovarianzfunktion beschreiben. Es wird sich zeigen, dass die Kovarianzfunktion eines orthogonal invarianten zufälligen zentrierten Polynoms aus $\mathcal{H}_d^n(\mathbb{R})$ durch $\lfloor d/2 \rfloor$-viele positive Zahlen beschrieben werden kann. Damit können insbesondere orthogonal invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ durch $\lfloor d/2 \rfloor$-viele positive Zahlen parametrisiert werden.

5.2.1 Klassifikation unitär invarianter zufälliger Polynome

Wir erinnern uns daran, dass wir uns auf orthogonal invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ beschränkt haben, da der komplexe Fall von geringem Interesse ist: Auf $\mathcal{H}_d^n(\mathbb{C})$ gibt es bis auf Skalierung nur eine Möglichkeit unitär invariante zufällige zentrierte Polynome zu wählen.

Sei $\langle \cdot, \cdot \rangle_{\mathcal{H}_d^n(\mathbb{C})}$ das Weylsche Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ (vgl. (5)). Aus Abschnitt 3.1.2 wissen wir, dass das Weylsche Skalarprodukt (bis auf Skalierung mit einem positiven Skalar) das einzige unitär invariante Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ ist. Bezüglich $\langle \cdot, \cdot \rangle_{\mathcal{H}_d^n(\mathbb{C})}$ sind verschiedene Monome X^α orthogonal zueinander. Insbesondere sind verschiedene Monome bezüglich jedem unitär invarianten Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ orthogonal zueinander.

Wir haben in (28) gesehen, dass eine unitär invariante zentrierte Normalverteilung auf $\mathcal{H}_d^n(\mathbb{C})$ zu einem unitär invarianten Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{C})$ korrespondiert. Daher erhalten wir aus den Eigenschaften des Weylschen Skalarprodukts, dass für ein unitär invariantes zufälliges zentriertes Polynom $f \in \mathcal{H}_d^n(\mathbb{C})$ Folgendes gilt: Die Koeffizienten von f sind paarweise unabhängig und für die Varianzen der Koeffizienten f_α gilt bis auf Skalierung $\mathrm{Var}(f_\alpha) = \binom{d}{\alpha}$.

5.2.2 Klassifikation indefiniter Skalarprodukte

Wir wissen, dass (positiv definite) Kovarianzmatrizen von orthogonal invarianten zufälligen zentrierten Polynomen aus $\mathcal{H}_d^n(\mathbb{R})$ zu orthogonal invarianten Skalarprodukten auf $\mathcal{H}_d^n(\mathbb{R})^*$ korrespondieren (vgl. (31)).

Sei $e_1, \ldots, e_N$ eine Basis von $\mathcal{H}_d^n(\mathbb{R})$, $N := \binom{d+n}{d}$ und C sei eine (positiv definite) Kovarianzmatrix eines orthogonal invarianten zufälligen zentrierten Polynoms. Setzen wir $v(x) := \big(e_1(x), \ldots, e_N(x)\big)^T$ für $x \in \mathbb{R}^{n+1}$, so definiert die Kovarianzmatrix C durch

$$v(x)^T \, C \, v(y), \quad x, y \in \mathbb{R}^{n+1} \tag{33}$$

eine orthogonal invariante (positiv definite) symmetrische Bilinearform auf $\mathcal{H}_d^n(\mathbb{R})^*$. Insbesondere definiert (33) ein orthogonal invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})^*$. Umgekehrt gilt auch, dass ein (positiv definites) Skalarprodukt eine symmetrische positiv definite Matrix C definiert.

Sei jetzt $e_1, \ldots, e_N$ die Monombasis von $\mathcal{H}_d^n(\mathbb{R})$, d.h. $e_i = X^{\alpha^{(i)}}$, $\alpha^{(i)} \in \mathbb{N}^{n+1}$ und $\left|\alpha^{(i)}\right| = d$, $i = 1, \ldots, N$.

Satz 5.14. *Sei $e_1, \ldots, e_N$ die Monombasis von $\mathcal{H}_d^n(\mathbb{R})$, $v(x) := \big(e_1(x), \ldots, e_N(x)\big)^T$ für $x \in \mathbb{R}^{n+1}$ und C die Kovarianzmatrix eines orthogonal invarianten zufälligen zentrierten Polynoms aus $\mathcal{H}_d^n(\mathbb{R})$. Dann gilt:*

$$v(x)^T \, C \, v(y) = \sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k \, \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k} \tag{34}$$

für gewisse $\beta_k \in \mathbb{R}$, $0 \leq k \leq \lfloor d/2 \rfloor$.

Beweis. Zunächst ist $v(x)^T \, C \, v(y)$ ein homogenes Polynom vom Grad d sowohl in x als auch in y. Nach Voraussetzung ist $v(x)^T \, C \, v(y)$ orthogonal invariant, d.h. es gilt $v(gx)^T \, C \, v(gy) = v(x)^T \, C \, v(y)$ für alle g aus $\mathcal{O}(n+1)$ und alle x, y aus $\mathbb{R}^{n+1}$. Es ist leicht einzusehen, dass $v(x)^T \, C \, v(y)$ als Funktion von $\|x\|$, $\|y\|$ und $\langle x,y \rangle$ geschrieben werden kann. Dabei bezeichent $\langle \cdot,\cdot \rangle$ das Standardskalarprodukt auf $\mathbb{R}^{n+1}$. Sei also $v(x)^T \, C \, v(y) = F(\|x\|, \|x\|, \langle x,y \rangle)$ eine Funktion von $\|x\|$, $\|y\|$ und $\langle x,y \rangle$ für alle $x, y \in \mathbb{R}^{n+1}$. Invariantentheorie [12, Cp. 13, Section 7] zeigt, dass F sogar ein reelles Polynom ist und wir können $v(x)^T \, C \, v(y)$ als ein Polynom in $\|x\|$, $\|y\|$ und $\langle x,y \rangle$ ausdrücken. Da $v(x)^T \, C \, v(y)$ ein homogenens Polynom vom Grad d sowohl in x als auch in y ist, muss die Potenz von $\|x\|$ und $\|y\|$ in F eine gerade Zahl sein. Wir erhalten, dass $v(x)^T \, C \, v(y)$ die folgende Form hat:

$$v(x)^T \, C \, v(y) = \sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k \, \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k}, \quad \beta_k \in \mathbb{R}.$$

$\square$

Wir haben also gezeigt, dass orthogonal invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})^*$ durch $\lfloor d/2 \rfloor$-viele reelle Zahlen beschrieben werden können. Aus der

Vorbemerkung (32) erhalten wir, dass Gleichung (34) auch als Kovarianzfunktion geschrieben werden kann.

Korollar 5.15. *Für jedes orthogonal invariante zufällige zentrierte Polynom $f \in \mathcal{H}_d^n(\mathbb{R})$ können wir die Kovarianzfunktion $r_f(x,y)$ in der folgenden Form schreiben:*

$$r_f(x,y) = \mathbb{E}\Big[f(x)f(y)\Big] = \sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k \, \|x\|^{2k} \, \|y\|^{2k} \, \langle x,y \rangle^{d-2k} \tag{35}$$

für gewisse $\beta_k \in \mathbb{R}$, $0 \leq k \leq \lfloor d/2 \rfloor$.

Korollar 5.15 klassifiziert damit alle orthogonal invarianten zufälligen zentrierten Polynome aus $\mathcal{H}_d^n(\mathbb{R})$. Aber nicht jede Wahl der $\beta_k \in \mathbb{R}$, $0 \leq k \leq \lfloor d/2 \rfloor$, in (35) ergibt eine Kovarianzfunktion eines invarianten zufälligen zentrierten Polynoms f aus $\mathcal{H}_d^n(\mathbb{R})$.

Beispiel 5.16. Wählen wir $\beta_0 = 1$ und $\beta_k = 0$ für $0 < k \leq \lfloor d/2 \rfloor$ in (35), so erhalten wir die Kovarianzfunktion der Kostlan-Verteilung.

Bemerkung 5.17. Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein orthogonal invariantes zufälliges zentriertes Polynom mit Kovarianzfunktion $r_f(x,y) = \beta_0 \, \langle x,y \rangle^d$, $\beta_0 \in \mathbb{R} \setminus \{0\}$. Dann gilt: β_0 ist größer als Null und die Koeffizienten f_α von f sind unabhängig und die Varianzen der Koeffizienten f_α sind bis auf Skalierung gleich $\binom{d}{\alpha}$. Mit anderen Worten bedeutet dies: Die Wahl $\beta_0 > 0$ und $\beta_k = 0$ für $0 < k \leq \lfloor d/2 \rfloor$ in (35) ist die einzige, bei der die Koeffizienten f_α von f unabhängig sind und deren Varianzen bis auf Skalierung gleich $\binom{d}{\alpha}$ sind. Jede andere Kombination führt dazu, dass die Koeffizienten abhängig sind.

Wir möchten den Parameter eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ berechnen, dessen Kovarianzfunktion $r_f(x,y)$ in der Form (35) gegeben ist. Sei dazu $r_f(x,y)$ wie in (35), ∂_{x_i} die partielle Ableitung nach x_i, $i = 1, \ldots, n$ und $q = (1, 0, \ldots, 0) \in \mathbb{S}^n$. Wir beobachten, dass der (i,j)-te Eintrag in der Kovarianzmatrix von $\mathrm{D}f(q)$ durch

$$\mathbb{E}\Big[\big(\partial_{x_i} f(q)\big)\big(\partial_{x_j} f(q)\big)\Big] = \partial_{x_i}\partial_{y_i} r_f(x,y)\Big|_{x=y=q} = \partial_{x_i}\partial_{y_i} r_f(q,q) \tag{36}$$

gegeben ist. Wir berechnen die partielle Ableitung von $r_f(x,y)$ nach y_j und erhalten

$$\begin{aligned}\partial_{y_j} r_f(x,y) = &\sum_{k=0}^{\lfloor d/2 \rfloor} 2k \, \beta_k \, y_j \, \|x\|^{2k} \|y\|^{2k-2} \langle x,y \rangle^{d-2k} \\ &+ \sum_{k=0}^{\lfloor d/2 \rfloor} (d-2k) \, \beta_k \, x_j \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k-1}.\end{aligned}$$

Analog berechnen wir $\partial_{x_i}\partial_{y_j} r_f(x,y)$. Für $\partial_{x_i}\partial_{y_j} r_f(q,q)$ erhalten wir:

$$\partial_{x_i}\partial_{y_i} r_f(q,q) = \sum_{k=0}^{\lfloor d/2 \rfloor} (d-2k)\,\beta_k\,\delta_{ij}.$$

Daher folgt: $\mathbb{E}\big[\big(\partial_{x_i} f(q)\big)^2\big] = \sum_{k=0}^{\lfloor d/2 \rfloor}(d-2k)\,\beta_k$.

Die Erwartung $\mathbb{E}\big[f(q)^2\big] = r_f(q,q)$ berechnen wir über

$$\mathbb{E}\Big[f(q)^2\Big] = r_f(q,q) = \sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k.$$

Wir haben die Erwartungen $\mathbb{E}\big[\big(\partial_{x_i} f(q)\big)^2\big]$ und $\mathbb{E}\Big[f(q)^2\Big]$ über die Kovarianzfunktion berechnet. Mithilfe von Korollar 5.11 können wir den Parameter $\delta(f)$ von f bestimmen.

Proposition 5.18. *Sei f ein invariantes zufälliges zentriertes Polynom $f \in \mathcal{H}_d^n(\mathbb{R})$, dessen Kovarianzfunktion in der Form (35) gegeben ist. Für den Parameter von f gilt:*

$$\delta(f) = \frac{\sum_{k=0}^{\lfloor d/2 \rfloor}(d-2k)\,\beta_k}{\sum_{k=0}^{\lfloor d/2 \rfloor}\beta_k} \tag{37}$$

Beispiel 5.19. Wir haben gesehen, dass wir durch die Wahl von $\beta_0 = 1$ und $\beta_k = 0$ für $0 < k \leq \lfloor d/2 \rfloor$ in (35) die Kostlan-Verteilung erhalten. Setzen wir diese β_k in (37) ein, so erhalten wir $\delta(f) = d$, wie in Beispiel 5.12.

5.2.3 Klassifikation invarianter zentrierter Normalverteilungen

Wir beobachten, dass nicht jede Kombination der β_k in (34) ein positiv definites Skalarprodukt bzw. eine positiv definite Kovarianzmatrix ergibt. Wir möchten aber diese Skalarprodukte als Kovarianzen verwenden. Daher sind wir an positiv definiten Skalarprodukten interessiert, da nur diese eine $\mathcal{O}(n+1)$-invariante Normalverteilung auf $\mathcal{H}_d^n(\mathbb{R})$ definieren.

Wir werden hier unser Ergebnis aus Abschnitt 3.2 verwenden, um invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ zu klassifizieren: $\mathcal{O}(n+1)$-invariante Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})$ sind durch $\lfloor d/2 \rfloor$-viele positive Zahlen c_k parametrisiert. Dabei entsprachen die c_k Gewichten der Skalarprodukte auf den Räumen $r^{2k}\mathcal{Y}_{d-2k}^n(\mathbb{R})$ (vgl. Gleichung (23)).

Zwar haben wir im vorigen Abschnitt schon eine Klassifikation angeben können, allerdings ist diese unzureichend, da wir nicht direkt sehen können, welche Kombination der β_k in (35) positiv definite Skalarprodukte erzeugt. Deshalb möchten wir die positiven Gewichte c_k, $1 \leq k \leq \lfloor d/2 \rfloor$, benutzen. Dazu folgen wir [10].

Kostlan [10] verwendet dafür ***Gegenbauer Polynome***. Dabei definieren wir die Gegenbauer Polynome $\mathcal{C}_m^\nu$ über ihre erzeugende Funktion

$$(1-2tz+z^2)^{-\nu} = \sum_{m=0}^{\infty} \mathcal{C}_m^{\nu}(t) z^m,$$

$t \in [-1,1], \nu \in \mathbb{R} \setminus \{0\}, |z| < 1$.

Sei im Folgenden $\nu \neq 0$ und $m \in \mathbb{N}$. Die Gegenbauer-Polynome $\mathcal{C}_m^{\nu}$ haben dann bekanntermaßen die Form

$$\mathcal{C}_m^{\nu}(t) = \frac{1}{\Gamma(\nu)} \sum_{j=0}^{\lfloor m/2 \rfloor} (-1)^j \frac{\Gamma(\nu+m-j)}{j!(m-2j)!} (2t)^{m-2j}. \tag{38}$$

Dabei bezeichnet $\Gamma(x)$ die ***Eulersche Gammafunktion***. Für $t=1$ berechnen wir

$$\mathcal{C}_m^{\nu}(1) = \frac{\Gamma(2\nu+m)}{\Gamma(2\nu)\, m!}.$$

Definition 5.20. Speziell definieren wir die normierten Gegenbauer-Polynome

$$\hat{\mathcal{C}}_m^{\nu}(t) \coloneqq \frac{\mathcal{C}_m^{\nu}(t)}{\mathcal{C}_m^{\nu}(1)}$$

für $\nu \neq 0$.

Für $\hat{\mathcal{C}}_m^{\nu}(t)$ erhalten wir die folgende explizite Formel:

$$\begin{aligned}
\hat{\mathcal{C}}_m^{\nu}(t) &= \frac{\mathcal{C}_m^{\nu}(t)}{\mathcal{C}_m^{\nu}(1)} \\
&= \frac{\Gamma(2\nu)\, m!}{\Gamma(2\nu+m)\,\Gamma(\nu)} \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j\,\Gamma(\nu+m-j)}{j!\,(m-2j)!} (2t)^{m-2j} \\
&= \frac{2^m\,\Gamma(2\nu)\, m!}{\Gamma(2\nu+m)\,\Gamma(\nu)} \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j\,\Gamma(\nu+m-j)}{2^{2j}\, j!\,(m-2j)!} t^{m-2j} \\
&= \frac{2^{m-1}\, m!\,\Gamma(2\nu+1)}{\Gamma(2\nu+m)\,\Gamma(\nu+1)} \sum_{j=0}^{\lfloor m/2 \rfloor} \frac{(-1)^j\,\Gamma(\nu+m-j)}{2^{2j}\, j!\,(m-2j)!} t^{m-2j}.
\end{aligned} \tag{39}$$

Dabei haben wir in der letzten Gleichung die Identität

$$\Gamma(\nu) = \frac{\Gamma(\nu+1)}{\nu}$$

verwendet.

Um (positiv definite) Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})^*$ der Form $v(x)^T C v(y)$ (vgl. (33)) zu klassifizieren, betrachten wir zunächst ein orthogonal invariantes Skalarprodukt der Form $v(x)^T C_i v(y)$ auf $\left(r^{2i} \mathcal{Y}_{d-2i}^n(\mathbb{R})\right)^*$ für $x, y \in \mathbb{R}^{n+1}$ und $0 \leq i \leq \lfloor d/2 \rfloor$. Nach Satz 5.14 können wir dieses Skalarprodukt in der folgenden Form schreiben:

$$v(x)^T\, C_i\, v(y) = \sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k\, \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k}$$

für gewisse $\beta_k \in \mathbb{R}$, $0 \le k \le \lfloor d/2 \rfloor$.

Fixieren wir $y \in \mathbb{R}^{n+1}$, so können wir $v(x)^T\, C_i\, v(y)$ als Polynom in $r^{2i}\mathcal{Y}^n_{d-2i}(\mathbb{R})$ auffassen. Wir wenden den linearen Operator $r^2\Delta$ auf $v(x)^T\, C_i\, v(y)$ an und erhalten

$$\begin{aligned} r^2\Delta &\sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k\, \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k} \\ &= \sum_{k=1}^{\lfloor d/2 \rfloor} \Big[2k\,(n+2d-2k-1)\,\beta_k \\ &\qquad + (d-2k+2)(d-2k+1)\,\beta_{k-1} \Big]\, \|x\|^{2k} \|y\|^{2k} \langle x,y \rangle^{d-2k}. \end{aligned} \tag{40}$$

Da außerdem nach Proposition 3.22 der Raum $r^{2i}\mathcal{Y}^n_{d-2i}(\mathbb{R})$ ein Eigenraum des Operators $r^2\Delta$ zum Eigenwert $\lambda_i := 2i\,(n+2d-2i-1)$ ist, gilt

$$r^2\Delta\, v(x)^T\, C_i\, v(y) = \lambda_i\, v(x)^T\, C_i\, v(y).$$

Setzen wir (40) in die obige Gleichung ein, so folgt

$$\big(2k\,(n+2d-2k-1) - \lambda_i\big)\,\beta_k + (d-2k+2)(d-2k+1)\,\beta_{k-1} = 0. \tag{41}$$

Wegen $\lambda_i = 2i\,(n+2d-2i-1)$ folgt daher, dass $\beta_k = 0$ für $0 \le k < i$ gilt. Insbesondere folgt aus (41), dass die β_k bis auf eine Konstante eindeutig bestimmt sind.

Kostlan [10] behauptet, dass die Koeffizienten bestimmter Gegenbauer-Polynome die selben Relationen wie die β_k in Gleichung (41) erfüllen.

Behauptung 5.21 ([10, Cp. 4.4])**.** *Sei $v(x)^T\, C_i\, v(y)$ eine orthogonal invariante symmetrische Bilinearform auf $\big(r^{2i}\mathcal{Y}^n_{d-2i}(\mathbb{R})\big)^*$ und Θ bezeichne den Winkel zwischen x und y. Dann gilt:*

$$v(x)^T\, C_i\, v(y) = c_i\, \|x\|^d\, \|y\|^d\, \hat{\mathcal{C}}^{\frac{n-1}{2}}_{d-2i}(\cos\Theta)\,, \quad c_i \in \mathbb{R}$$

und die Bilinearform ist genau dann positiv definit und daher ein Skalarprodukt, wenn $c_i > 0$ gilt.

Angenommen, wir haben eine zentrierte $\mathcal{O}(n+1)$-invariante Normalverteilung auf $\mathcal{H}^n_d(\mathbb{R})$, mit einem orthogonal invarianten Skalarprodukt $v(x)^T C v(y)$ auf $\mathcal{H}^n_d(\mathbb{R})^*$ assoziiert (vgl. (33)). Dann gilt: $v(x)^T C v(y)$ ist eine Linearkombination von orthogonal invarianten Skalarprodukten $v(x)^T\, C_i\, v(y)$ auf $\big(r^{2i}\mathcal{Y}^n_{d-2i}(\mathbb{R})\big)^*$.

Satz 5.22 ([10, Cp. 4.5]). *Sei $v(x)^T Cv(y)$ ein orthogonal invariantes Skalarprodukt auf $\mathcal{H}_d^n(\mathbb{R})^*$ und Θ der Winkel zwischen x und y, $x, y \in \mathbb{R}^{n+1}$. Dann gilt:*

$$v(x)^T Cv(y) = \|x\|^d \|y\|^d \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \, \hat{\mathcal{C}}_{d-2k}^{\frac{n-1}{2}} (\cos\Theta), \quad c_k > 0. \tag{42}$$

Satz 5.22 parametrisiert damit alle positiv definiten Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})^*$ durch $\lfloor d/2 \rfloor$-viele positive Zahlen. Da wir positiv definite Skalarprodukte auf $\mathcal{H}_d^n(\mathbb{R})^*$ mit orthogonal invarianten zentrierten Normalverteilungen auf $\mathcal{H}_d^n(\mathbb{R})$ identifizieren können, folgt, dass invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ durch $\lfloor d/2 \rfloor$-viele positive Zahlen parametrisiert sind.

Seien die c_k im Weiteren positiv.

Korollar 5.23. *Sei Θ der Winkel zwischen x und y, $x, y \in \mathbb{R}^{n+1}$ und $f \in \mathcal{H}_d^n(\mathbb{R})$ ein invariantes zufälliges zentriertes Polynom. Für die Kovarianzfunktion $r_f(x, y)$ von f gilt:*

$$r_f(x, y) = \mathbb{E}\Big[f(x)f(y)\Big] = \|x\|^d \|y\|^d \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \, \hat{\mathcal{C}}_{d-2k}^{\frac{n-1}{2}} (\cos\Theta). \tag{43}$$

Bemerkung 5.24. Die c_k aus Satz 5.22 entsprechen den positiven Gewichten aus (23).

Sei $x = y \neq 0$. Vergleichen wir Gleichung (42) mit (34), so erhalten wir mit $\hat{\mathcal{C}}_{d-2k}^{\frac{n-1}{2}}(1) = 1$:

$$\sum_{k=0}^{\lfloor d/2 \rfloor} \beta_k = v(x)^T Cv(x) = \sum_{k=0}^{\lfloor d/2 \rfloor} c_k.$$

Bemerkung 5.25 ([10, Cp. 4.5]). Wir können die Parameter β_k aus (34) in Abhängigkeit der Parameter c_k aus (42) angeben. Sei $0 \leq k, i \leq \lfloor d/2 \rfloor$. Es gilt:

(i) $$\beta_k = \frac{2^{d-2k-1}\,(n-1)!}{\Gamma\left(\frac{n+1}{2}\right)\,(d-2k)!} \sum_{i=0}^{k} \frac{(-1)^{k-i}\,(d-2i)!\,\Gamma\left(\frac{n-1}{2} + d - i - k\right)}{\Gamma\,(n+d-2i-1)\;(k-i)!}\, c_i.$$

(ii) $$c_i = \frac{\Gamma\left(\frac{n+1}{2}\right)\,\left(\frac{n-1}{2} + d - 2i\right)\Gamma\,(n-1+d-2i)}{(n-1)!\,(d-2i)!} \cdot \sum_{k=0}^{i} \frac{(d-2k)!}{2^{d-2k-1}\,(i-k)!\,\Gamma\left(\frac{n+1}{2} + d - k - i\right)}\, \beta_k.$$

Analog zum vorigen Abschnitt möchten wir den Parameter in Abhängigkeit der positiven Gewichte c_k angeben, die invariante zufällige zentrierte Polynome aus $\mathcal{H}_d^n(\mathbb{R})$ parametrisieren.

Sei $f \in \mathcal{H}_d^n(\mathbb{R})$ ein orthogonal invariantes zufälliges zentriertes Polynom, dessen Kovarianzfunktion r_f in der Form (43) gegeben ist. Wir berechnen die Erwartungen $\mathbb{E}\big[\big(\partial_{x_l} f(q)\big)^2\big]$ und $\mathbb{E}\big[f(q)^2\big] = r_f(q,q)$ für $q = (1,0,\ldots,0) \in \mathbb{S}^n$ und $1 \le l \le n$.

Wir erinnern uns daran, dass Θ den Winkel zwischen $x \in \mathbb{R}^{n+1}$ und $y \in \mathbb{R}^{n+1}$ bezeichnet. Daher können wir $\cos(\Theta)$ auch in der Form $\cos(\Theta) = \frac{\langle x,y \rangle}{\|x\|\|y\|}$ schreiben. Eingesetzt in (43) ergibt dies (vgl. (39))

$$\begin{aligned} r_f(x,y) = \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \frac{2^{d-2k-1}\,(d-2k)!\,\Gamma(n)}{\Gamma(n+d-2k-1)\;\Gamma\left(\frac{n+1}{2}\right)} \\ \cdot \sum_{i=k}^{\lfloor d/2 \rfloor} \frac{(-1)^{i-k}\,\Gamma\left(\frac{n-1}{2}+d-k-i\right)}{2^{2(i-k)}\,(i-k)!\,(d-2i)!} \|x\|^{2i}\,\|y\|^{2i} \langle x,y \rangle^{d-2i}. \end{aligned}$$

Die Erwartung $\mathbb{E}\big[\big(\partial_{x_l} f(q)\big)^2\big]$ berechnen wir über $\mathbb{E}\big[\partial_{x_l} f(q) \partial_{x_m} f(g)\big] = \partial_{x_l}\partial_{y_m} r_f(q,q)$ für $1 \le l,m \le n$. Es gilt:

$$\begin{aligned} \partial_{x_l}\partial_{y_m} r_f(q,q) = \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \frac{2^{d-2k-1}\,(d-2k)!\,\Gamma(n)}{\Gamma(n+d-2k-1)\;\Gamma\left(\frac{n+1}{2}\right)} \\ \cdot \sum_{i=k}^{\lfloor d/2 \rfloor} \frac{(-1)^{i-k}\,\Gamma\left(\frac{n-1}{2}+d-k-i\right)}{2^{2(i-k)}\,(i-k)!\,(d-2i-1)!} \delta_{lm}. \end{aligned}$$

Die obige Gleichung können wir vereinfachen. Wir erhalten:

$$\partial_{x_l}\partial_{y_m} r_f(q,q) = \sum_{k=0}^{\lfloor d/2 \rfloor} c_k \frac{(d-2k)\,(n+d-2k-1)}{n} \delta_{lm}.$$

Für die Erwartung $\mathbb{E}\big[f(q)^2\big]$ gilt:

$$\mathbb{E}\Big[f(q)^2\Big] = r_f(q,q) = \sum_{k=0}^{\lfloor d/2 \rfloor} c_k\,\hat{\mathcal{C}}_{d-2k}^{\frac{n-1}{2}}(1) = \sum_{k=0}^{\lfloor d/2 \rfloor} r_k^2,$$

da $\hat{\mathcal{C}}_{d-2k}^{\frac{n-1}{2}}(1) = 1$ nach Konstruktion gilt.

Wir wenden Korollar 5.11 an, um den Parameter $\delta(f)$ eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ in Abhängigkeit der positiven c_k anzugeben.

Proposition 5.26. *Sei die Kovarianzfunktion eines invarianten zufälligen zentrierten Polynoms $f \in \mathcal{H}_d^n(\mathbb{R})$ in der Form (43) gegeben. Für den Parameter $\delta(f)$ von f gilt:*

$$\delta(f) = \frac{\sum_{k=0}^{\lfloor d/2 \rfloor} c_k\,(d-2k)(n+d-2k-1)}{n \sum_{k=0}^{\lfloor d/2 \rfloor} c_k}.$$

6 Die erwartete Anzahl reeller Nullstellen von Systemen invarianter zufälliger Polynome

In diesem Kapitel untersuchen wir die erwartete Anzahl reeller Nullstellen von Systemen $f_1, \ldots, f_n$ von invarianten zufälligen Polynomen $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$ für $i = 1, \ldots, n$. Wir werden zeigen, dass die erwartete Anzahl reeller Nullstellen von Systemen $f_1, \ldots, f_n$ paarweise unabhängiger invarianter zufälliger zentrierter Polynome $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$ für $i = 1, \ldots, n$, nur vom Parameter $\delta(f_i)$ (vgl. Definition 5.6) der einzelnen Polynome abhängt (vgl. Theorem 6.2).

Wir beschränken uns zunächst auf zentrierte Polynome. Am Ende von Abschnitt 6.3 werden wir sehen, dass es leicht ist, den nicht-zentrierten Fall auf den zentrierten Fall zu reduzieren: Die erwartete Anzahl reeller Nullstellen von nicht-zentrierten Polynomen unterscheidet sich nur um einen Faktor vom zentrierten Fall.

Bemerkung 6.1. Wir beschränken uns auf homogene Polynome und bemerken, dass wir von homogenen Polynomen zu inhomogenen Polynomen kommen, indem wir eine Variable durch eine Konstante ersetzen. Dies führt unter Umständen zu einer Skalierung der Koeffizienten, hat jedoch keinen Einfluss auf die Verteilung der Nullstellen.

Seien $f_1, \ldots, f_n$ homogene Polynome vom Grad $d_1, \ldots, d_n$, $d_i \geq 1$, in $n+1$ Variablen. Das System $f = (f_1, \ldots, f_n)$ definiert dann eine differenzierbare Abbildung $f : \mathbb{S}^n \to \mathbb{R}^n$. Für $x \in \mathbb{S}^n$ bezeichen wir mit $\mathrm{D}f(x) : \mathrm{T}_x\,\mathbb{S}^n \to \mathbb{R}^n$ die Ableitung von f an der Stelle x. Speziell gilt:

$$\mathrm{D}f(q) = [\mathrm{D}f_1(q), \ldots, \mathrm{D}f_n(q)] \tag{44}$$

für $q = (1, 0, \ldots, 0) \in \mathbb{S}^n$. Anders gesagt bedeutet dies, dass die Vektoren $\mathrm{D}f_1(q), \ldots, \mathrm{D}f_n(q)$ mit $\mathrm{D}f_i(q) = (\partial_1 f_i(q), \ldots, \partial_n f_i(q))$ die Spalten der Matrix $\mathrm{D}f(q)$ bilden (vgl. Abschnitt 5.1).

Betrachten wir die Nullstellen $x \in \mathbb{S}^n$ des Systems $f(x) = 0$, d.h. die gemeinsamen Nullstellen von $f_1(x) = \ldots = f_n(x) = 0$, so ist die Anzahl der Nullstellen im Allgemeinen endlich (n Gleichungen in $n+1$ Variablen im n-dimensionalen reellen projektiven Raum).

Wir bezeichen mit $N^f_{\mathbb{P}^n}(V)$ die Anzahl von Nullstellen des Systems

$$f_1(x) = \ldots = f_n(x) = 0,$$

die in der Teilmenge $V \subseteq \mathbb{P}^n$ liegen. Dabei bezeichnet $\mathbb{P}^n$ den projektiven Raum des $\mathbb{R}^n$. Dabei können wir $\mathbb{P}^n$ mit $\mathbb{S}^n/_\sim$ (modulo der Äquivalenzrelation, die antipodale Punkte miteinander identifiziert) repräsentieren. Ist $V = \mathbb{P}^n$, so definieren wir $N^f_{\mathbb{P}^n} := N^f_{\mathbb{P}^n}(\mathbb{P}^n)$.

Da das System f aus homogenen Polynomen besteht, sind die Nullstellen Geraden im $\mathbb{R}^{n+1}$, die durch den Nullpunkt gehen, also Punkte im $\mathbb{P}^n$. Diese entsprechen gerade den Punkten $(\xi, -\xi)$ auf der Einheitsphäre $\mathbb{S}^n$. Daraus erhalten wir für die Einschränkung auf die Einheitssphäre $\mathbb{S}^n$:

$$N^f_{\mathbb{P}^n}(V) = \frac{N^f_{\mathbb{S}^n}(V|_{\mathbb{S}^n})}{2}.$$

Insbesondere gilt:

$$N^f_{\mathbb{P}^n} = \frac{N^f_{\mathbb{S}^n}}{2}.$$

6.1 Haupttheorem

In diesem Abschnitt möchten wir die erwartete Anzahl reeller Nullstellen $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big]$ eines Systems $f = (f_1, \ldots, f_n)$ von paarweise (stochastisch) unabhängigen invarianten zufälligen zentrierten Polynomen $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$, bestimmen.

Sei $f = (f_1, \ldots, f_n)$ ein System von paarweise unabhängigen invarianten zufälligen zentrierten Polynomen f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$. Der Vektor $f(x) \in \mathbb{R}^n$ ist dann für alle $x \in \mathbb{S}^n$ normalverteilt und zentriert. Wir bezeichnen mit $\rho_{f(x)} : \mathbb{R}^n \to \mathbb{R}$ die Dichtefunktion des zufälligen Vektors $f(x) \in \mathbb{R}^n$.

Schreiben wir das Polynom f_i in der Form $f_i = \sum_{|\alpha|=d} f^{(i)}_\alpha X^\alpha$, so impliziert die paarweise Unabhängigkeit der f_i, dass die Koeffizienten $f^{(i)}_\alpha$ von f_i für verschiedene i unabhängig sind. Die $f^{(i)}_\alpha$ können aber durchaus bei festem i für verschiedene α korreliert sein.

Die erwartete Anzahl reeller Nullstellen $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big]$ des System $f = 0$ ergibt sich dann als die Quadratwurzel des Produkts der Parameter $\delta(f_i)$.

Theorem 6.2 (Haupttheorem)**.** *Sei $f = (f_1, \ldots, f_n)$ ein System von paarweise (stochastisch) unabhängigen invarianten zufälligen zentrierten Polynomen f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$. Dann gilt:*

$$\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \sqrt{\delta(f_1) \cdots \delta(f_n)}.$$

Dabei bezeichnet $\delta(f_i)$ den Parameter des invarianten zufälligen zentrierten Polynoms $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$.

Einen Beweis geben wir Abschnitt 6.3. Der Beweis beruht größtenteils auf der ***Rice-Formel***, die wir im nächsten Abschnitt vorstellen werden.

Am Haupttheorem erkennen wir auch, warum wir die Quadratwurzel des Parameters eines invarianten zufälligen zentrierten Polynoms aus $\mathcal{H}^n_d(\mathbb{R})$ als der erwartete Grad dieses Polynoms interpretieren können: Wir schreiben das Ergebnis von Theorem 6.2 in der Form $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big] = \sqrt{\delta(f_1)} \cdots \sqrt{\delta(f_n)}$. Vergleichen wir mit der Aussage vom Satz von Bézout, der besagt, dass die Anzahl komplexer Nullstellen n komplexer homogener Polynome h_i aus $\mathcal{H}^n_{d_i}(\mathbb{C})$ für $i = 1, \ldots, n$ im Allgemeinen das Produkt der Grade $d_1 \cdots d_n$ ist, so können wir $\sqrt{\delta(f_1)} \cdots \sqrt{\delta(f_n)}$ als das Produkt der erwarteten Grade der invarianten zufälligen zentrierten Polynome $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$ verstehen.

6.2 Die Rice-Formel

Die Rice-Formel wird in Azaïs & Wschebor [11] vorgestellt.

Satz 6.3 (Rice-Formel, [11, Ch. 2, Theorem 1]). *Sei V eine kompakte Teilmenge des $\mathbb{R}^n$, $Z : V \to \mathbb{R}^n$ ein zufälliges Feld, d.h. $Z(x) = \big(Z_1(x), \ldots, Z_n(x)\big)^T$, $x \in V$, mit zufälligem $Z_i : V \to \mathbb{R}$ und $u \in \mathbb{R}^n$ fest. Angenommen es gelten die folgenden Aussagen:*

(i) *Z ist normalverteilt.*

(ii) *$x \mapsto Z(x)$ ist fast sicher stetig differenzierbar.*

(iii) *Für jedes $x \in V$ hat $Z(x)$ eine nicht-degenerierte Verteilung mit Dichte $\rho_{Z(x)}$.*

(iv) $\operatorname{Prob}\Big\{\exists x \in \mathring{V} \mid Z(x) = u, \det\big(\mathrm{D}Z(x)\big)\Big\} = 0$, *wobei $\mathring{V}$ das Innere von V bezeichnet.*

(v) *$\lambda_n(\partial V) = 0$, wobei ∂V den Rand von V bezeichnet und λ_n das Lebesgue-Maß auf $\mathbb{R}^n$. Dabei schreiben wir $\mathrm{d}x$ anstelle von $\lambda_n(\mathrm{d}x)$.*

Bezeichnen wir mit $N_u^Z(V) := \#\big\{x \in V \mid Z(x) = u\big\}$ die Anzahl der Nullstellen des Systems $Z(x) = u$, so gilt:

$$\mathbb{E}\Big[N_u^Z(V)\Big] = \int\limits_V \mathbb{E}\Big[\big|\det\big(\mathrm{D}Z(x)\big)\big|\Big|Z(x) = u\Big]\,\rho_{Z(x)}(u)\,\mathrm{d}x$$

und beide Seiten sind endlich.

Wir möchten einen kurzen Beweis angeben, der in [11, Cp. 2] vorgestellt wird.

Beweis von Satz 6.3. Angenommen u sei kein kritischer Punkt von Z, d.h. es gilt $Z(x) = u$ und $\mathrm{D}Z(x) \neq 0$ (dies ist richtig mit Wahrscheinlichkeit 1, unter den Voraussetzungen des Satzes). Wir setzen $m := N_u^Z(V)$. Da V kompakt ist, ist m endlich. Seien $x_1, \ldots, x_m \in V$ die Nullstellen von $Z(x) = u$, für $m \geq 1$. Man kann zeigen, dass fast sicher $x_i \notin \partial V$ gilt, $i = 1, \ldots, m$. Wir können den ***Satz von der Umkehrabbildung*** anwenden. Ist $\delta > 0$ klein genug, so können wir offene Umgebungen $U_1, \ldots, U_m$ von $x_1, \ldots, x_m$ in V finden, sodass gilt:

(i) Z ist ein $\mathcal{C}^1$-Diffeomorphismus von U_i nach $B_n(u, \delta)$, für alle $i = 1, \ldots, m$, wobei $B_n(u, \delta)$ die offene Kugel im $\mathbb{R}^n$ mit Zentrum u und Radius δ bezeichnet.

(ii) $U_1, \ldots, U_m$ sind paarweise disjunkt.

(iii) Ist $x \notin \bigcup_{i=1}^m U_i$, dann folgt $Z(x) \notin B_n(u, \delta)$.

Aus den obigen Eigenschaften folgt zunächst, dass

$$\int_V \left|\det\left(\mathrm{D}Z(x)\right)\right| \mathbb{1}_{\{\|Z(x)-u\|<\delta\}}\,\mathrm{d}x = \sum_{i=1}^{m} \int_{U_i} \left|\det\left(\mathrm{D}Z(x)\right)\right| \mathrm{d}x.$$

gilt. Wir wenden den Transformationssatz aus der Analysis an. Es gilt:

$$\sum_{i=1}^{m} \int_{U_i} \left|\det\left(\mathrm{D}Z(x)\right)\right| \mathrm{d}x = \lambda_n\big(B_n(u,\delta)\big)\, m.$$

Für die Anzahl reeller Nullstellen von $Z(x) - u$, die in V liegen, erhalten wir den folgenden Ausdruck:

$$N_u^Z(V) = m = \lim_{\delta \downarrow 0} \frac{1}{\lambda_n\big(B_n(u,\delta)\big)} \int_V \left|\det\left(\mathrm{D}Z(x)\right)\right| \mathbb{1}_{\{\|Z(x)-u\|<\delta\}}\,\mathrm{d}x.$$

Wir verwenden den Satz von der majorisierten Konvergenz um $\lim_{\delta\downarrow 0}$ vor das Integral verschieben zu können. Eine Rechnung ergibt dann die erwartete Anzahl reeller Nullstellen $\mathbb{E}\big[N_u^Z(V)\big]$ von $Z(x) - u$, die in V liegen. Es gilt:

$$\begin{aligned}
&\mathbb{E}\Big[N_u^Z(V)\Big] \\
&= \lim_{\delta\downarrow 0} \int_V \frac{1}{\lambda_n\big(B_n(u,\delta)\big)} \int_{B_n(u,\delta)} \mathbb{E}\Big[\left|\det\left(\mathrm{D}Z(x)\right)\right| \Big| Z(x) = y\Big] \rho_{Z(x)}(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_V \mathbb{E}\Big[\left|\det\left(\mathrm{D}Z(x)\right)\right| \Big| Z(x) = u\Big] \rho_{Z(x)}(u)\,\mathrm{d}x.
\end{aligned}$$

□

Für normalverteilte Vektoren impliziert Unkorreliertheit auch Unabhängigkeit. Diese Eigenschaft werden wir für den Beweis des Haupttheorems ausnutzen, um die bedingte Erwartung in der Rice-Formel zu vereinfachen.

6.3 Beweis des Haupttheorems

In diesem Abschnitt möchten wir das Haupttheorem (Theorem 6.2) beweisen. Sei daher $f = (f_1, \ldots, f_n)$ ein System von paarweise (stochastisch) unabhängigen invarianten zufälligen zentrierten Polynomen f_i aus $\mathcal{H}_{d_i}^n(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$.

Sei $q = (1, 0, \ldots, 0) \in \mathbb{S}^n$. Wir können ohne Beschränkung der Allgemeinheit $\mathbb{E}\big[f_i(q)^2\big] = 1$ annehmen, $i = 1, \ldots, n$. Andernfalls betrachten wir das System $\hat{f} := (\hat{f}_1, \ldots, \hat{f}_n)$ mit

$$\hat{f}_i(x) = \frac{f_i(x)}{\sqrt{\mathbb{E}\Big[f_i(q)^2\Big]}}, \quad i = 1, \ldots, n.$$

Dann gilt $\mathbb{E}\big[\hat{f}_i(q)^2\big] = 1$ und für den Parameter $\delta(\hat{f}_i)$ von $\hat{f}_i$ erhalten wir $\delta(\hat{f}_i) = \delta(f_i)$ für alle $i = 1, \dots, n$, da Skalierungen den Parameter nicht verändern. Außerdem ändert sich die erwartete Anzahl an Nullstellen beim Übergang zum System $\hat{f}$ ***nicht***, da aus $f(x) = 0 \iff \hat{f}(x) = 0$ folgt, dass $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big] = \mathbb{E}\big[N^{\hat{f}}_{\mathbb{P}^n}\big]$ gilt.

Also seien die $f_i(q)$ standard-normalverteilt und paarweise unabhängig für alle $i = 1, \dots, n$. Für die Dichtefunktion $\rho_{f(q)}(x)$ von $f(q)$ gilt:

$$\rho_{f(q)}(x) = (2\pi)^{-n/2} \exp\left(-\frac{1}{2}(x_1^2 + \cdots + x_n^2)\right), \quad x \in \mathbb{R}^n. \tag{45}$$

Insbesondere gilt $\rho_{f(q)}(0) = (2\pi)^{-n/2}$.

Beweis des Haupttheorems. Wir nehmen nach der obigen Erklärung $\mathbb{E}\big[f_i(q)^2\big] = 1$ an, $i = 1, \dots, n$. Um $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big]$ zu bestimmen, verwenden wir die Rice-Formel (Satz 6.3). Durch Einsetzen in die Rice-Formel erhalten wir:

$$\begin{aligned}
\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] &\underset{(6)}{=} \frac{1}{2}\,\mathbb{E}\Big[N^f_{\mathbb{S}^n}\Big] \\
&= \frac{1}{2}\int\limits_{\mathbb{S}^n} \mathbb{E}\Big[\big|\det\big(\mathrm{D}f(x)\big)\big| \,\Big|\, f(x) = 0\Big]\,\rho_{f(x)}(0)\,\mathrm{d}\mathbb{S}^n(x).
\end{aligned}$$

Aufgrund der $\mathcal{O}(n+1)$-Invarianz der f_i reduziert sich die obige Gleichung auf

$$\begin{aligned}
\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] &= \frac{1}{2}\,\mathbb{E}\Big[N^f_{\mathbb{S}^n}\Big] \\
&= \frac{1}{2}\int\limits_{\mathbb{S}^n} \mathbb{E}\Big[\big|\det\big(\mathrm{D}f(q)\big)\big| \,\Big|\, f(q) = 0\Big]\,\rho_{f(q)}(0)\,\mathrm{d}\mathbb{S}^n(x) \\
&= \frac{1}{2}\,(2\pi)^{-n/2}\int\limits_{\mathbb{S}^n} \mathbb{E}\Big[\big|\det\big(\mathrm{D}f(q)\big)\big| \,\Big|\, f(q) = 0\Big]\,\mathrm{d}\mathbb{S}^n(x).
\end{aligned} \tag{46}$$

Dabei ist $q = (1, 0, \dots, 0) \in \mathbb{S}^n$ und $\mathrm{D}f(q) = [\mathrm{D}f_1(q), \dots, \mathrm{D}f_n(q)]$ bezeichne die Matrix mit den Spalten $\mathrm{D}f_1(q), \dots, \mathrm{D}f_n(q)$ (vgl. (44)).

Nach Korollar 5.9 ist $f_i(q)$ unabhängig von $\mathrm{D}f_i(q)$ für alle $i = 1, \dots, n$. Außerdem sind die f_i paarweise unabhängig. Daher folgt, dass $f(q)$ und $\mathrm{D}f(q)$ unabhängig sind. Damit vereinfacht sich die bedingte Erwartung zu

$$\mathbb{E}\Big[\big|\det\big(\mathrm{D}f(q)\big)\big| \,\Big|\, f(q) = 0\Big] = \mathbb{E}\Big[\big|\det\big(\mathrm{D}f(q)\big)\big|\Big].$$

Eingesetzt in Gleichung (46) ergibt dies

$$\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \frac{1}{2}\,\mathbb{E}\Big[N^f_{\mathbb{S}^n}\Big] = \frac{1}{2}\,(2\pi)^{-n/2}\int\limits_{\mathbb{S}^n} \mathbb{E}\Big[\big|\det\big(\mathrm{D}f(q)\big)\big|\Big]\mathrm{d}\mathbb{S}^n(x). \tag{47}$$

Als nächstes werden wir $\mathbb{E}\big[\big|\det\big(\mathrm{D}f(q)\big)\big|\big]$ berechnen. Für die Kovarianzmatrix von $\mathrm{D}f_i(q)$ gilt nach Proposition 5.10

$$\operatorname{Cov}\big(\mathrm{D}f_i(q)\big) = \delta(f_i)\, I_n.$$

Wir setzen

$$T_i := \frac{\mathrm{D}f_i(q)}{\sqrt{\delta(f_i)}}, \quad i = 1, \ldots, n$$

und T bezeichne die zufällige Matrix mit den Spalten T_i, d.h. es gilt

$$T := [T_1, \ldots, T_n].$$

Dann gilt $\operatorname{Cov}\big(T_i\big) = I_n$ für alle $i = 1, \ldots, n$. Mit anderen Worten bedeutet dies, dass die $T_1, \ldots, T_n$ standard-normalverteilt und unabhängig sind.

Für die Determinante $\big|\det\big(\mathrm{D}f(q)\big)\big|$ von $\mathrm{D}f(q)$ erhalten wir

$$\big|\det\big(\mathrm{D}f(q)\big)\big| = \big|\det\big(T\big)\big| \prod_{i=1}^{n} \sqrt{\delta(f_i)}.$$

Die Matrix $T \in \mathbb{R}^{n\times n}$ ist eine zufällige standard-normalverteilte Matrix und nach Lemma 4.7 gilt $\big|\det\big(T\big)\big| = L_n$. Folglich gilt:

$$\big|\det\big(\mathrm{D}f(q)\big)\big| = \big|\det\big(T\big)\big| \prod_{i=1}^{n} \sqrt{\delta(f_i)} = L_n \sqrt{\delta(f_1)\cdots\delta(f_n)}.$$

Eingesetzt in (47) ergibt dies

$$\begin{aligned}\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \frac{1}{2}\,\mathbb{E}\Big[N^f_{\mathbb{S}^n}\Big] &= \frac{1}{2}\,(2\pi)^{-n/2}\, L_n \sqrt{\delta(f_1)\cdots\delta(f_n)} \int\limits_{\mathbb{S}^n} \mathrm{d}\mathbb{S}^n(x) \\ &= \frac{1}{2}\,(2\pi)^{-n/2}\, L_n \sqrt{\delta(f_1)\cdots\delta(f_n)}\, \mathrm{Vol}_{\mathbb{S}^n}(\mathbb{S}^n).\end{aligned}$$

Für das Oberflächenvolumen $\sigma_n := \mathrm{Vol}_{\mathbb{S}^n}(\mathbb{S}^n)$ der Einheitssphäre $\mathbb{S}^n$ im $\mathbb{R}^{n+1}$ gilt die wohlbekannte Formel (vgl. [8, Cp. 1])

$$\sigma_n = 2\,\frac{\pi^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)}. \tag{48}$$

Wir schreiben L_n in der Form (26) und $\mathrm{Vol}_{\mathbb{S}^n}(\mathbb{S}^n)$ in der Form (48). Eingesetzt in die obige Formel ergibt dies

$$\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \frac{1}{2}\,\mathbb{E}\Big[N^f_{\mathbb{S}^n}\Big] = \sqrt{\delta(f_1)\cdots\delta(f_n)}.$$

□

Das Haupttheorem (Theorem 6.2) zeigt, dass die erwartete Anzahl reeller Nullstellen von Systemen $f = (f_1, \ldots, f_n)$ paarweise unabhängiger invarianter zufälliger zentrierter Polynome f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$ die Quadratwurzel aus dem Produkt der Parameter $\delta(f_i)$ der Polynome f_i ist.

Wir wollen jetzt den Fall untersuchen, in dem die Polynome f_i nicht notwendigerweise zentriert sind. Sei daher $f = (f_1, \ldots, f_n)$ im Folgenden ein System von paarweise (stochastisch) unabhängigen invarianten zufälligen und nicht notwendigerweise zentrierten Polynomen f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$.

Bemerkung 6.4. Aus der Definition 5.1 erhalten wir direkt, dass ein invariantes zufälliges Polynom aus $\mathcal{H}^n_d(\mathbb{R})$ zentriert ist, falls d ungerade ist.

Bezeichne $\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{R}^n$ den Erwartungsvektor von $f(q)$ und $\sigma \in \mathbb{R}^{n\times n}$ die Kovarianzmatrix von $f(q)$, $q = (1, 0, \ldots, 0) \in \mathbb{S}^n$. Dann gilt: μ_i ist die Erwartung des Koeffizienten vom Monom $X_0^{d_i}$ von f_i. Außerdem gilt aufgrund der paarweisen Unabhängigkeit der f_i, dass die Kovarianzmatrix σ eine Diagonalmatrix mit den Varianzen σ_i des Koeffizienten vom Monom $X_0^{d_i}$ der f_i auf der Diagonalen ist. D.h. es gilt $\sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. Für die Dichtefunktion $\rho_{f(q)}$ gilt dann (vgl. (24)):

$$\begin{aligned}\rho_{f(q)}(x) &= \frac{1}{\sqrt{(2\pi)^n \det \sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \sigma^{-1}(x-\mu)\right) \\ &= (2\pi)^{-n/2}\,(\sigma_1 \cdots \sigma_n)^{-1/2} \exp\left(-\frac{1}{2}(x-\mu)^T \sigma^{-1}(x-\mu)\right).\end{aligned}$$

Insbesondere erhalten wir für die Dichtefunktion $\rho_{f(q)}(x)$ an der Stelle $x = 0$:

$$\rho_{f(q)}(0) = (2\pi)^{-n/2}\,(\sigma_1 \cdots \sigma_n)^{-1/2} \exp\left(-\frac{1}{2}\left(\frac{\mu_1^2}{\sigma_1} + \cdots + \frac{\mu_n^2}{\sigma_n}\right)\right). \tag{49}$$

Wir bemerken, dass der Parameter $\delta(f)$ (vgl. Definition 5.6) eines invarianten zufälligen Polynoms $f \in \mathcal{H}^n_d(\mathbb{R})$ nicht von der Erwartung des Polynoms abhängt.

Wollen wir das Hauptheorem für das System f von paarweise unabhängigen orthogonal invarianten zufälligen und nicht notwendigerweise zentrierten Polynomen beweisen, so brauchen wir daher nur eine Stelle im Beweis ändern, nämlich die Dichtefunktion $\rho_{f(q)}$ in Gleichung (46), denn $f(q)$ muss nicht notwendigerweise standard-normalverteilt sein. Mit anderen Worten beudeutet dies, dass wir die Dichtefunktion in (46) durch die Dichtefunktion in (49) ersetzen müssen (alles andere kann unverändert bleiben).

Vergleichen wir (45) mit (49) in der Gleichung (46), so sehen wir, dass wir die erwartete Anzahl reeller Nullstellen aus dem Haupttheorem (Theorem 6.2) nur mit dem Faktor $(\sigma_1 \cdots \sigma_n)^{-1/2} \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{\mu_i^2}{\sigma_i}\right)$ skalieren zu brauchen, um die erwartete Anzahl reeller Nullstellen von Systemen $f = (f_1, \ldots, f_n)$ paarweise unabhängiger invarianter zufälliger Polynome f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$ zu erhalten.

Korollar 6.5. *Sei $f = (f_1, \ldots, f_n)$ ein System von paarweise (stochastisch) unabhängigen invarianten zufälligen Polynomen f_i aus $\mathcal{H}^n_{d_i}(\mathbb{R})$, $d_i \geq 1$, $i = 1, \ldots, n$. Weiter bezeichne μ_i die Erwartung und σ_i die Varianz des Koeffizienten des Monoms $X_0^{d_i}$ vom Polynom f_i. Dann gilt:*

$$\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \sqrt{\delta(f_1)\cdots\delta(f_n)}\,(\sigma_1\cdots\sigma_n)^{-1/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\frac{\mu_i^2}{\sigma_i}\right).$$

6.4 Das Ergebnis von Shub & Smale

Zum Abschluss möchten wir das Resultat von Shub & Smale mit den Methoden aus dieser Arbeit erarbeiten.

Shub & Smale [13] haben schon 1993 gezeigt, dass die erwartete Anzahl reeller Nullstellen eines Systems $f_1(x) = 0, \ldots, f_n(x) = 0$ von paarweise unabhängigen Kostlan-verteilten homogenen Polynomen $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$ mit $d_i \geq 1$ für alle $i = 1, \ldots, n$ gleich $\sqrt{d_1 \cdots d_n}$ ist, d.h. die Quadratwurzel aus dem Produkt der Grade der Polynome $f_1, \ldots, f_n$. Der Beweis dieses Resultats erfolgte bei ihnen jedoch nicht über die Rice-Formel.

Wir erinnern nochmal daran, dass wir ein Polynom $f \in \mathcal{H}^n_d(\mathbb{R})$ ***Kostlan-verteilt*** nennen, falls die Koeffizienten $f_\alpha \in \mathbb{R}$ von $f = \sum_{|\alpha|=d} f_\alpha X^\alpha$ normalverteilt, zentriert und paarweise unabhängig sind mit Varianz $\mathrm{Var}(f_\alpha) = \binom{d}{\alpha}$.

Sei $f_1(x) = 0, \ldots, f_n(x) = 0$ ein System von paarweise unabhängigen Kostlan-verteilten homogenen Polynomen $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$ mit $d_i \geq 1$ für alle $i = 1, \ldots, n$. Schreiben wir $f_i = \sum_{|\alpha|=d} f^{(i)}_\alpha X^\alpha$ mit $f^{(i)}_\alpha \in \mathbb{R}$, so folgt aus der paarweisen Unabhängigkeit der f_i, dass die $f^{(i)}_\alpha$ paarweise unabhängig sind für alle $i = 1, \ldots, n$ und alle α mit $|\alpha| = d$.

Korollar 6.6 (Shub & Smale)**.** *Sei $f = (f_1, \ldots, f_n)$ ein System von paarweise unabhängigen Kostlan-verteilten homogenen Polynomen $f_i \in \mathcal{H}^n_{d_i}(\mathbb{R})$ mit $d_i \geq 1$ für alle $i = 1, \ldots, n$. Für die erwartete Anzahl von reellen Nullstellen $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big]$ vom System $f = 0$ gilt:*

$$\mathbb{E}\Big[N^f_{\mathbb{P}^n}\Big] = \sqrt{d_1 \cdots d_n}.$$

Beweis. Nach Beispiel 5.12 erhalten wir $\delta(f_i) = d_i$ für den Parameter von f_i. Setzen wir dies in das Haupttheorem (Theorem 6.2) ein, so erhalten wir sofort $\mathbb{E}\big[N^f_{\mathbb{P}^n}\big] = \sqrt{d_1 \cdots d_n}$. □

Damit ist das Haupttheorem eine echte Verallgemeinerung des Ergebnisses von Shub & Smale.

Literatur

[1] C. Teleman, Lent, *Representation Theory*, 2005, `https://math.berkeley.edu/~teleman/math/RepThry.pdf`

[2] Gerd Fischer, *Lineare Algebra*, Vieweg+Teubner, 17. Auflage, 2010

[3] James E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer New York, 1972

[4] James E. Humphreys, *Linear Algebraic Groups*, Springer New York, 1975

[5] Hanspeter Kraft, *Geometrische Methoden in der Invariantentheorie*, Vieweg+Teubner Verlag, 2. Auflage, 1985

[6] Mark R. Sepanski, *Compact Lie Groups*, Springer New York, 2007

[7] Peter Bürgisser, Felipe Cucker, *Condition - The Geometry of Numerical Algorithms*, Springer-Verlag Berlin Heidelberg, 2013

[8] Kendall Atkinson, Weimin Han, *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, Springer Berlin Heidelberg, 2012

[9] Peter Bürgisser, *Average volume, curvatures, and Euler characteristics of random real algebraic varieties*, arXiv:math/0606755v2, 2013

[10] Eric Kostlan, *On the Expected Number of Real Roots of a System of Random Polynomial Equations*, Foundations of Computational Mathematics: Proceedings of the Smalefest 2000, pp 149–188, 2001

[11] Jean-Marc Azaïs, Mario Wschebor, *On the roots of a random system of equations. The theorem of Shub & Smale and some extensions*, Foundations of Computational Mathematics, Vol. 5, No. 2, pp 125–144, 2005

[12] Michael Spivak, *Comprehensive Introduction to Differential Geometry*, Vol. 5, Publish or Perish Berkley, 2. Auflage, 1979

[13] M. Shub and S. Smale, *Complexity of Bezout's Theorem II Volumes and Probabilities*, Computational Algebraic Geometry, pp 267–285, Birkhäuser Boston, 1993

[14] M. Kac, *On the average number of real roots of a random algebraic equation*, Bull. Amer. Math. Soc., Vol. 49, No. 4, pp 314–320, 1943, `http://projecteuclid.org/euclid.bams/1183505112`

[15] S. S. Podkorytov, *The Euler Characteristic of a Random Algebraic Hypersurfacee*, Journal of Mathematical Sciences, Vol. 104, No. 4, pp 1394–1398, 2001

[16] Andreas Gathmann, *Algebraic Geometry*, 2014, http://www.mathematik.uni-kl.de/~gathmann/class/alggeom-2014/main.pdf

[17] Siegfried Bosch, *Algebra*, Springer Berlin Heidelberg, 8. Auflage, 2013